Estudio de los efectos de mutaciones naturales y modificaciones postraduccionales sobre la estabilidad y funcionalidad de la enzima asociada a cáncer NQO1

Programa de Doctorado en Química

Juan Luis Pacheco García

Departamento de Química Física

Facultad de Ciencias

Universidad de Granada

Directores: Ángel Luis Pey Rodríguez y Rogelio Jesús Palomino Morales

ÍNDICE

ABREVIACIONES

ADN	Ácido desoxirribonucleico
AGT	Alanina:glioxilato aminotransferasa
ARE	Elemento de respuesta antioxidante
ARN	Ácido ribonucleico
ASA	Área de superficie accesible
CLIPS	Chaperonas ligadas a la síntesis de proteínas
CoQ	Coenzima Q
COSMIC	Catálogo de Mutaciones Somáticas de Cáncer
CTD	Dominio C-terminal
DBS	Sitio de unión del dicumarol
DIC	Dicumarol
FAD	Flavín adenín dinucleótido
FADH$_2$	Flavín adenín dinucleótido reducido
FBS	Sitio de unión del FAD
FMN	Flavín mononucleótido
gnomAD	Base de Datos de Agregación del Genoma
HDX-MS	Intercambio hidrógeno-deuterio acoplado a espectrometría de masas
HSPs	Proteínas de choque térmico
IDPs	Proteínas intrínsecamente desordenadas
IDRs	Regiones intrínsecamente desordenadas
MMC	Mitomicina C
MMI	Interfaz monómero-monómero
NADH	Nicotín adenín dinucleótido
NAD(P)$^+$	Nicotín adenín dinucleótido 2'-fosfato oxidado
NAD(P)H	Nicotín adenín dinucleótido 2'-fosfato
NQO1	NAD(P)H:quinona oxidorreductasa 1
Nrf2	Factor nuclear eritroide similar al factor 2
NTD	Dominio N-terminal
ORFs	Marcos abiertos de lectura
PDB	Protein Data Bank
PTMs	Modificaciones postraduccionales

SNP	Polimorfismo de un solo nucleótido
TCS	Sitio de corte de la termolisina
Tm	Temperatura de desnaturalización
TQ	α-tocoferol quinona
TTR	Transtiretina

RESUMEN

La estabilidad y función de las proteínas puede verse afectada por el efecto de mutaciones que provocan la sustitución de un aminoácido por otro (mutación de cambio de sentido) así como por modificaciones postraduccionales. Cuando ese efecto es heredable y se pierde la estabilidad y función de la proteína, a menudo se desarrollan enfermedades conocidas como enfermedades de pérdida de función. Por ejemplo, las flavoproteínas, que tienen múltiples funciones, están asociadas a distintas enfermedades hereditarias debido a mutaciones que provocan la pérdida de su función. Concretamente, en esta tesis vamos a estudiar el efecto de mutaciones naturales y modificaciones postraduccionales en la proteína NAD(P)H:quinona oxidorreductasa 1 (NQO1), una flavoproteína multifuncional cuya pérdida de función por el polimorfismo P187S se ha asociado con un posible incremento del riesgo de padecer cáncer.

En primer lugar, hemos estudiado distintas variantes naturales que han sido detectadas en estudios de secuenciación masiva en la población y en muestras de cáncer, pero cuyo efecto sobre la estabilidad y función de NQO1 se desconocía. Estas variantes presentaban mutaciones de cambio de sentido en el extremo N-terminal y en el sitio activo de la proteína, y el efecto de esas mutaciones se propagó a regiones alejadas del sitio mutado, provocando en algunos casos una menor estabilidad térmica, una reducción en la afinidad por FAD y una menor actividad catalítica.

En segundo lugar, comparamos como se asemeja la desestabilización causada por una fosforilación a la causada por el polimorfismo asociado a enfermedad P187S, utilizando la mutación fosfomimética S82D. Para ello, estudiamos como ambas mutaciones afectaban a la estabilidad estructural de la proteína en distintas regiones mediante intercambio de hidrógeno-deuterio acoplado a espectrometría de masas (HDX-MS) combinándolo con estudios de proteólisis limitada, ensayos enzimáticos y experimentos en cultivos celulares. Los resultados mostraron que tanto S82D como P187S tienen varios efectos sobre NQO1, como la reducción de la actividad catalítica o la estabilidad intracelular, que se produce por la propagación estructural de su efecto desestabilizante a varios sitios funcionales.

En tercer y último lugar, comparamos las consecuencias de la fosforilación específica de sitio en tres residuos (S40, S82 y T128) que presentan distinta exposición al solvente

siguiendo la metodología utilizada con S82. En este caso, vimos que cada mutación fosfomimética (S40D, S82D y T128D) tuvo un efecto diferente. S82D causó la mayor desestabilización afectando fuertemente a la afinidad por FAD y su actividad enzimática y a la estabilidad conformacional e intracelular. Mientras que S40D afectó principalmente a la estabilidad conformacional e intracelular y T128D no tuvo un efecto apreciable sobre la estabilidad de NQO1, pero sí sobre la actividad enzimática.

ABSTRACT

The stability and function of proteins can be affected by mutations that result in the substitution of one amino acid for another (missense mutation) as well as by post-translational modifications. When this effect is hereditary and leads to the loss of protein stability and function, it often results in diseases known as loss-of-function diseases. For example, flavoproteins, which have multiple functions, are associated with various hereditary diseases due to mutations that lead to the loss of their function. Specifically, in this thesis, we are going to study the effect of natural mutations and post-translational modifications in the NAD(P)H:quinone oxidoreductase 1 (NQO1) protein, a multifunctional flavoprotein, whose loss of function due to the P187S polymorphism has been linked to a potential increase in the risk of cancer.

First, we have studied various natural variants that have been detected in population and cancer sample massive sequencing studies, but whose effect on the stability and function of NQO1 was unknown. These variants had missense mutations at the N-terminal end and in the active site of the protein, and the effect of these mutations propagated to distant regions from the mutated site, causing, in some cases, decreased thermal stability, reduced affinity for FAD, and reduced catalytic activity.

Secondly, we compared how the destabilization caused by phosphorylation resembled that caused by the disease-associated P187S polymorphism, using the phosphomimetic mutation S82D. To do this, we studied how both mutations affected the structural stability of the protein in different regions using hydrogen-deuterium exchange coupled with mass spectrometry (HDX-MS), combined with limited proteolysis studies, enzymatic assays, and cell culture experiments. The results showed that both S82D and P187S had various effects on NQO1, such as reduced catalytic activity or intracellular stability, which occurred due to the structural propagation of their destabilizing effect to various functional sites.

Lastly, we compared the consequences of site-specific phosphorylation at three residues (S40, S82, and T128) with varying solvent exposure following the methodology used with S82. In this case, we observed that each phosphomimetic mutation (S40D, S82D, and T128D) had different effects. S82D caused the most destabilization, significantly affecting FAD affinity, enzymatic activity, and conformational and intracellular

stability. While S40D mainly affected conformational and intracellular stability, and T128D had no appreciable effect on the stability of NQO1 but did impact enzymatic activity.

1. INTRODUCCIÓN

1.1 ESTABILIDAD PROTEICA Y EL ORIGEN DE MUCHAS ENFERMEDADES

La exitosa ejecución de todos los procesos celulares que tienen lugar en nuestro cuerpo depende de la correcta interacción de las proteínas. Las proteínas son moléculas grandes y complejas que desempeñan una enorme variedad de funciones de una manera muy eficiente y altamente regulada. Para ser funcionalmente activa una proteína tiene que adquirir una estructura tridimensional determinada que depende en la mayoría de los casos de su estructura primaria, es decir, de la secuencia de aminoácidos y del entorno celular en el que se encuentra[1].

El paradigma de la relación entre la secuencia de la proteína y su estructura tridimensional fue establecido por Christian Anfinsen tras sus estudios con la ribonucleasa. Anfinsen demostró que las proteínas aisladas en disolución pueden recuperar su conformación nativa después de ser desnaturalizadas, por lo tanto la información necesaria para adquirir esa conformación específica debía estar codificada en la secuencia de aminoácidos[2].

1.1.1 Plegamiento de proteínas

El plegamiento proteico es el proceso por el que una proteína recién sintetizada adquiere su estructura tridimensional característica. Para ser funcional, una proteína tiene que plegarse en esa conformación nativa a partir de la cadena lineal de aminoácidos[3]. No obstante, existen proteínas intrínsecamente desordenadas (IDPs, del inglés *intrinsically disordered proteins*) que son activas sin tener una estructura específica. Las IDPs carecen de estructura secundaria o terciaria estable y existen en un conjunto de conformaciones, adaptando su estructura en función de la molécula a la que se unen[4].

El plegamiento es un proceso muy complejo y todavía no se conoce con exactitud los mecanismos moleculares que definen la estructura nativa de una proteína[5]. Aunque, en los últimos años ha habido grandes avances en el desarrollo de herramientas computacionales para predecir la estructura de proteínas a partir de la secuencia de aminoácidos. Es el caso de AlphaFold2[6] y RoseTTaFold[7], redes neuronales que han

sido entrenadas con miles de estructuras conocidas y permiten predecir con gran exactitud la estructura de una proteína.

Según la teoría del paisaje energético, el plegamiento de una proteína no sigue un único camino, sino que ocurre a través de una ruta de plegamiento en la que la proteína adopta un conjunto de estados intermediarios como si se tratase de un embudo rugoso dirigido hacia el estado de mínima energía que corresponde a la conformación nativa (Figura 1A)[8]. En el paisaje energético encontramos muchos estados de alta energía y pocos estados de baja energía. En los niveles de energía más altos la proteína carece de estructuras ordenadas, pero conforme la proteína se pliega y adquiere las propiedades de la conformación nativa pasa a niveles de energía más bajos. Como resultado de las distintas rutas de plegamiento que puede seguir una proteína y del aspecto rugoso del paisaje energético, una proteína puede quedar atrapada de un estado de baja energía en el que está mal plegada o parcialmente plegada. Los estados mal plegados o parcialmente plegados normalmente exponen aminoácidos hidrofóbicos y regiones del esqueleto polipeptídico no estructuradas al solvente que favorecen la agregación y que puede llevar a la formación de agregados amorfos o fibrillas amiloides (Figura 1A)[5,8,9].

El plegamiento *in vivo* ocurre en condiciones muy diferentes a las estudiadas *in vitro*, y tiene que hacer frente a varios retos[10]. Por una parte, el entorno celular tiene una densidad macromolecular mucho mayor que favorece la agregación proteica al aumentar la afinidad entre las macromoléculas que interactúan, incluyendo los intermediarios de plegamiento[10,11]. Además, el plegamiento *in vivo* puede ocurrir simultáneamente con la traducción en los ribosomas lo que aumenta las restricciones en el plegamiento por la limitación de espacio hasta que la proteína ha salido del ribosoma. Mientras que proteínas pequeñas de un solo dominio no se pliegan hasta que la traducción se ha completado, las proteínas de varios dominios pueden empezar a plegarse por el extremo N-terminal mientras continúa la traducción, pero los residuos del extremo C-terminal no pueden participar en las interacciones de larga distancia necesarias para el plegamiento cooperativo de los dominios[10,12].

Por otra parte, el plegamiento dentro de las células cuenta con la ayuda de unas proteínas llamadas chaperonas que interactúan con la cadena polipeptídica recién sintetizada ayudándola a alcanzar su conformación nativa[13]. Si el plegamiento falla y la proteína está mal plegada, existen mecanismos que forman parte del sistema de control

de calidad de proteínas por los que la proteína puede ser degradada, como es el caso de las vías de degradación proteosomal 20S/26S[14].
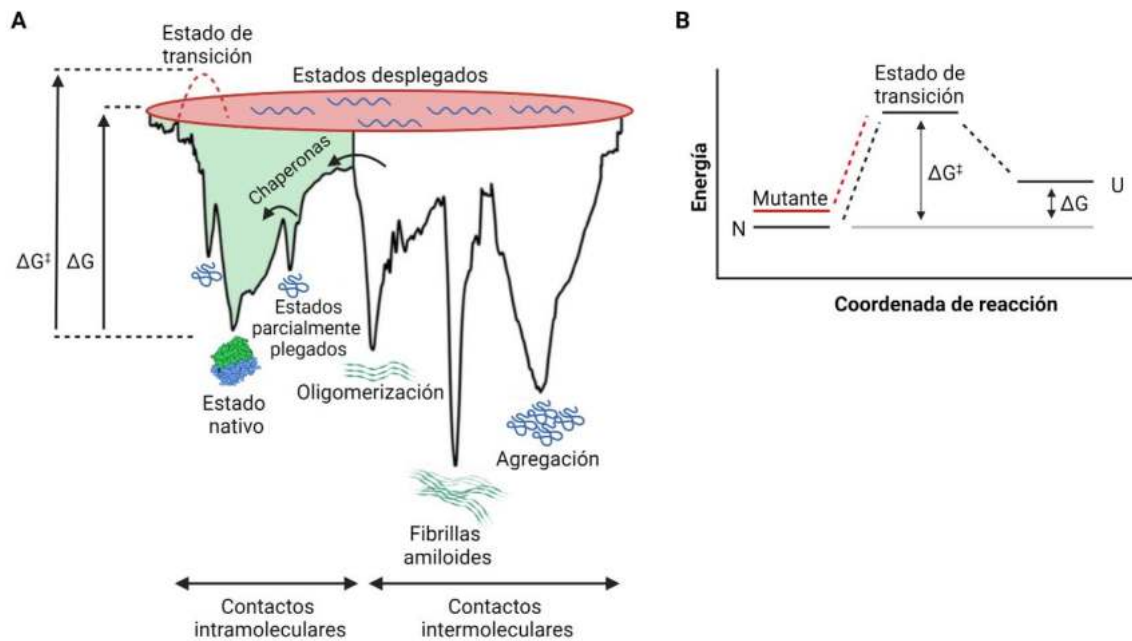

Figura 1. Representación del paisaje energético del plegamiento de una proteína (A) y del diagrama de energía libre de la desnaturalización de una proteína (B). (A) Las proteínas se pliegan poblando varias conformaciones en un paisaje energético de plegamiento. Las interacciones intramoleculares energéticamente favorables (verde) se estabilizan a medida que la proteína avanza en el paisaje hacia el estado nativo (estado de mínima energía). Las conformaciones energéticamente favorables pero que no son nativas dan como resultado poblaciones de estados cinéticamente atrapados que ocupan pozos de baja energía (estados parcialmente plegados o mal plegados). Estos estados pueden formar interacciones intermoleculares que conducen a la agregación, oligomerización y formación de fibrilla amiloides. Las chaperonas previenen esos contactos intermoleculares, ayudando a los estados parcialmente plegados a superar las barreras de energía libre, promoviendo así el plegamiento en el estado nativo. (B) Representación de la estabilidad termodinámica (ΔG) de una proteína como la diferencia de energía libre de Gibbs entre el estado desplegado (U) y el estado nativo (N); y la estabilidad cinética como la barrera de energía libre ($\Delta G^{\ddagger}$) entre el estado nativo y el estado de transición. Las mutaciones pueden favorecer un estado de más energía que disminuye la diferencia de energía libre, y por tanto la estabilidad, en ambos casos.

1.1.2 Sistema de control de calidad de proteínas

La homeostasis proteica (o proteostasis) hace referencia al correcto funcionamiento de todas las proteínas que forman el proteoma celular (entre 20000 y 25000 en humanos) y que permiten que las funciones biológicas se ejecuten exitosamente[13]. Para mantener la proteostasis las células cuentan con una compleja maquinaria conocida como sistema de control de calidad de las proteínas. Se trata de un conjunto de mecanismos dirigidos a replegar o degradar proteínas mal plegadas y evitar así la formación de agregados tóxicos[1,13].

Todos esos mecanismos están centrados en las chaperonas mencionadas anteriormente[15]. Se considera chaperona cualquier proteína que interactúa y ayuda en el plegamiento y ensamblaje de una proteína sin formar parte de su estructura final[16]. Las chaperonas no sólo ayudan durante la síntesis de la cadena polipeptídica como hemos mencionado en el plegamiento[13], también reconocen proteínas mal plegadas y son capaces de ayudarlas a replegarse correctamente[17] (Figura 1A) o, si no es posible, dirigirlas hacia su degradación[18] así como secuestrarlas en compartimentos especializados[19].

En las células eucariotas encontramos distintos tipos de chaperonas que actúan en distintos momentos. Por un lado están las chaperonas ligadas a la síntesis de proteínas (CLIPS, del inglés *chaperones linked to protein synthesis*) que están unidas funcional y físicamente a la maquinaria de traducción y se encargan de ayudar durante el plegamiento de las nuevas proteínas sintetizadas[20]. Y por otro lado, están las proteínas de choque térmico (HSPs, del inglés *heat shock proteins*) que pueden ser inducidas por factores de estrés celular y se encargan de proteger el proteoma[21]. Los distintos tipos de chaperonas cooperan funcionalmente para promover el plegamiento y la degradación. Además de ayudar en el plegamiento, las chaperonas pueden dirigir hacia la degradación a proteínas mal plegadas a través de la vía ubiquitina-proteosoma y la vía autofágica[13,22]. La vía proteosomal dependiente de ubiquitina requiere chaperonas que mantengan la proteína en estados no agregados para que el proteosoma 26S pueda actuar[13], así como chaperonas que interactúen con E3 ubiquitín ligasas para reconocer y marcar las proteínas mal plegadas para su degradación[22,23].

Además del proteosoma 26S, las proteínas parcialmente plegadas pueden ser reconocidas y degradadas por el proteosoma 20S independiente de ubiquitina[24], un mecanismo común tanto en eucariotas como en bacterias[25].

Cuando las proteínas agregadas no se pueden desplegar para ser degradadas por el proteosoma, actúa la vía autofágica con la ayuda de otras chaperonas que separan y concentran pequeños agregados proteicos en sitios específicos del citosol formando el agregosoma que será dirigido a su degradación en lisosomas[19,22].

1.1.3 Estabilidad y mutaciones de cambio de sentido

La secuencia de aminoácidos depende directamente de la secuencia genética que la codifica y puede verse afectada, entre otros factores, por mutaciones que ocurren durante la división celular en la replicación de la secuencia de ADN[26].Cuando estas mutaciones afectan a uno o varios nucleótidos se conocen como mutaciones puntuales. Las mutaciones puntuales pueden suponer la inserción o deleción de nucleótidos en la secuencia de ADN que puede provocar el cambio del marco de lectura, o pueden suponer el cambio de un nucleótido por otro. Este cambio en la secuencia de nucleótidos puede introducir un codón de parada prematuro, dando lugar a una proteína truncada. Por otra parte, puede dar lugar a mutaciones silenciosas cuando el nuevo triplete de nucleótidos codifica el mismo aminoácido, o puede provocar la sustitución de un aminoácido por otro (mutación de cambio de sentido) que puede tener un efecto más o menos neutro sobre la estabilidad y la función de la proteína[26]. Si esas mutaciones afectan a un solo nucleótido y se encuentran en al menos un 1% de la población se consideran polimorfismos de un sólo nucleótido (SNPs, del inglés *single nucleotide polymorphisms*)[27].

La estabilidad proteica es una de las propiedades más básicas de las proteínas y puede verse afectada por esas mutaciones de cambio de sentido. Como la mayoría de las proteínas necesitan estar plegadas para llevar a cabo su función, la pérdida de estabilidad puede llevar a la pérdida de función[28]. Desde el punto de vista bioquímico o biofísico, cuando hablamos de estabilidad nos referimos a estabilidad termodinámica o estabilidad cinética entre el estado plegado o nativo (N) y el estado desplegado (U) de una proteína. La estabilidad termodinámica, asumiendo el modelo de dos estados[29], hace referencia a la diferencia de la energía libre de Gibbs (ΔG) entre el estado desnaturalizado y el estado nativo de una proteína (Figura 1), que está relacionado con

una baja cantidad de estados desplegados o parcialmente desplegados en equilibrio (K_U) con la proteína funcional activa a una determinada temperatura (T):

$$\Delta G\,(T) = \Delta H\,(T) - T \cdot \Delta S\,(T) = - R \cdot T \cdot \ln\,(K_U)$$

La variación de entalpía (ΔH) y entropía (ΔS) junto a la variación de la capacidad calorífica de desnaturalización (ΔCp) proporciona una descripción completa de la estabilidad termodinámica. Una proteína con alta estabilidad termodinámica posee una energía libre de Gibbs negativa, lo que indica que su conformación nativa es termodinámicamente favorable y más estable que otras conformaciones alternativas. La estabilidad cinética está relacionada con la escala de tiempo en la que una proteína puede permanecer en su estado nativo, y de acuerdo a la teoría del estado de transición[30], está relacionada con las barreras energéticas ($\Delta G^{\ddagger}$) y los estados de transición que la proteína debe superar durante su proceso de desplegamiento (Figura 1). La constante de velocidad puede describirse como:

$$k = k_0 \cdot \exp\left(-\frac{\Delta G^{\ddagger}}{R \cdot T}\right)$$

donde k_0 es el factor pre-exponencial y $\Delta G^{\ddagger}$ es la energía libre de activación entre el estado nativo y el de transición. Una proteína con alta estabilidad cinética tiene una barrera energética alta, lo que implica que el desplegamiento ocurre de manera más lenta. Tal barrera puede garantizar que la función biológica de la proteína se mantenga, al menos durante una escala de tiempo fisiológicamente relevante, incluso si el estado nativo no es termodinámicamente estable con respecto a las formas no funcionales[31].

La alteración en la secuencia de aminoácidos puede tener un efecto patogénico por al menos dos vías diferentes[32]. Por una parte, pueden provocar la ganancia de función en la proteína por distintos mecanismos, ya sea aumentando su actividad[33], cambiando la especificidad por otros sustratos o moléculas a las que unirse[34], o favoreciendo la agregación[35]. Por otra parte, el cambio de aminoácidos puede favorecer una conformación de más energía que disminuye la barrera energética frente al estado desplegado, pudiendo provocar la pérdida de estabilidad y por tanto de función de la proteína (Figura 1B)[36–39]. Los análisis que relacionan el efecto previsto de estos cambios con la estabilidad termodinámica de una proteína, con su estabilidad celular y patogenicidad sugieren que la pérdida de estabilidad podría ser un factor principal y el

origen de ciertas enfermedades, a menudo conocidas como enfermedades conformacionales o enfermedades por plegamiento incorrecto[28].

Ya en 1941, Linus Pauling estableció el concepto de enfermedad molecular. La definió como "Una enfermedad en la que hay una anormalidad o deficiencia de una molécula determinada" tras demostrar que el origen de la anemia falciforme era debido a la alteración sufrida por una proteína por un cambio alélico en un solo gene[40]. En 1957, Vernon Ingram descubrió que era la sustitución de un ácido glutámico por una valina en la cadena β de la hemoglobina el que provocaba la pérdida de solubilidad de la proteína y como consecuencia el desarrollo de la enfermedad[41].

Desde entonces un gran número de estudios han asociado mutaciones de cambio de sentido en distintas proteínas como el origen de distintas enfermedades. Es el caso por ejemplo de la fenilcetonuria, la hiperoxaluria primaria, la galactosemia o el cáncer, entre muchas otras[36,42–45].

En el contexto celular y del desarrollo de una enfermedad no influye únicamente la estabilidad termodinámica o cinética entre el estado nativo y el estado completamente desnaturalizado[28]. Existen muchos otros factores y conformaciones proteicas que determinan que una proteína pueda llevar a cabo su función. Entre esos factores están las interacciones con el sistema de control de calidad proteica[15], las interacciones proteína-proteína[46], el tráfico celular[47] y las modificaciones postraduccionales[48]. Una mejor comprensión de la compleja relación entre la secuencia de la proteína, la estructura, el plegamiento y la estabilidad celular podría brindar nuevas posibilidades para el diagnóstico e incluso el tratamiento[28].

1.1.4 Modificaciones postraduccionales

La complejidad del proteoma humano va más allá de los aproximadamente 20000 marcos abiertos de lectura (ORFs, del inglés *open reading frames*) que hay en el genoma que codifican proteínas[49]. Las modificaciones postraduccionales (PTMs, del inglés *post-translational modifications*) son alteraciones químicas que ocurren tras la síntesis de las proteínas en el ribosoma y que dan lugar a una mayor variedad estructural y funcional de proteínas, modificando la relación genotipo-fenotipo[50]. Estas modificaciones son introducidas por enzimas muy específicas, de una forma muy selectiva y dinámica, provocando una alteración de las propiedades bioquímicas de la proteína diana[48]. Un desequilibrio de estas modificaciones está asociado con el

desarrollo de varias enfermedades como cáncer, diabetes o enfermedades neurodegenerativas[51–53].

Se han detectado más de 400 tipos de modificaciones postraduccionales que se emplean para regular procesos biológicos alterando la actividad de la proteína, la señalización, la localización, la conformación, la unión o la renovación proteica[48]. Entre estas modificaciones encontramos fosforilación, glicosilación, metilación, ubiquitinación, acetilación y muchos otros tipos (Figura 2) que pueden consultarse en www.uniprot.org[54].
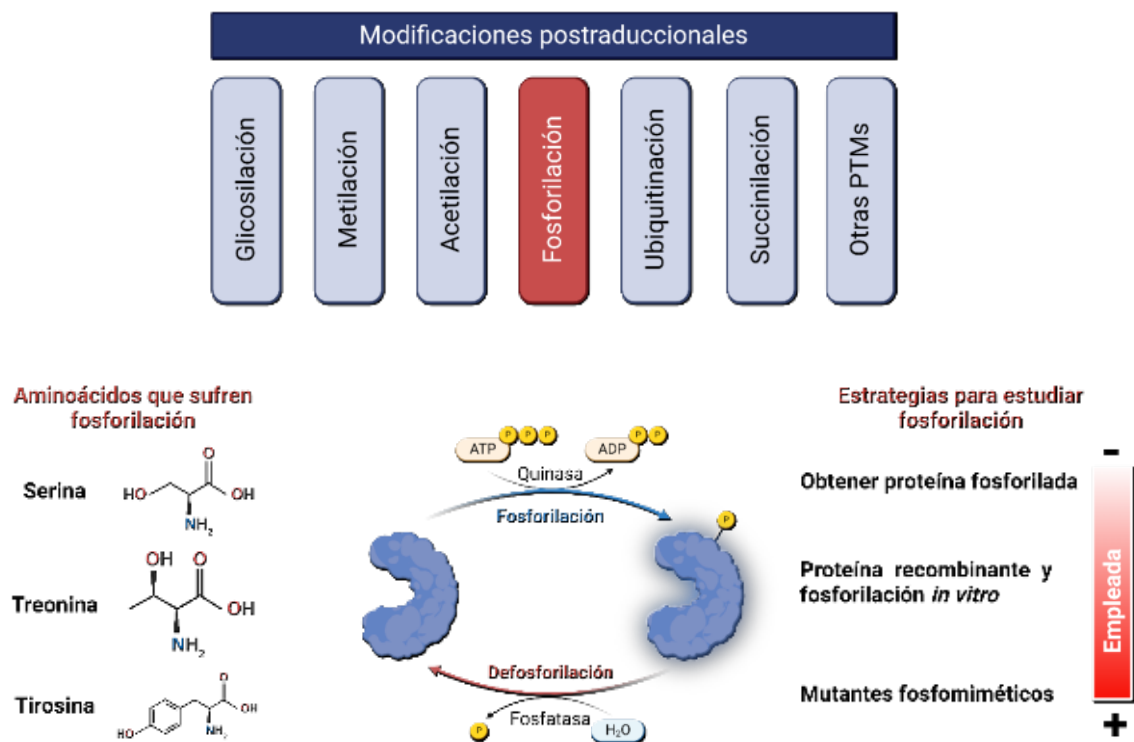
Figura 2. Modificaciones postraduccionales. El panel superior muestra los principales tipos de modificaciones postraduccionales. El panel inferior muestra la regulación por fosforilación indicando los principales aminoácidos que pueden ser fosforilados y las estrategias usadas para estudiarla.

Entre ellas, la fosforilación es la principal modificación postraduccional involucrada en procesos de regulación celular gracias a su versatilidad y rápida reversibilidad, una característica esencial en la señalización celular[55]. Esta modificación, generalmente controlada por quinasas y fosfatasas, consiste principalmente en la adición de un grupo fosforilo a residuos que contengan grupos hidroxilo (serina, treonina y tirosina) mediante la formación de un enlace éster (Figura 2)[56]. Aunque también se ha observado

que otros residuos como la histidina o la arginina puedes fosforilarse, pero es menos frecuente en eucariotas[57]. Los sitios de fosforilación pueden encontrarse en regiones intrínsecamente desordenadas (IDRs, del inglés *intrinsically disordered regions*) o dominios estructurados[58]. Dado que algunos sitios de fosforilación se encuentran enterrados en la estructura cristalina donde son inaccesibles para las quinasas correspondientes, distintos estudios han sugerido que la fosforilación es posible gracias a la dinámica de las proteínas, es decir, a su movilidad molecular intrínseca, que expone transitoriamente esos residuos permitiendo la interacción con las quinasas[59,60]. Los análisis sugieren que la fosforilación de serinas y treoninas ocurre predominantemente en IDRs, mientras que la fosforilación de tirosinas ocurre tanto en IDRs como en regiones estructuradas. En cualquier caso, la fosforilación puede inducir la transición de regiones desordenadas a ordenadas y viceversa[58,61].

El estudio de la fosforilación presenta una serie de desafíos significativos debido a su naturaleza dinámica y compleja[55]. Por una parte, obtener la fosfoproteína de fuentes naturales es complicado debido a la baja abundancia de las fosfoproteínas en comparación con las proteínas no fosforiladas[62]. Además, la fosforilación es una modificación reversible, lo que implica que los eventos de fosforilación y desfosforilación ocurren de manera dinámica y rápida en respuesta a señales celulares[56]. Esto dificulta aún más el estudio y la detección precisa de las fosfoproteínas en un momento determinado. Además, existen múltiples sitios de fosforilación dentro de una misma proteína, lo que agrega una capa adicional de complejidad al análisis de la fosforilación a nivel de sitios específicos[62]. A pesar de estos desafíos, los avances tecnológicos en técnicas de espectrometría de masas y anticuerpos específicos para la detección de fosfoproteínas han permitido grandes avances en el estudio de este proceso[63]. Sin embargo, obtener fosfoproteínas endógenas de fuentes naturales no es óptimo en estudios biofísicos o estructurales ya que se necesita una gran cantidad de proteína. Por eso, la estrategia alternativa es sobre expresar la proteína deseada en bacterias o levaduras seguido de una fosforilación *in vitro* del residuo deseado utilizando la quinasa apropiada[61]. No obstante, la falta de información sobre las quinasas específicas para la mayoría de fosforilaciones lleva a la adopción de otras aproximaciones metodológicas como es la utilización de mutantes fosfomiméticos (Tabla 1)[61].

La fosforilación de la serina y la treonina puede imitarse a través de mutaciones que sustituyen esos aminoácidos por aspartato y glutamato, que introducen una carga negativa equivalente a la fosforilación y tienen un volumen similar[64,65] Por otra parte, el glutamato se ha utilizado en ocasiones para imitar la fosforilación de la tirosina[66]. Sin embargo, el glutamato no es capaz de imitar completamente ni el volumen ni la carga de la tirosina fosforilada, y en muchos casos, esta mutación falla al imitar el efecto de la fosforilación[67]. Una alternativa entre los mutantes fosfomiméticos es el uso de aminoácidos no canónicos que también imitan las propiedades del fosfato del grupo fosfato. Para ello, esos análogos de los aminoácidos fosforilados se añaden al medio de cultivo y se modifica genéticamente la cepa donde se produce la proteína para que pueda incorporarlos[68].

Residuo fosforilado	Aminoácido canónico fosfomimético	Aminoácido no canónico fosfomimético
Serina	Ácido aspártico[69] Ácido glutámico[65]	Fosfoserina fotoprotegida[70] Ácido 2-amino-4-phosphonobutírico[71]
Treonina	Ácido aspártico[72] Ácido glutámico[73]	Fosfotreonina fotoprotegida[70] O-fosfo-L-treonina[74]
Tirosina	Ácido glutámico[66]	Fosfotirosina fotoprotegida[70] O-fosfotirosina[75] 4-fosfometil-L-fenilalanina[76] 4'-carboximetiloxi-3'-fosfonofenilalanina[77]

Tabla 1. Ejemplos de aminoácidos canónicos y no canónicos que imitan el efecto de la fosforilación de los residuos serina, treonina y tirosina.

1.1.5 Chaperonas farmacológicas

Distintos estudios han descrito el uso de compuestos de bajo peso molecular para abordar los problemas de plegamiento en proteínas asociadas a enfermedades conformacionales[78–80]. Estas pequeñas moléculas, a menudo también llamadas chaperonas químicas, pueden tener un impacto directo en la estabilidad y solubilidad de la proteína, y también pueden regular o sustituir las actividades de plegamiento de las chaperonas moleculares[81]. El efecto de estas moléculas puede ser particularmente relevante en el contexto de enfermedades por plegamiento incorrecto/conformacionales,

especialmente cuando el rescate parcial de la proteína alterada es suficiente para obtener la actividad mínima requerida o al menos para atenuar parcialmente los defectos observados[80].

Las chaperonas químicas pueden estabilizar el estado nativo, prevenir la agregación o degradación, dirigir la vía de plegamiento o promover el replegamiento de proteínas mal plegadas/desestabilizadas[82]. Cabe señalar que, dado que su mecanismo de acción no es específico, el mismo compuesto puede tener diferentes efectos si está presente en diferentes etapas del proceso de plegamiento o efectos opuestos en dos proteínas diferentes. Por el contrario, tenemos las chaperonas farmacológicas, que sí tienen una acción específica sobre una proteína diana determinada[80] y que a menudo se han desarrollado a partir de inhibidores competitivos de la proteína[83].

Se han probado chaperonas farmacológicas para muchas enfermedades, sobre todo frente a los trastornos de almacenamiento lisosomal como la enfermedad de Gaucher[84–86]. Pero también frente a la fibrosis quística, el Parkinson o la fenilcetonuria entre otras[47–49].

Las dianas más frecuentes para las que actualmente existen chaperonas farmacológicas son enzimas transferasas, seguidas de transportadores y receptores; y entre las chaperonas farmacológicas encontramos inhibidores competitivos, ligandos, agonistas/antagonistas y cofactores de proteínas como metales y flavinas[80]. En particular, en esta tesis nos interesará el papel de chaperona farmacológica de las flavinas como el flavín adenin dinucleótido (FAD). Se ha observado que la falta de FAD desestabiliza globalmente el flavoproteoma humano, mostrando que este cofactor no es sólo importante para la función de las flavoproteínas sino que también para su estabilidad[90,91]. Además, se ha visto que la suplementación con FAD (o su precursor, la vitamina B2 o riboflavina) puede estabilizar a aquellas flavoproteínas que presentan alguna mutación que reduce la afinidad por FAD, y que por lo tanto son más inestables.[91–94].

1.1.6 Variabilidad genética y medicina personalizada

Los progresos tecnológicos en la secuenciación del ADN han permitido que la secuenciación a gran escala del genoma humano sea posible tanto desde el punto de vista técnico como económico. Como resultado de esta avalancha de datos, se ha destacado el desafío de comprender las implicaciones fenotípicas de las variaciones

genéticas[26,28]. La interpretación de variantes representa un desafío notable debido a que más del 50% de las diferentes variantes identificadas en un análisis de más de 60,000 exomas humanos solo se han observado en un único individuo[95]. Esto dificulta su análisis e interpretación sobre el impacto que puede tener en el desarrollo de una enfermedad, dada la falta de información clínica y la posibilidad de llevar a cabo estudios de asociación en la población[96]. Además, esta tarea se vuelve aún más complicada debido a que muchas enfermedades tienen un origen poligénico complejo. Esto implica que el riesgo de padecer la enfermedad depende de varios genes, cuyo efecto puede ser acumulativo. Por lo que, en función de los otros genes, una variante puede tener un efecto o no[97]. Aunque el problema es complicado, el potencial para mejorar la comprensión, el diagnóstico y el tratamiento de las enfermedades humanas es enorme[28].

Las variantes de cambio de sentido constituyen más del 40% de las variantes únicas registradas en la Base de Datos de Agregación del Genoma (gnomAD, del inglés *Genome Aggregation Database)*[95], y predecir sus consecuencias fenotípicas suele ser complicado. A diferencia de las variantes sin sentido o de cambio de marco de lectura, que resultan en alteraciones significativas en la proteína codificada y suelen ser perjudiciales. Distintos estudios han demostrado que la mayoría de cambios de un único aminoácido en la secuencia de una proteína tienen un impacto menor en su función[98,99]. Sin embargo, algunas variantes son muy perjudiciales y provocan la completa pérdida de función[100–102]. Aunque estos efectos pueden variar bastante dependiendo del estado celular y de las condiciones. Así, en un contexto biológico y clínico puede haber una gran variabilidad en el número y tipo de mutaciones toleradas en un gen[28,103].

En un entorno clínico, sería útil contar con métodos sólidos y datos suficientes para la interpretación de variantes genéticas y poder clasificarlas de forma precisa como perjudiciales o benignas[104]. Además, para ampliar nuestra comprensión de los orígenes de la enfermedad, sería muy valioso contar con herramientas predictivas de los mecanismos subyacentes por los cuales las variantes conducen a la enfermedad[28].

Se están siguiendo varias estrategias para estudiar, modelar y predecir las consecuencias fenotípicas y patogénicas de las variantes proteicas. Por una parte, se están realizando estudios de alto rendimiento en los que se hacen todos los cambios posibles de aminoácidos de una proteína para ver el efecto en tanto en su función como en sus propiedades[105,106]. Estos estudios generan datos que permiten desarrollar herramientas

bioinformáticas y de aprendizaje automático. Estos algoritmos permitan predecir qué variación en la secuencia es compatible con la función de la proteína, mediante la integración de los datos existentes sobre conservación de secuencias, estructura de proteínas, función y plegamiento[107,108].

1.2. NAD(P)H:QUINONA OXIDORREDUCTASA 1

El flavoproteoma humano cuenta con alrededor de 100 proteínas que necesitan FAD o FMN para su función catalítica[109]. La mayoría de las enzimas dependientes de flavina catalizan procesos de oxidación-reducción en rutas metabólicas primarias como el ciclo del ácido cítrico, b-oxidación de ácidos grasos y degradación de aminoácidos. Además, tienen un papel importante en procesos como la biosíntesis de otros cofactores y hormonas esenciales como la coenzima A, la coenzima Q, el grupo hemo, el piridoxal 5'-fosfato, esteroides y tiroxina, y en la regulación del metabolismo del folato[109]. Debido a su participación en tantos procesos metabólicos, aproximadamente 2/3 de las flavoproteínas humanas, cuando están alteradas por el efecto de mutaciones de cambio de sentido, están asociadas a enfermedades genéticas[93,109,110]. Ese es el caso de la NAD(P)H:quinona oxidorreductasa 1 (NQO1)[111], y por eso vamos a estudiar en esta tesis cómo se ve afectada por esas mutaciones.

1.2.1 Expresión, localización y estructura

La proteína humana NQO1 (EC 1.6.5.2) es una flavoenzima codificada por el gen *NQO1* localizado en el brazo largo del cromosoma 16 (16q22.1) (Figura 3)[112]. Esta proteína se expresa en muchos tejidos humanos, principalmente en células epiteliales y endoteliales de riñón y del tracto gastrointestinal, localizándose principalmente en el citosol celular, aunque también se encuentra en bajos niveles en el núcleo, en el retículo endoplasmático y en mitocondrias[113]. Además, se han observado altos niveles de expresión de NQO1 en algunas células tumorales, que podría estar protegiéndolas de las especies reactivas de oxígeno[114].

Su tasa de expresión está controlada tanto de forma basal como en condiciones de estrés oxidativo por el elemento de respuesta antioxidante (ARE, del inglés *antioxidant response element*), que se encuentra en la región promotora del gen y que está regulado por el factor nuclear eritroide similar al factor 2 (Nrf2) (Figura 3)[115]. Nrf2 induce la expresión del gen *NQO1* junto a otras enzimas desintoxicantes adicionales en respuesta a xenobióticos, agentes oxidantes, antioxidantes, electrófilos, radiación ionizante y metales pesados[116].
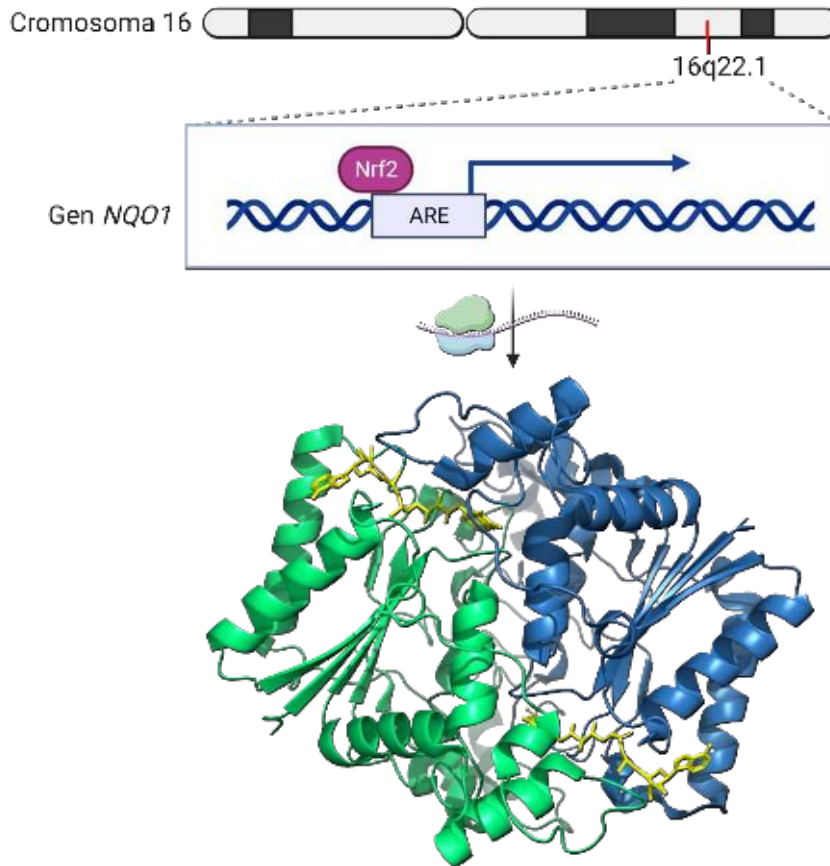
Figura 3. Representación esquemática de la localización y regulación del gen *NQO1* y de la estructura de la proteína. El panel superior muestra la localización del gen *NQO1* en el locus 16q22.1 del cromosoma 16. El panel medio muestra el control de su expresión por el elemento de respuesta antioxidante (ARE) a través del factor nuclear eritroide similar al factor 2 (Nrf2). El panel inferior muestra la estructura dimérica de NQO1 con las moléculas de FAD (amarillo) (PDB 2F1O)[117] creada con PyMol (www.pymol.org).

Estructuralmente, NQO1 es una enzima homodimérica que cuenta con una molécula de FAD unida por monómero (Figura 3), cada uno de los cuales está formado por 274 aminoácidos y tiene una masa molecular aproximada de 31 kDa[117]. Cada monómero está formado por dos dominios: un dominio catalítico N-terminal (NTD, del inglés *N-terminal domain*) (residuos 1-224) donde se une el FAD, y un dominio C-terminal (CTD, del inglés *C-terminal domain*) (residuos 225-274) que se encarga de estabilizar al dímero y es crítico para la unión del nicotin adenin dinucleótido (NADH) y el sustrato. El dominio catalítico presenta una hoja plegada beta, formada por 5 hebras paralelas, que se encuentra en el centro de 5 hélices alfa, mientras que el dominio C-terminal está

formado por una horquilla beta antiparalela, una hélice alfa y varios lazos. La NQO1 presenta dos sitios activos que se encuentran cerca de la interfase de las dos subunidades, estando una molécula de FAD unida en cada uno. Cada sitio activo tiene tres lugares de unión: uno sitio de unión para el FAD (FBS, del inglés *FAD binding site*), uno para la parte adenina y ribosa del NAD(P)H y otro para la parte nicotinamida del NAD(P)H o para la molécula aceptora de electrones[117]. El FAD está unido de forma no covalente en el sitio activo con una alta afinidad, con cierto grado de cooperatividad negativa entre ambos sitios en la unión[118].

1.2.2 Mecanismo enzimático

NQO1 es una enzima que lleva a cabo reacciones de reducción de dos electrones por un mecanismo ping-pong. En primer lugar, el NAD(P)H se une a NQO1 para formar un complejo NQO1-NAD(P)H. A continuación, la molécula de FAD es reducida por NAD(P)H para formar $FADH_2$ y $NAD(P)^+$. La salida del $NAD(P)^+$ del sitio activo permite la unión del sustrato, a menudo una quinona, que es reducido por el $FADH_2$ regenerando de nuevo el FAD (Figura 4)[119]. Esta reacción enzimática es inhibida de manera competitiva por el dicumarol, impidiendo la unión del NAD(P)H[117,120].
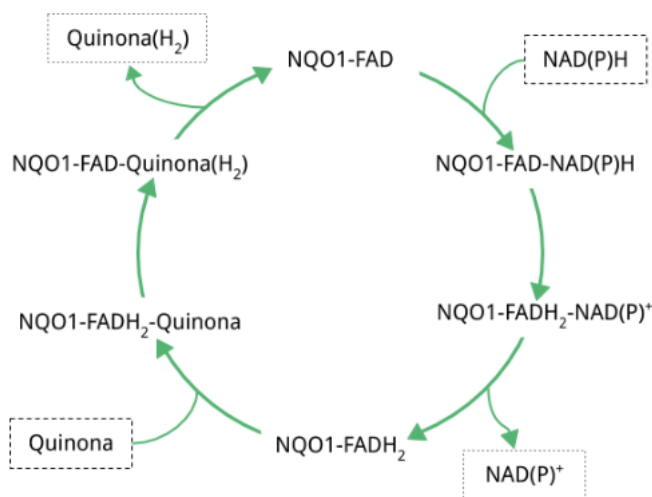

Figura 4. Mecanismo ping-pong de NQO1.

1.2.3 Funciones de NQO1

NQO1 es una proteína muy versátil que tiene distintos papeles biológicos y que presenta tanto funciones enzimáticas como no enzimáticas[121].

- **Detoxificación de quinonas y otros compuestos**

La detoxificación de quinonas es la función más conocida y estudiada de NQO1. NQO1 cataliza la reducción de dos electrones de quinonas a hidroquinonas, evitando las vías de reducción de un electrón que producen semiquinonas[122]. Las semiquinonas son muy reactivas y pueden reaccionar con el oxígeno molecular para formar radicales superóxido. De esta manera NQO1 actúa como una enzima detoxificante que protege a la célula del estrés oxidativo y previene la carcinogénesis, ya que las hidroquinonas son excretadas más fácilmente y evita que se produzcan radicales libres por el ciclo redox[123]. La enzima NQO1 también cataliza la reducción de dos electrones de compuestos nitroaromáticos, colorantes azoicos y quinoneiminas, evitando la peroxidación lipídica de la membrana celular y el daño oxidativo de proteínas y del ADN a través de las especies reactivas de oxígeno que generan[121,124].

- **Enzima antioxidante**

NQO1 defiende a la célula contra la oxidación al mantener la forma reducida y activa de antioxidantes endógenos. Principalmente, la ubiquinona (coenzima Q (CoQ)) y la vitamina E (α-tocoferol quinona (TQ)), dos antioxidantes esenciales solubles en lípidos que protegen la membrana celular evitando la peroxidación de lípidos[125–127].

NQO1 también puede contribuir directamente a la protección antioxidante. Como se ha indicado anteriormente, es capaz de evitar la formación de radicales superóxido al reducir las quinonas a hidroquinonas[123]. No obstante, también presenta actividad superóxido reductasa y es capaz de eliminar directamente los radicales superóxido en una reacción mediada por el cofactor[121,127]. Aunque, la velocidad de reacción de NQO1 con el superóxido es muy baja en comparación a otras enzimas que llevan a cabo esa función como la superóxido dismutasa, y se cree sólo tiene relevancia biológica en aquellas condiciones y tipos celulares en los que se encuentran niveles de expresión de NQO1 muy altos[128,129].

- **Unión a proteínas y ARN**

Una función importante de NQO1 es su capacidad para interactuar físicamente con otras proteínas y existen numerosos ejemplos en los que tales interacciones son relevantes para la función celular. NQO1 se asocia con proteínas supresoras de tumores como p53, p33 y p73 en una interacción proteína-proteína, estabilizándolas contra la degradación por el proteosoma 20S independiente de ubiquitina[130–133]. Esta estabilización es dependiente de la unión de NAD(P)H[134]. La cantidad de proteínas conocidas con las que interactúa NQO1 es cada vez mayor. Además de las proteínas supresoras tumorales, se ha visto que también estabiliza entre otras proteínas a la ornitina descarboxilasa[135], a PGC-1α[136] y, recientemente, se ha visto que estabiliza al factor inducible por hipoxia 1 alfa (HIF-1α) que está implicado en el crecimiento tumoral y la metástasis[137].

Además de las interacciones directas de NQO1 con estas proteínas, se ha demostrado que NQO1 está interactuando directamente con la porción central (20S) del proteosoma ,sugiriendo que NQO1 modula la actividad deproteosomal[138,139]. Sin embargo, es un proceso complejo y los mecanismos moleculares exactos por los que NQO1 inhibe la degradación de proteínas no se conocen[128].

Por otra parte, distintos estudios han demostrado que NQO1 se une a moléculas de ARN, posiblemente regulando su traducción[140]. Es una función poco estudiada aún, pero cada vez más estudios apoyan este papel biológico, ya que aparte de unirse a moléculas de ARN se ha observado que también interactúa directamente con la proteína L13A de la subunidad 60S del ribosoma[128,141].

- **Activación de fármacos contra el cáncer**

La reducción de quinonas por NQO1 puede utilizarse para obtener hidroquinonas que son biológicamente activas frente a células tumorales. Se ha observado que NQO1 se sobre expresa en distintos tipos de cáncer en comparación con tejidos normales. Por eso, la activación de quinonas como mitomicina C, E09, RH1 y β-lapachona a compuestos citotóxicos que llevan a la muerte celular[142–144], es una estrategia prometedora frente a aquellos tumores que presentan niveles altos de expresión de NQO1[145].

1.2.4 Polimorfismos de NQO1

Los estudios de secuenciación han mostrado la variabilidad genética de NQO1 que existe en la población. Entre las variantes encontradas, un alto porcentaje corresponde a

mutaciones de cambio de sentido que pueden tener efectos más o menos deletéreos en la proteína. En gnomAD, existen actualmente 106 mutaciones de cambio de sentido (consultado en julio de 2023), mientras que en el Catálogo de Mutaciones Somáticas en Cáncer (COSMIC, del inglés *Catalogue Of Somatic Mutations In Cancer*)[146] existen actualmente 57 mutaciones de cambio de sentido (consultado en julio de 2023). De las mutaciones recogidas en estas bases de datos, las dos variantes alélicas más frecuentes han sido estudiadas en profundidad:

- **P187S**

La variante P187S (rs1800566) es el resultado de la sustitución de una citosina por una timina en la secuencia de ADN, que da como resultado la sustitución de una prolina, en la posición 187 en la secuencia de aminoácidos, por una serina [147]. Se trata de la variante de NQO1 más estudiada y distintos estudios la han asociado con una mayor susceptibilidad de desarrollar algunos tipos de cáncer[147,148] así como con un mal pronóstico cuando ocurre en homocigosis[146]. Este polimorfismo se ha encontrado en varias líneas celulares, muchas de ellas líneas celulares tumorales, que presentaban niveles de ARN mensajero normales, pero en las que no se detectaban ni actividad ni expresión proteica[149–151].

Estructuralmente, P187S parece que no afecta a la conformación global de la proteína, ya que, los análisis de cristalografía de rayos X de NQO1 P187S (PDB: 4CF6) no han mostrado diferencias frente a los realizados en NQO1 WT (PDB: 2F1O) cuando estas enzimas están formando complejos ternarios con FAD y dicumarol[152]. Sin embargo, en ausencia de dicumarol, se ha demostrado que P187S tiene un CTD muy dinámico[153]. Esta variante provoca cambios locales en la estructura y dinámica de la proteína en su forma apo (en ausencia de FAD unido) que se propagan a través de la estructura proteica a sitios funcionales que se encuentran alejados[154,155]. Entre esos cambios, destaca la alteración estructural y dinámica del sitio de unión del FAD, principalmente del bucle 57-66 y de la región entre la tirosina 127 y la tirosina 129, que favorece conformaciones no competentes para la unión del FAD[153,156]. Esto conlleva que P187S presente una afinidad por FAD 10-40 veces menor que la proteína WT[94,152]. Además, hay una desestabilización del núcleo de la proteína que disminuye la estabilidad conformacional de NQO1, haciendo que P187S sea cinéticamente inestable a temperatura fisiológica[94]. La diferente estabilidad y dinámica del CTD en la variante

P187S es crítica para una mayor ubiquitinación y degradación incluso como holoproteína (es decir, cuando la proteína tiene unido el cofactor FAD), mientras que la degradación de la enzima WT ocurre principalmente cuando se encuentra como apoproteína, dependiendo por lo tanto de los niveles de flavina intracelulares[91,153]. Además, este CTD termodinámicamente inestable de P187S favorece su degradación por la vía proteosomal independiente de ubiquitina[132].

El dicumarol se une al sitio activo de NQO1 de manera competitiva respecto al NAD(P)H y sólo provoca pequeños cambios en la estructura proteica[117]. Sin embargo, sí tiene un fuerte efecto en la dinámica de NQO1, afectando a la estabilidad local del núcleo de NQO1 que se propaga a regiones alejadas en la estructura proteica[157]. La unión del dicumarol a NQO1 se traduce en un aumento de la estabilidad conformacional al estabilizar el núcleo de la proteína, y ese efecto se propaga hasta el NTD aumentando la resistencia a la digestión con termolisina, así como al CTD[153,154]. Como hemos indicado P187S tiene un dominio C-terminal muy inestable y flexible, que tras la unión de dicumarol se estabiliza y rigidifica[154]. Esto se traduce en una estabilización intracelular de P187S, que provoca una disminución de su ubiquitinación y un aumento de los niveles de proteína dentro de la célula[91,153].

- **R139W**

La variante p.R139W (rs1131341) surge del cambio de una citosina por una timina, que como resultado provoca la sustitución de una arginina por un triptófano en la posición 139 de la secuencia aminoacídica. Algunos estudios han asociado este polimorfismo con un mayor riesgo de desarrollar algunos tipos de cáncer como la leucemia linfoblástica aguda infantil[158]. Este polimorfismo se ha aislado de líneas celulares de cáncer resistentes al tratamiento con mitomicina C (MMC)[159,160]. La MMC es reducida a su forma activa por NQO1 y es utilizada como fármaco frente a ciertos tumores, y se ha observado que las líneas con R139W presentan una menor sensibilidad a MMC por lo que supone un peor pronóstico en algunos tipos de cáncer[159,161]. Este polimorfismo tiene dos efectos distintos: 1) por una parte, altera el *splicing* del ARN que lleva a la pérdida del exón 4, que como consecuencia genera una proteína inestable que carece de parte del sitio de unión del FAD[153,162]; 2) por otra parte, la sustitución de la arginina por el triptófano, que tiene lugar en un bucle expuesto al solvente, sólo produce efectos leves en la estabilidad térmica, función catalítica o unión del FAD. Por lo tanto, la

mayor consecuencia patogénica de este polimorfismo es la alteración del *splicing* del ARN[94,153,162].

1.2.5 Enfermedades asociadas a NQO1

La naturaleza multifuncional de NQO1 implica que su alteración puede estar asociada al desarrollo de una gran variedad de enfermedades (Figura 5)[163]. Entre ellas, se ha sugerido su asociación con el cáncer. Por una parte, la falta de actividad de NQO1 como en el polimorfismo P187S se ha relacionado con una mayor riesgo de desarrollar distintos tipos de cáncer a lo largo de la vida[147,150]. Este polimorfismo reduce la capacidad antioxidante de la célula y aumenta la probabilidad de que se formen semiquinonas reactivas que pueden dañar el ADN, proteínas y lípidos[164]. Además, dado que NQO1 interacciona y estabiliza a proteínas supresoras de tumores como p53 y p73[131,132], una reducción en los niveles de NQO1 puede evitar que se produzca la apoptosis en aquellas células dañadas que pueden transformarse en células cancerosas[131]. Aunque esta asociación no se ha demostrado, sí se ha asociado este polimorfismo con un mal pronóstico en algunos tipos de cáncer, debido a una peor respuesta a los tratamientos[165,166]. Por otra parte, muchas células tumorales tienen niveles de NQO1 incrementados que pueden estar protegiendo a las células cancerígenas de las especies reactivas de oxígeno[114]. Por eso, la inhibición de NQO1 se ha propuesto como una posible terapia frente al cáncer. aunque actualmente no hay ningún tratamiento basado en la inhibición de la NQO1[167,168].
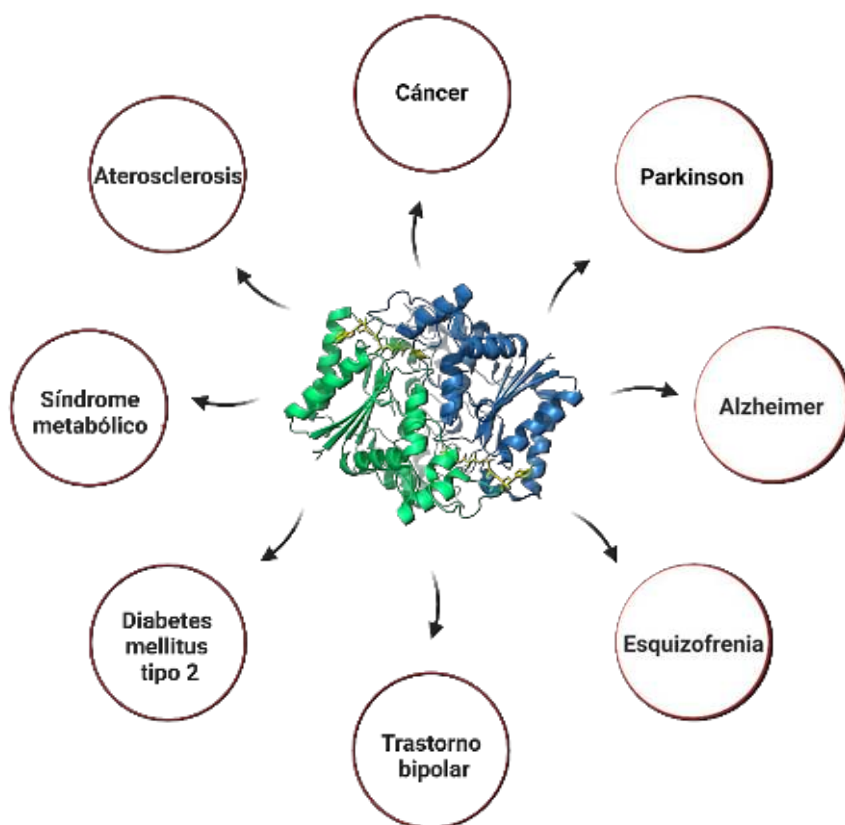
Figura 5. Enfermedades asociadas a NQO1.

NQO1 también ha sido asociada a trastornos del sistema nervioso. Se ha observado que se expresa en neuronas en la sustancia negra cuando está en condiciones de estrés oxidativo[168,169] y que su función es necesaria para la correcta mielinización de las neuronas[170]. De hecho, los ratones que carecen del gen *NQO1* presentan convulsiones y alteración del comportamiento[170]. Como enzima antioxidante, NQO1 juega un papel importante en este órgano donde hay un alto consumo de oxígeno y el potencial de generación de estrés oxidativo es alto[171,172]. Estudios han sugerido que la disminución de la actividad de la NQO1 puede estar asociada con enfermedades neurodegenerativas como la esclerosis múltiple, el Parkinson o el Alzheimer, y trastornos psiquiátricos como la esquizofrenia y el trastorno bipolar[168]. Aunque no hay una asociación clara entre el polimorfismo P187S y un mayor riesgo de padecer estas enfermedades[168,173,174], sí se ha observado que este polimorfismo puede estar asociado a un mayor riesgo de desarrollar Alzheimer[175], y se cree que su reducida actividad puede favorecer una mayor neurotoxicidad y formación de agregados amiloides[168].

NQO1 es una enzima presente en el sistema cardiovascular que protege contra la inflamación y el estrés oxidativo[176]. Su expresión puede ser inducida por diversos

estímulos, y su sobreexpresión en células vasculares y cardiacas proporciona una defensa adaptativa[177]. Se ha observado que el polimorfismo P187S está asociado tanto con una mayor prevalencia de placas ateroscleróticas como con un menor riesgo de accidentes cerebrovasculares en pacientes con aterosclerosis[178,179]. También se ha asociado con enfermedades como la diabetes mellitus tipo 2 o el síndrome metabólico, dado que reduce el estrés oxidativo y esto protege frente a la resistencia a insulina o la dislipemia[180]. Sin embargo, se requiere más investigación para comprender su papel en estas enfermedades y sus posibles aplicaciones terapéuticas[163].

1.2.6 Modificaciones postraduccionales en NQO1

La estabilidad y función de NQO1 puede modularse por la acción de modificaciones postraduccionales[181]. Actualmente, hay identificadas 43 modificaciones postraduccionales de NQO1 en PhosphoSitePlus® (www.phosphosite.org[182], consultado en julio de 2023) que incluye 12 sitios de fosforilación, 9 sitios de acetilación, 17 sitios de ubiquitinación, 3 sitios de succinilación y 2 sitios de sumoilación (Figura 6). Estos sitios se encuentran distribuidos por toda la estructura de NQO1 y en ocasión un mismo sitio puede ser objeto de más de una modificación postraduccional. La mayoría de estas modificaciones han sido detectados en estudios proteómicos por espectrometría de masas, pero apenas hay estudios que hayan caracterizado el efecto de esas modificaciones y que hayan determinado su papel biológico. No obstante, algunas de estas modificaciones postraduccinales han sido caracterizadas directamente[183–185], así como la estabilidad local de algunos de los sitios que experimentan las modificaciones postraduccionales mencionadas anteriormente[157,181].
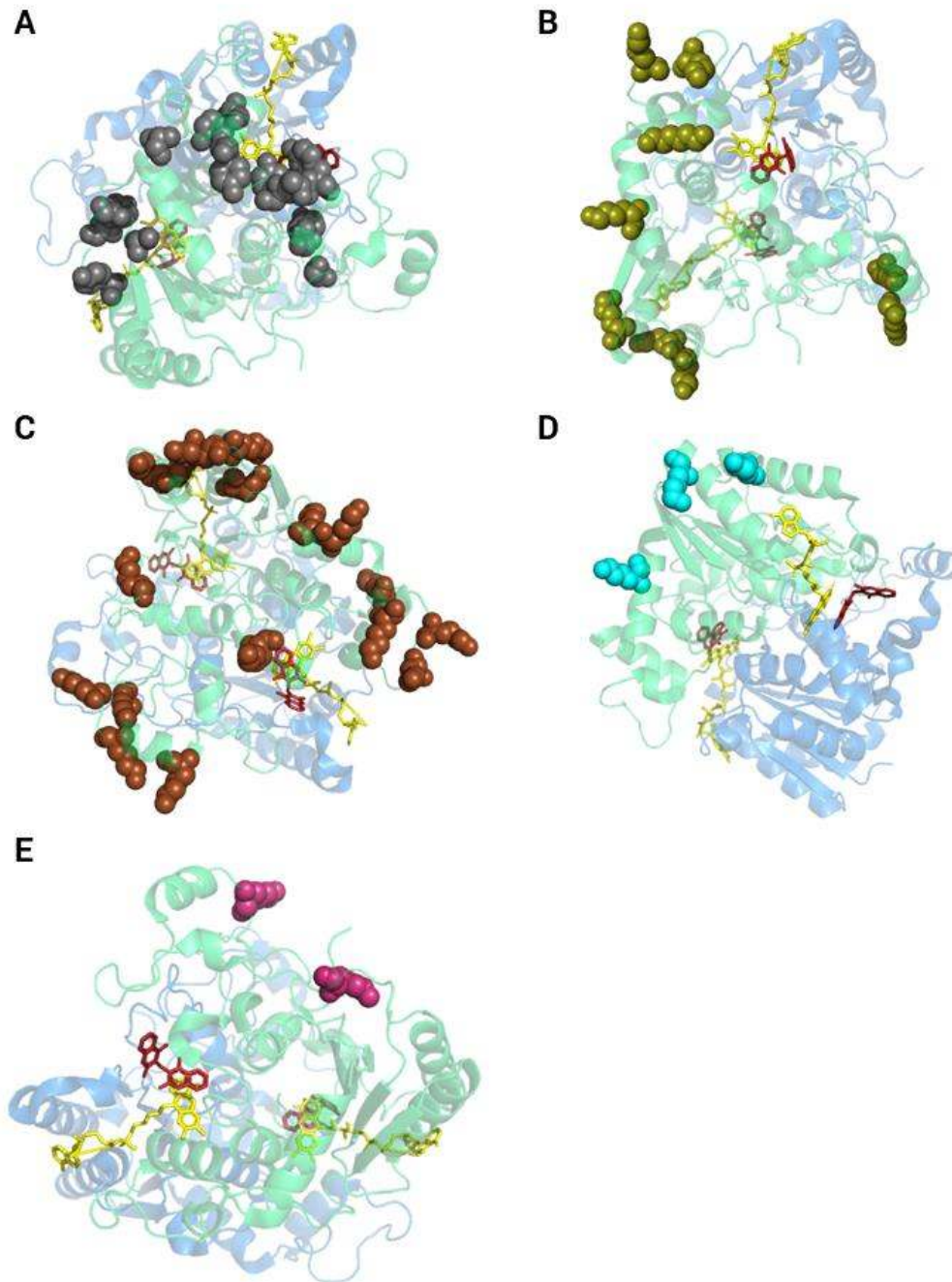
Figura 6. Localización estructural de los residuos afectados por modificaciones postraduccionales (PDB 2F1O[117]**).** (A) Sitios de fosforilación. (B) Sitios de acetilación. (C) Sitios de ubiquitinación. (D) Sitios de succinilación. (E) Sitios de sumoilación.

La mayoría de los sitios de ubiquitinación se encuentran en regiones expuestas al solvente y presentan una estabilidad estructural media-baja, facilitando el acceso para la adición de ubiquitinas[157]. Este proceso está asociado a la degradación proteosomal dependiente de ubiquitina, sobre todo para aquellas lisinas localizadas en el extremo C-

terminal, que se considera como sitio de iniciación de la degradación del polimorfismo P187S y de la proteína WT cuando se encuentra en su forma apo al presentar una mayor dinámica[154,186].

Los sitios de acetilación, que en algunos casos se ha visto que se solapan con los de ubiquitinación, se encuentran también en regiones expuestas al solvente[157,181]. Recientemente, se ha estudiado su acetilación *in vitro*, y se ha observado que la acetilación de aquellos lisinas que se encuentran cerca del sitio de unión al FAD disminuyen la actividad catalítica de la proteína[185], posiblemente alterando la red de interacciones electrostáticas que son especialmente importantes para la unión del FAD[156,183].

Como se ha indicado anteriormente, se han identificado 12 sitios de fosforilación en NQO1. Actualmente, sólo existen estudios sobre el efecto que tiene la fosforilación en tres de ellos (S40, S82 y T128)[183,184]. Resulta interesante que las regiones en las que se encuentran algunos de los sitios de fosforilación presentan una estabilidad estructural media-baja, incluso cuando están enterrados en la estructura de la proteína[157]. Además, la estabilidad de esas regiones depende del estado de ligación de NQO1, es decir, de si tiene unido FAD y dicumarol[157]. Esto sugiere que la fosforilación puede ser dependiente de la unión de ligandos y de los cambios que eso conlleva en la estabilidad local de esas regiones de la proteína, o que puede ocurrir simultáneamente con el proceso de traducción antes de que esos residuos queden enterrados en la estructura proteica[181].

Se ha observado que la fosforilación de S40 y T128 ocurre a través de la quinasa Akt, favoreciendo la ubiquitinación de NQO1 por Parkin y conduciéndola a su degradación, aunque no afecta a su actividad enzimática. El papel de T128 parece tener una mayor relevancia, siendo necesaria su fosforilación para desencadenar la fosforilación en S40[184]. Por el contrario, se desconoce la quinasa encargada de la fosforilación de S82, por lo que el efecto de la fosforilación de este residuo se ha estudiado usando la mutación fosfomimética S82D[183]. Se ha observado que esta variante disminuye la estabilidad conformacional de NQO1, así como la afinidad por FAD y la capacidad catalítica de la enzima. Esa reducción en la afinidad por el FAD favorece el estado apo de la proteína dentro de la célula, disminuyendo su estabilidad al facilitar su degradación por el proteosoma[91,183].

2. OBJETIVOS

Esta tesis está centrada en el efecto de mutaciones de cambio de sentido naturales y modificaciones postraduccionales sobre la enzima NQO1. Por ello, los objetivos principales de esta tesis doctoral son:

1. Estudiar el impacto de mutaciones naturales sobre la estabilidad y la función de NQO1.

2. Evaluar el resultado de la "fosforilación" específica de sitio sobre la estabilidad y función de NQO1.

3. Estudiar como como el efecto se propaga a lo de la estructura y comprender los efectos alostéricos y pleiotrópicos de estas variantes.

3. RESULTADOS

Los resultados obtenidos en esta tesis doctoral se presentan en este capítulo por la modalidad de agrupación de publicaciones.

Publicación 1:

Pacheco-García, J. L., Cano-Muñoz, M., Sánchez-Ramos, I., Salido, E., & Pey, A. L. (2020). Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein. Journal of Personalized Medicine, 10(4), 207.

Impact Factor (2020): 4.945

Category name / Quartile in Category: Health Care Sciences & Services / Q1

Publicación 2:

Pacheco-García, J. L., Anoz-Carbonell, E., Loginov, D. S., Kavan, D., Salido, E., Man, P., Medina, M. & Pey, A. L. (2023). Counterintuitive structural and functional effects due to naturally occurring mutations targeting the active site of the disease-associated NQO1 enzyme. The FEBS Journal, 290(7), 1855-1873.

Impact Factor (2022): 5.4

Category name / Quartile in Category: Biochemistry & Molecular Biology / Q2

Publicación 3:

Pacheco-Garcia, J. L., Anoz-Carbonell, E., Vankova, P., Kannan, A., Palomino-Morales, R., Mesa-Torres, N., Salido, E., Man, P., Medina, M., Naganathan, A. N. & Pey, A. L. (2021). Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD (P) H: quinone oxidoreductase 1 and their pharmacological rescue. Redox Biology, 46, 102112.

Impact Factor (2021): 10.787

Category name / Quartile in Category: Biochemistry & Molecular Biology / D1

Publicación 4:

Pacheco-Garcia, J. L., Anoz-Carbonell, E., Loginov, D. S., Vankova, P., Salido, E., Man, P., Medina, M., Palomino-Morales, R. & Pey, A. L. (2022). Different phenotypic outcome due to site-specific phosphorylation in the cancer-associated NQO1 enzyme studied by phosphomimetic mutations. Archives of Biochemistry and Biophysics, 729, 109392.

Impact Factor (2022): 3.9

Category name / Quartile in Category: Biophysics / Q1

PUBLICACIÓN 1

Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein.

Naturally-Occurring Rare Mutations Cause Mild to Catastrophic Effects in the Multifunctional and Cancer-Associated NQO1 Protein

Juan Luis Pacheco-García [1] , **Mario Cano-Muñoz** [1] , **Isabel Sánchez-Ramos** [1] , **Eduardo Salido** [2] **and Angel L. Pey** [3,*]

[1] Departamento de Química Física, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain; juanlupacheco@correo.ugr.es (J.L.P.-G.); mariocano@ugr.es (M.C.-M.); isanchezb@correo.ugr.es (I.S.-R.)

[2] Centre for Biomedical Research on Rare Diseases (CIBERER), Hospital Universitario de Canarias, 38320 Tenerife, Spain; edsalido@gmail.com

[3] Departamento de Química Física y Unidad de Excelencia de Química Aplicada a Biomedicina y Medioambiente (UEQ), Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

* Correspondence: angelpey@ugr.es; Tel.: +34-958243173

Abstract: The functional and pathological implications of the enormous genetic diversity of the human genome are mostly unknown, primarily due to our unability to predict pathogenicity in a high-throughput manner. In this work, we characterized the phenotypic consequences of eight naturally-occurring missense variants on the multifunctional and disease-associated NQO1 protein using biophysical and structural analyses on several protein traits. Mutations found in both exome-sequencing initiatives and in cancer cell lines cause mild to catastrophic effects on NQO1 stability and function. Importantly, some mutations perturb functional features located structurally far from the mutated site. These effects are well rationalized by considering the nature of the mutation, its location in protein structure and the local stability of its environment. Using a set of 22 experimentally characterized mutations in NQO1, we generated experimental scores for pathogenicity that correlate reasonably well with bioinformatic scores derived from a set of commonly used algorithms, although the latter fail to semiquantitatively predict the phenotypic alterations caused by a significant fraction of mutations individually. These results provide insight into the propagation of mutational effects on multifunctional proteins, the implementation of *in silico* approaches for establishing genotype-phenotype correlations and the molecular determinants underlying loss-of-function in genetic diseases.

Keywords: missense mutation; genetic disease; protein structure-function; genotype–phenotype correlations; multifunctional proteins

1. Introduction

Due to technological advances in DNA sequencing technology, we are begining to realize the tremendous genetic variability in the human genome [1,2] and to accurately map the genetic alterations associated with certain diseases [3]. A single human genome carries thousands of missense mutations [1,2,4]. However, the potential physio-pathological implications of this genetic diversity are unclear, partly due to our limited ability to predict the effects of missense variants on predisposition to disease [5,6]. Our limited capacity for establishing large-scale genotype–phenotype correlations can be explained by different reasons: (i) the relationships between molecular, pathogenic and organismal effects are complex, even for diseases with *simple* Mendelian inheritance [1,7,8]; (ii) genotype–phenotype correlations are improved when experimental data for molecular effects of

mutations are available, but current predictive tools still underperform experimental characterization, particularly for mild to moderate phenotypes [8–11]; (iii) although many diseases are caused by loss-of-function mutations [12], an operational definition of *loss-of-function* is difficult because human proteins display many functional and regulatory traits. In this context, genotype–phenotype correlations require an integrated understanding on how a single mutation may affect multiple molecular functions *simultaneously* and, consequently, how these molecular effects translate into pathogenic and fitness consequences [1,13–23].

The molecular mechanisms by which missense mutations cause loss-of-function are many-fold [24] and include accelerated protein degradation [5,9,11,25,26], enhanced protein aggregation [14,15,21], catalytic and regulatory alterations [21,27–29] and altered biomacromolecular interactions [10,30]. Importantly, the common molecular origin for all these coexisting mechanisms appears to be the structural and energetic perturbation caused by missense mutations [5,8,11,19,24,26]. These perturbations may differently contribute to the molecular phenotype associated with a given missense variant [19]. Therefore, a fundamental issue to understand the effect of missense variants on multifunctional proteins will come from a deep knowledge on how mutational effects are propagated through the protein structure affecting multiple functional features [19,31].

There are several strategies to develop tools for estimating the impact of missense mutations on protein stability and establishing the *potential pathogenicity* of missense variants that can be broadly classified into: (i) *sequence-based* methods that use multiple sequence analysis combined with simple physical estimations of the mutational effect [5,32]—a clear advantage of these methods is that high-resolution structural information on the protein of interest is not required; (ii) *structure-based* methods that (quantitatively) evaluate mutational effects on protein thermodynamic stability (as a change in the unfolding free energy, ΔG_U). These are often trained and tested using large datasets of experimentally determined mutational effects on ΔG_U [33–36] and work rather well overall but fail for individual mutations [37]. Importantly, mutational effects are universally distributed across protein structures, with solvent-exposed mutations causing much milder effects than those affecting buried residues [35,36]. These studies have three important implications to understand the functional consequences of the vast genetic diversity present in the human genome. First, between $^1/_2$ and $^2/_3$ of all possible mutations have minimal effects on conformational stability [35,36] and mild functional effects in vivo [38]. Second, rather small effects on conformational stability ($\Delta\Delta G_U \sim -2$ kcal·mol^{-1}) can subtantially accelerate protein degradation and cause loss-of-function [5,9,11,26]. Third, missense mutations may differently affect *native* vs. *partially folded* states and, consequently, operate through different loss-of-function mechanisms [19,39].

We have recently used the enzyme NAD(P)H quinone oxidoreductase 1 (NQO1; EC 1.6.5.2) to investigate how missense mutations may affect different functional features in a multifunctional human protein [19,39–45]. NQO1 folds into functional dimers of 62 kDa [46,47] and each monomer has 274 residues divided into two domains: an N-terminal domain (NTD, residues 1–225) that contains a tightly bound FAD and most of the catalytic site and the monomer:monomer interface (MMI), and a C-terminal domain (CTD, residues 225–274) that completes the MMI and the active site [44,46,48–51]. FAD binding is essential for NQO1 function and intracellular stability, since the wild-type (WT) protein lacking bound cofactor (NQO1$_{apo}$) is highly flexible and rapidly targeted for degradation [41,44,49,52]. NQO1 is a multifunctional protein with enzymatic and non-enzymatic functions [47,53]. NQO1 catalyzes the two electron reduction of quinones, acting in the antioxidant defense, maintaining several vitamers and biomolecules in their reduced state and scavenging superoxide radicals [54–58]. The enzymatic cycle contains two half-reactions following a *ping-pong* mechanism: a first step in which NAD(P)H binds to the FAD-bound enzyme (NQO1$_{holo}$), reducing the flavin to FADH$_2$ and releasing NAD(P)$^+$, and a second half-reaction in which FADH$_2$ reduces the bound substrate regenerating the NQO1$_{holo}$ state [47,49,59]. Dicoumarol (Dic) acts as a competitive inhibitor of NAD(P)H in the first half-reaction and also inhibitis the oxidative (second) half-reaction [53,59,60]. In addition, NQO1 develops key macromolecular interactions with regulatory functions, particularly protein:protein (PPI) and protein:nucleic acid

interactions [53,58]. Particularly relevant, NQO1 drives PPI with over 20 partners, including key factors in carcinogenesis such as p53, p73α and HIF-1α and these interactions have chaperoning roles protecting protein partners towards intracellular degradation [53,61,62]. Furthermore, these PPIs are very sensitive to ligand binding to NQO1: reduction of FAD by NAD(P)H generally enhances the interaction, while Dic binding usually prevents it [53,62]. NQO1 seems to play a dual role in different pathological states [47]: first, NQO1 is overexpressed in cancer cells, enhancing the cellular antioxidant capacity. This has prompted research towards inhibiting NQO1 as a cancer therapy (e.g., Dic effectively inhibits cancer cells growth [63,64]). Second, two common single-nucleotide polymorphisms (SNPs, p.P187S and p.R139W) have been found in cancer lines and associated with increased cancer risk. The molecular consequences of these two SNPs have been characterized in some detail. P187S (c.559 C > T/p.P187S, allele frequency of $2.47 \cdot 10^{-1}$ in human population; gnomAD v.2.1.1.; https://gnomad.broadinstitute.org/gene/ENSG00000181019?dataset=gnomad_r2_1) is strongly associated with cancer development and poor prognosis when it occurs in homocygosis [65]. This variant strongly destabilizes the protein intracellularly due to thermodynamic destabilization of the CTD that enhances its ubiquitination and acts as very efficient site for proteasomal degradation [41,43,44,49,52,66–68]. In addition, P187S strongly destabilizes the FAD binding site in the apo-state (NQO1$_{apo}$), decreasing by 10–40 fold the affinity for the flavin cofactor [43–45,49]. R139W (c.465 C > T/p.R139W, allele frequency of $3.11 \cdot 10^{-2}$, gnomAD v.2.1.1) is associated with increased risk of developing acute lymphoblastic leukemia in children [69]. This SNP causes minimal perturbations in the function and stability of the NQO1 protein [40,41,70] but promotes aberrant RNA processing, leading to exon 4 skipping that results in an unstable and inactive NQO1 protein [71]. In addition, the consequences of a rare mutation found in a cancer cell line (COSMIC database, c.718 A > C/p.K240Q) have been characterized in some detail, showing quite mild phenotypic consequences [19,39]. By March 2020, there were 125 missense variants in the gnomAD v.2.1.1 database, whereas in the COSMIC database, there were 41 missense mutations compiled. Interestingly, only five mutations were found in both databases. Therefore, it is striking that so little is known about the functional implications of the genetic diversity of NQO1 in the human population.

In this work, we have investigated the effects of eight rare mutations on the stability and functional features of the NQO1 enzyme using biophysical, structural and bioinformatic approaches. Five mutations were found through high-throughput whole-genome/exome sequencing initiatives (gnomAD) and not particularly associated with cancer development and three are derived from cancer cell lines (COSMIC) (Table 1). These mutations affect residues located in the N-terminal region of NQO1 and show different degrees of evolutionary conservation and expected impact on NQO1 stability and function. L7, T16 and Y20 are *fully* conserved among mammalian species (Figure 1) while V9 and A29 are highly conserved and G3 is not. Actually, the mutation A29T is found in some mammalian sequences while G3 is not even the preferred (*consensus*) residue. We may only consider the mutation V9I as *conservative*, since the rest of mutations will cause large changes in either backbone conformation/flexibility or side-chain size, polarity or charge. Structurally, residues affected by these mutations are located in the most N-terminal loop (L1, Gly3), a fully buried short β-sheet (β1, Leu7 and Val9), a solvent-exposed loop (L2, Thr16) and a long α-helix (α1, solvent-exposed Tyr20 and buried Ala29) (Figure 1A,B). Initial structure-based assessments for their impact on NQO1 stability support that V9I and T16M have neutral effects ($\Delta\Delta G < -0.5$ kcal·mol^{-1}) and G3S, G3D, Y20N and A29T are destabilizing ($\Delta\Delta G \sim -1$–1.5 kcal·mol^{-1}), while L7P and L7R are very destabiling ($\Delta\Delta G \geq -2$ kcal·mol^{-1}) (note that these algorithms do not consider the effect of bound ligands in the structure).
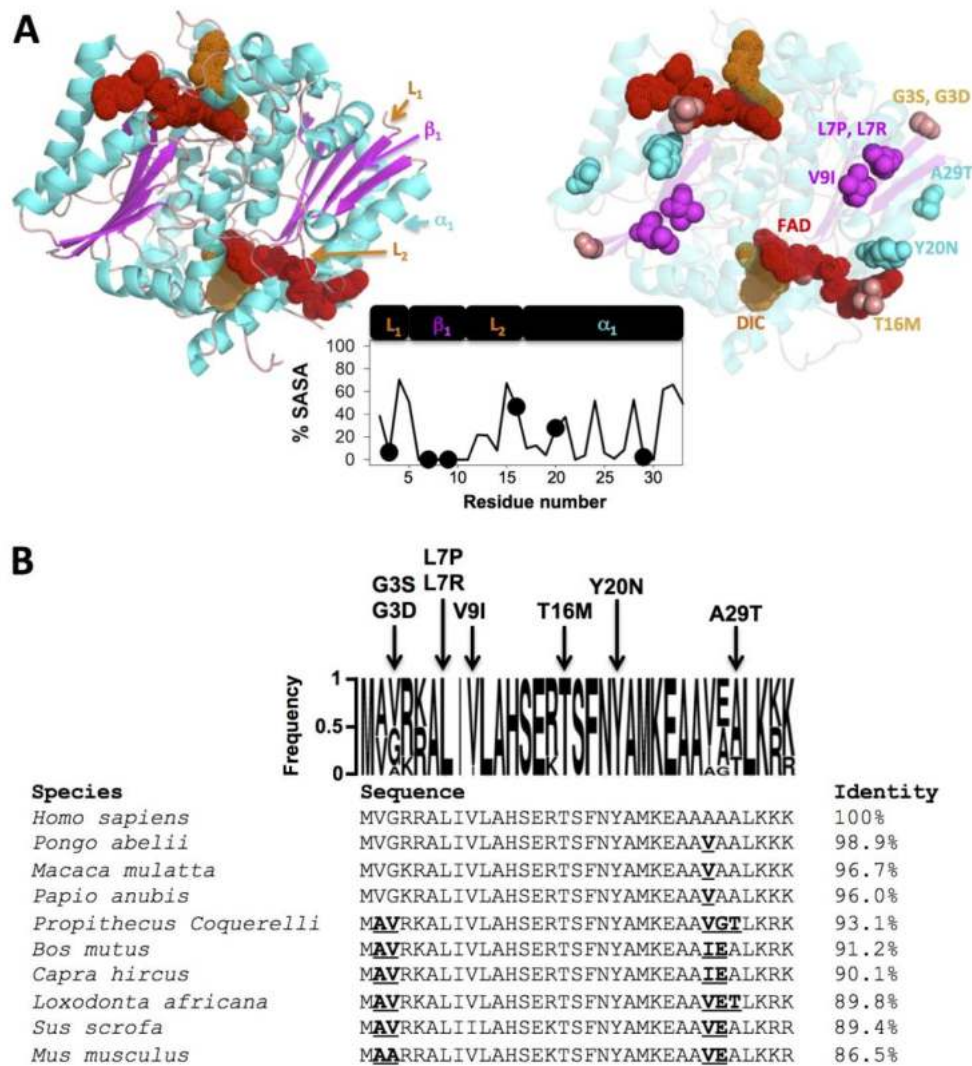
Figure 1. Structural location and sequence conservation of the NQO1 mutations studied in this work. (**A**) Structural representation on the dimeric structure of NQO1 (PDB code 2F1O; [72]). The left panel shows the secondary structure elements to which mutated residues belong (L_1, residues 1–4; β_1, residues 5–10; L_2, residues 11–16 and α_1, residues 17–33). The right panel shows the location of the mutated residues as well as the FAD and dicoumarol (Dic) molecules. Residues are colored according to the secondary structure. The plot in the middle shows the solvent accessibility (% SASA) of this region (determined using GETAREA on the PDB code 2F1O as the average of the two monomers in the dimer). (**B**) Sequence alignment of 10 selected NQO1 mammalian proteins. Residues in bold-underlined indicate non-conservative mutations vs. the human sequence (note that K-to-R or V-to-I are considered as *conservative*). The frequency plot over the NQO1 alignment (generated using WebLogo, https://weblogo.berkeley.edu/) also shows the identity of the missense mutations investigated in this work.

Table 1. NQO1 mutations experimentally characterized in this work.

| Trivial Name | Amino Acid Change | Nucleotide Change | COSMIC | gnomAD V.2.1.1 | | | $\Delta\Delta G$ (kcal·mol^{-1}) [1] |
				All Samples (AS)	Non-Cancer (NC)	Ratio (AS:NC)	
G3S	p.G3S	c.7G > A	-	1.2×10^{-5}	1.3×10^{-5}	0.9	-1.5 ± 1.0
G3D	p.G3D	c.8G > A	+	-	-	-	-1.4 ± 0.9
L7P	p.L7P	c.20T > C	+	-	-	-	-4.3 ± 1.2
L7R	p.L7R	c.20T > G	-	7.8×10^{-5}	7.5×10^{-5}	1.0	-2.1 ± 0.8
V9I	p.V9I	c.25G > A	-	3.9×10^{-5}	4.1×10^{-5}	0.9	-0.4 ± 0.6
T16M	p.T16M	c.47C > T	-	2.8×10^{-5}	2.6×10^{-5}	1.1	-0.1 ± 0.2
Y20N	p.Y20N	c.58T > A	-	2.1×10^{-5}	1.9×10^{-5}	1.1	-1.1 ± 0.6
A29T	p.A29T	c.85G > A	+	-	-	-	-1.3 ± 1.3

[1] $\Delta\Delta G$ is the average ±s.d. from the stability effects provided by Site Directed Mutator (SDM), PopMuSiC, DynaMut, MAESTRO and CUPSAT (see Supplementary Materials Table S1).

2. Materials and Methods

2.1. Protein Expression and Purification

Mutations were introduced by site-directed mutagenesis on wild-type (WT) NQO1 cDNA cloned into the pET-15b vector (pET-15b-NQO1) by GenScript (Leiden, The Netherlands). Codons were optimized for expression in *Escherichia coli* and mutagenesis was confirmed by sequencing the entire cDNA. pET-15b-NQO1 plasmids were used to transform BL21(DE3) cells for protein expression. Typically, 40 mL of Luria–Bertani (LB) medium containing 0.1 mg·mL^{-1} ampicillin was inoculated with transformed cells and grown for 16 h at 37 °C. These cultures were diluted into 800 mL of LB containing 0.1 mg·mL^{-1} ampicillin (Canvax Biotech), grown at 37 °C for 3 h to reach an optical density of about 1.0 and then these were transferred to 25 °C and induced with 0.5 mM isopropyl β-d-1-thiogalactopyranoside (IPTG, Canvax Biotech). After 6 h, cells were harvested by centrifugation and frozen at −80 °C for 16 h. Cells were resuspended in binding buffer, BB (20 mM Na-phosphate, 300 mM NaCl, 50 mM imidazole, pH 7.4) plus 1 mM phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich) and sonicated in an ice bath. These *total extracts* were centrifuged (2000× *g*, 30 min, 4 °C) and the supernatants (*soluble extracts*) were loaded into immobilized-metal affinity chromatography (IMAC) columns (Ni-Sepharose, GE Healthcare), washed with 30–40 volumes of BB and eluted in this buffer containing 500 mM imidazole. These eluates were immediately buffer exchanged using PD-10 columns to the storage buffer (50 mM HEPES-KOH pH 7.4) and concentrated to 1 mL (*purified NQO1 proteins*). After centrifugation for 20 min at 21,000× *g* at 4 °C, the amount of protein (in NQO1 subunit) and the FAD content were determined from their UV-visible spectra using: $\varepsilon_{(280)} = 47{,}900$ M^{-1}·cm^{-1} for NQO1 and $\varepsilon_{(450)} = 11{,}300$ M^{-1}·cm^{-1} for FAD [43]. Spectra were collected in a Cary 50 spectrophotometer (Agilent) using 0.3 cm path-length quartz cuvettes. The amount of FAD per NQO1 monomer was determined from the ratio between the absorbance at 450 nm from normalized spectra (in M^{-1}·cm^{-1}) and the extinction coefficient of FAD (thus assuming that the FAD spectral properties bound to NQO1 resemble those of the free flavin). The lack of significant light scattering (due protein aggregation) was routinely checked. Each NQO1 variant was expressed and analyzed at least three times. The purity of purified NQO1 proteins was checked by SDS-PAGE in 12% acrylamide gels.

The amount of NQO1 in total and soluble extracts was determined by Western-blot. Samples were resolved in 12% acrylamide SDS-PAGE gels and electrotransfered to polyvinylidene difluoride (PVDF) membranes (GE Healthcare) using standard procedures. Immunoblotting was carried out using a primary antibody against NQO1 (F8, sc-393736, at 1:500 dilution) and, as secondary antibody, an anti-mouse IgGκ BP-HRP antibody (sc-516102, at 1:2000 dilution) (both antibodies were purchased from Santa Cruz Biotechnology). Samples were visualized using luminol-based enhanced

chemiluminiscence (from BioRad Laboratories). Densitometric analyses were performed using Image Lab (BioRad Laboratories) and ImageJ (http://rsb.info.nih.gov/ij/).

To obtain apo-proteins, FAD was removed by treatment with 2 M KBr and 2 M urea, 1 mM β-mercapto-ethanol in BB and PMSF 1 mM, as previously described [43,44]. Apo-proteins were buffer exchanged to 50 mM HEPES-KOH pH 7.4. All purified proteins were stored at −80 °C upon flash-freezing in liquid nitrogen.

2.2. In Vitro Characterization of Purified Proteins

For thermal denaturation experiments, purified NQO1 proteins were prepared at a 2 μM final concentration in monomer in the presence of a 10-fold excess of FAD in 50 mM HEPES-KOH, pH 7.4. Triplicate samples were loaded into 3 × 3 mm path-length quartz cuvettes. Thermal denaturation was carried out in a Cary Eclipse (Varian) spectrofluorimeter equipped with a Peltier element. The instrument was calibrated to correct the operational temperature provided by the instrument and the real one using a thermocouple. Samples were equilibrated at 20 °C for 3 min and temperature was increased up to 70 °C at a 2 °C·min^{-1} scan rate. Fluorescence emission was recorded at 350 nm (slit 10 nm) upon excitation at 280 nm (slit 5 nm). Experimental curves were normalized using pre- and post-transition linear baselines to yield the apparent half-denaturation temperatures (T_m). For each variant, triplicate measurements were carried out and analyzed using two different protein batches. Thus, data are presented as mean ± s.d. from six experiments for each variant.

For partial proteolysis experiments, thermolysin from *Geobacillus stearothermophilus* (Sigma-Aldrich, Madrid, Spain) was prepared in 50 mM HEPES-KOH, pH 7.4, with 100 mM CaCl$_2$. The concentration of thermolysin stock solutions was determined spectrophotometrically using ε_{280} = 66,086 M^{-1}·cm^{-1} and small aliquots were stored at ~ 20 μM until used. For experiments, purified NQO1 proteins were prepared at ~10 μM NQO1 with 100 μM FAD in 50 mM HEPES-KOH, pH 7.4 (all concentrations were final concentrations), incubated for 5 min at 25 °C and reactions were initiated by addition of thermolysin (to 0.2–1.2 μM, final concentration of 10 mM CaCl$_2$). Aliquots were withdrawn at different times, mixed with 25 mM ethylenediaminetetraacetic acid (EDTA), pH 8, and denatured at 95 °C with Laemmli buffer. Controls without thermolysin were prepared similarly and considered as samples for time zero. Samples were resolved in SDS-PAGE gels (12% acrylamide) and analyzed by densitometry using ImageJ. The decay of the full-length protein was fitted using a single exponential function to provide the first-order kinetic constant k_{obs}, whereas the second-order rate constant k_{prot} was determined by calculating the slope of the linear fit of k_{obs} vs. thermolysin concentration. Changes in local stability of the primary cleavage site upon mutation $\Delta\Delta G_{prot(WT-mut)}$ were determined by the following equation [73]:

$$\Delta\Delta G_{prot(WT-mut)} \;=\; R{\cdot}T{\cdot}ln\,\frac{k_{prot(mut)}}{k_{prot(WT)}}$$

where R is the ideal gas constant, T is the absolute temperature and $k_{prot(mut)}$ and $k_{prot(WT)}$ are the second-order rate constants for the mutant and the wild-type protein, respectively. Errors in $\Delta\Delta G_{prot(WT-mut)}$ are those determined by linear propagation of the fitting errors for the k_{prot} values.

Fluorescence titrations were carried out at 25 °C using 1 × 0.3 cm path-length cuvettes in a Cary Eclipse spectrofluorimeter (Agilent Technologies, Madrid, Spain). Experiments were carried out in 20 mM K-phosphate, pH 7.4, to decrease the binding affinity of NQO1 for FAD and to allow more accurate characterization of the affinity in tightly binding variants [42]. Then, 20 μL of a 12.5 μM NQO1 stock solution (in subunit) was mixed with 0–500 μL of FAD 10 μM and the corresponding volume of buffer was added to yield a 1 mL final volume. Samples were incubated at 25 °C in the dark for at least 10 min before measurements. Fluorescence spectra were acquired in the 340–360 nm range upon excitation at 280 nm (slits 5 nm), and spectra were averaged over 10 scans registered at a scan rate of 200 nm·min^{-1}. Blanks were also measured similarly (containing only buffer) and subtracted.

The intensity of the fluorescence at 350 nm (I) vs. total concentration of FAD ([FAD]) was used to determine the apparent dissociation constant $K_{d(FAD)}$ using the following equation:

$$I \;=\; I_{apo} + \left(I_{holo} - I_{apo}\right)\cdot\left(\frac{\left[NQO1_{apo}\right] + [FAD] + K_{d(FAD)} - \sqrt{\left([NQO1] + [FAD] + K_{d(FAD)}\right)^2 - 4\cdot[NQO1]\cdot[FAD]}}{2\cdot[NQO1]}\right)$$

where I_{holo} and I_{apo} are the fluorescence intensity of NQO1$_{holo}$ and NQO1$_{apo}$, respectively, and [NQO1] is the total protein concentration (250 nM). Data from at least two independent titrations using different preparations of apo-proteins for each NQO1 variant were used in the fitting. The apparent change in binding free energy, $\Delta G_{(FAD)}$, was determined as follows: $\Delta G_{(FAD)} = R \cdot T \cdot \ln K_{d(FAD)}$. The error associated with $\Delta G_{(FAD)}$ was determined by linear propagation from the errors of $K_{d(FAD)}$.

Titrations of NQO1$_{holo}$ proteins with dicoumarol (Dic) were carried out by isothermal titration calorimetry (ITC) in an ITC$_{200}$ microcalorimeter (Malvern, Iesmat, Alcobendas, Spain). Purified NQO1 samples were prepared in 50 mM HEPES-KOH, pH 7.4, in the presence of 100 μM FAD and loaded into the calorimetric cell. Dic solutions were prepared from 10 mM stocks (in 100 mM NaOH) and diluted into 50 mM HEPES-KOH pH 7.4 to a final concentration in Dic of 120 μM (and 100 μM FAD) and loaded in the titrating syringe. Experiments were typically carried out by performing an initial injection of 0.5 μL followed by 20–22 injections of 1.75 μL, spaced by 100 s. The instrument operated in the high-feedback mode. Data analysis was carried out upon manual integration of the experimental enthalpograms and fittings were done using a single type of independent binding sites model with the software provided by the manufacturer. Dilution heats were included as a fitting parameter. This analysis yields the number of bindings per NQO1 monomer (N) as well as the association binding constant (K_a, being $K_a = 1/K_d$), enthalpy (ΔH) and entropy (ΔS). Experiments were performed at least three times at 25 °C and using two different protein preparations of each variant. Data at this temperature are presented as mean ± s.d. Experiments at 10, 15, 17.5 and 20 °C were performed only once and reported errors were those from fittings. The apparent change in heat capacity (ΔC_p) was determined from linear dependence of ΔH on temperature.

Structure–energetic relationships were combined with Dic binding calorimetric data to determine the magnitude of the conformational change associated with binding (N$_{conf}$, in number of residues) using previously described procedures [39]. Briefly, two experimental apparent thermodynamic binding variables, ΔH and ΔC_p, were assumed to be the sum of two separate contributions: one arising from the energetics of rigid-body interactions between *NQO1$_{holo}$* and Dic (*intrinsic binding*) and the other arising from the *conformational change* associated to binding. The former can be estimated from changes in solvent accesible surface upon Dic binding as determined from the crystal structure of the NQO1$_{holo}$ complex with Dic (NQO1$_{dic}$; PDB code 2F1O, [72]), yielding values of −5.8 kcal·mol^{-1} and −0.14 kcal·mol^{-1}·K^{-1} for intrinsic ΔH (at 25 °C) and ΔC_p, respectively. Thus, the contribution from the conformational change (ΔH_{conf} and $\Delta C_{p,conf}$) can be obtained from the difference between the experimental variables and those calculated for intrinsic binding. This contribution can be parametrized in terms of the number of residues folded upon binding (i.e., those involved in the conformational change) from well-known structure–energetics relationships for protein folding as follows [39,74]:

$$\Delta H_{conf} = 0.215 \cdot N_{conf}$$

$$\Delta C_{p,conf} = 0.0138 \cdot N_{conf}$$

where ΔH_{conf} is given in kcal·mol^{-1} and $\Delta C_{p,conf}$ in kcal·mol^{-1}·K^{-1}. This approach provides two ways of determining N$_{conf}$ from experimental and structural variables.

2.3. In Silico Mutagenesis and Structural Analysis

In silico mutagenesis of the dimeric NQO1 protein was performed using Rosetta [75]. The atomic coordinate files for the NQO1$_{holo}$ state (PDB code 1D4A) and NQO1$_{dic}$ state (PDB code 2F1O) of the

WT protein were processed to remove non-canonical amino acids and water molecules. Subsequently, each PDB file was refined with fast relax constrained to native coordinates using Cartesian-space refinement as described elsewhere [76] and the Rosetta energy function REF15. After relaxation and refinement, the structural model with the lowest energy was used as the starting structure for mutagenesis. For *in silico* mutagenesis, the cartesian version of Rosetta's ddG protocol [77] was followed and five structures were generated for each variant (WT and eight mutants), both in the $NQO1_{holo}$ and $NQO1_{dic}$ states, following the procedures described in [76]. Finally, the energies (expressed in $kcal·mol^{-1}$) of every structure were calculated and averaged for each variant in a given ligation state. $\Delta\Delta G$ values were calculated by subtracting the average energy of the corresponding mutant ensemble to the average energy of the WT ensemble. As cut-off, $\Delta\Delta G$ values > 0.5 $kcal·mol^{-1}$ were considered as *stabilizing*, $\Delta\Delta G$ values < −0.5 $kcal·mol^{-1}$ were considered as *destabilizing* and $\Delta\Delta G$ values between −0.5 and +0.5 $kcal·mol^{-1}$ were considered as *neutral*. Structural analysis and representations were generated with YASARA molecular modeling software [78].

2.4. Structure- and Sequence-Based Analysis of Mutational Effects on Protein Stability and Potential Pathogenicity

To determine changes in protein (thermodynamic stability), we used the following methods: (i) SDM2 (Site Directed Mutator 2, http://marid.bioc.cam.ac.uk/sdm2/prediction) [79] analyses were carried out using chains A and B from three high-resolution crystal structures (PDB codes 2F1O, 1D4A and 5FUQ, [43,72,80]). For each mutant, the value provided was the mean ± s.d. from six analyses; (ii) PoPMuSiC (https://soft.dezyme.com/) [81] analyses were carried out using all chains from three high-resolution crystal structures (PDB codes 2F1O, 1D4A and 5FUQ, [43,72,80]). For each mutant, the value provided was the mean ± s.d. from three analyses; (iii) DynaMut (http://biosig. unimelb.edu.au/dynamut/) [82] analyses were carried out in chains A and C (PDB codes 2F1O and 1D4A, [72,80]) or chains A and B (PDB code 5FUQ, [43]). For each mutant, the value provided was the mean ± s.d. from six analyses; (iv) MAESTRO (https://pbwww.che.sbg.ac.at/) [83,84] analyses were carried out in chains A and C (PDB codes 2F1O and 1D4A, [72,80]) or chains A and B (PDB code 5FUQ, [43]). For each mutant, the value provided was the mean ± s.d. from six analyses; (v) CUPSAT (http://cupsat.tu-bs.de/) [85] analyses were carried in the *Thermal denaturation method* using chains A and B from three high-resolution crystal structures (PDB codes 2F1O, 1D4A and 5FUQ, [43,72,80]). For each mutant, the value provided was the mean ± s.d. from six analyses.

For bioinformatic prediction of pathogenicity, we used: (i) PolyPhen-2 (http://genetics.bwh. harvard.edu/pph2/index.shtml) analyses, which yielded the potential effect of mutations as *Benign*, *Possibly damaging* or *Probably damaging*, using information on multiple sequence alignments and simple physical estimations of mutational effects [32]; (ii) SIFT (https://sift.bii.a-star.edu.sg/) analyses, which used multiple sequence alignment tools to estimate evolutionary tolerance of mutations (*Tolerated* vs. *Not tolerated*) [86]; (iii) PROVEAN (http://provean.jcvi.org) analyses, which used a sequence-based (multiple sequence alignment) approach to predict functional effects of mutations (*Neutral* vs. *Deleterious* effect) [87]; (iv) PON-P2 (http://structure.bmc.lu.se/PON-P2) analyses, which provided mutational effects as *Pathogenic*, *Neutral* or *Unknown* using evolutionary sequence conservation, properties of amino acids and physical and functional annotations of replacements sites [88]; (v) MutationTaster (http:// www.mutationtaster.org) analyses, which yielded mutational effects as *Polymorphisms* or *Disease-causing* using DNA sequence alterations and information from human genome variability, evolutionary conservation, splice-site changes, loss of protein features and changes that might affect the amount of mRNA [89]. For methods i–iv, the input was the NQO1 protein sequence, and for method v, the NQO1 cDNA sequence.

3. Results and Discussion

3.1. Expression Analysis of NQO1 Variants Reveals Dramatic Effects of the Mutations L7R and L7P on Protein Stability and/or Solubility

To characterize the effects of rare NQO1 mutations on the stability and functional features of this protein, we expressed them in *E. coli* and purified the WT and mutant proteins. Two of the mutants (L7R and L7P) showed largely reduced expression levels as soluble protein and were not amenable for purification, supporting that these two mutations severely perturb NQO1 folding and prevent the formation of stably folded dimers (Figure 2A,B). Accordingly, structure-based analysis of protein destabilization supported that these two mutations should be the most deleterious (Table 1). Previous expression analyses have also shown similar behavior for largely disrupting mutations at the P187 site (a residue buried in the structure and close to the MMI), such as P187R, P187E and P187L [19,39]. The remaining six mutants were amenable for purification and showed comparable yields to those of WT NQO1 (Figure 2B,C).
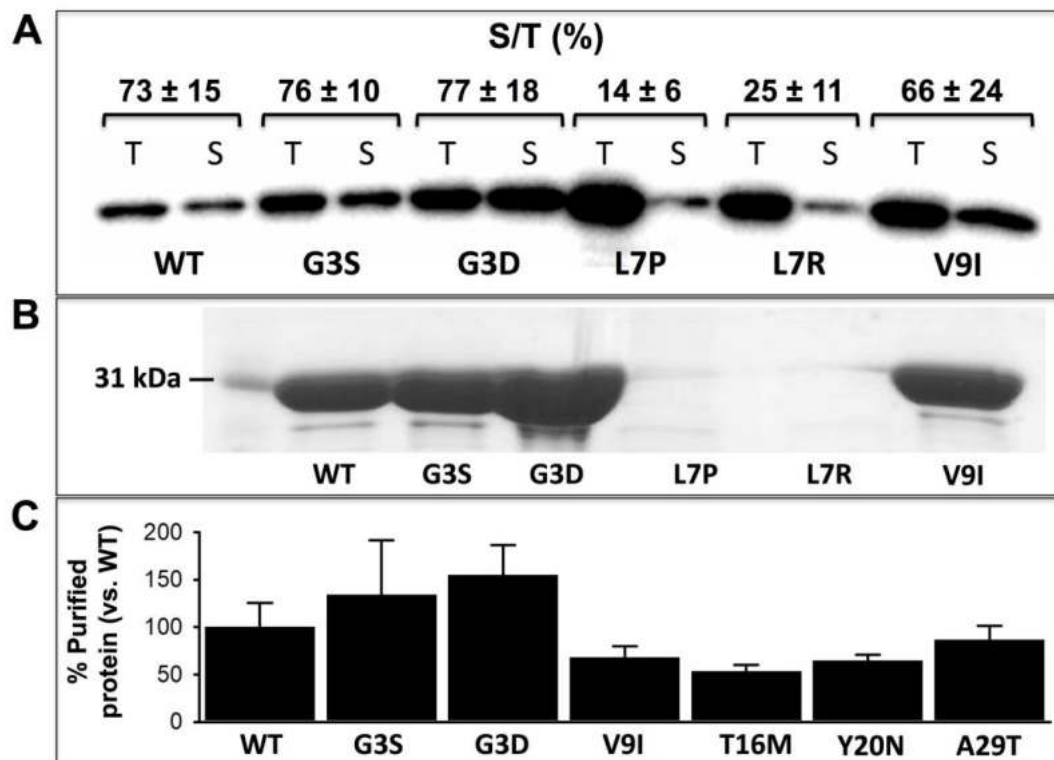
Figure 2. Expression and solubility of NQO1 variants in *E. coli* cells. (**A**) Western-blot analysis of NQO1 content upon expression in *E. coli* cells. Cells were sonicated to obtain total extracts (T), while soluble extracts (S) were obtained upon centrifugation at $21,000\times g$ for 20 min at 4 °C. Samples were denatured with Laemmli buffer and submitted for Western-blot analysis using anti-NQO1 antibody, F-8 antibody (Santa Cruz Biotechnology). The Western-blot is representative from three different purifications. The fraction of the total protein found in the soluble extract upon densitometric analysis (S/T, as %) is indicated (as mean ± s.d. from three independent experiments). (**B**) Purified NQO1 proteins from three different purifications were concentrated ~10-fold and samples analyzed by SDS-PAGE. Note that during purification and storage, the remaining soluble protein of L7P and L7R was negligible. (**C**) Yield in NQO1 protein variants after immobilized-metal affinity chromatography (IMAC) purification. Data were the mean ± s.d. from 3–4 different purifications for each NQO1 variant. Wild-type (WT) levels were 1.45 ± 0.38 mg·L^{-1} of culture and used to normalize yields.

3.2. Thermal Stability Analyses Revealed Significant Perturbation of the MMI by Mutations T16M and Y20N

For the six mutants amenable for purification, we analyzed and compared their stability with that of the WT protein in the presence of an excess of FAD (NQO1$_{holo}$) and using fluorescence-monitored thermal denaturation (Figure 3). To some extent, the thermal stability of NQO1 variants allows to capture the perturbation caused by the mutations on the MMI since the dimer dissociates prior to the rate-limiting step of the irreversible denaturation [15,39,41]. T16 and Y20 are within or in very close proximity to the MMI and mutations at these residues are potentially the most damaging for dimer stability (Figure 3A). Consequently, the mutations T16M and Y20N, which may introduce mild to moderate structural perturbations (Table 1), led to significant changes in thermal stability (i.e., their T$_m$ were lower than that of the WT protein by 4–5 °C) (Figure 3B,C). L7 and V9 are at the shortest distance of 8 Å from the MMI (Figure 3A). Accordingly, the severely perturbing L7P and L7R mutation may significantly destabilize the MMI, thus explaining the dramatic effects of these mutations that prevent dimer formation or severely affect its stability (Figure 2). The mild effects of the mutation V9I likely reflect a small structural perturbation in some proximity to the MMI (reducing thermal stability by 1.5 °C; Figure 3B,C). G3 and A29 are over 15 Å away from the MMI (Figure 3A), and consequently, the effects of the mutations G3S, G3D and A29T in thermal stability were marginal (Figure 3B,C).

Residue	MMI (Å)
Gly3	16.9
Leu7	8.0
Val9	7.2
Thr16	0
Tyr20	2.9
Ala29	15.0
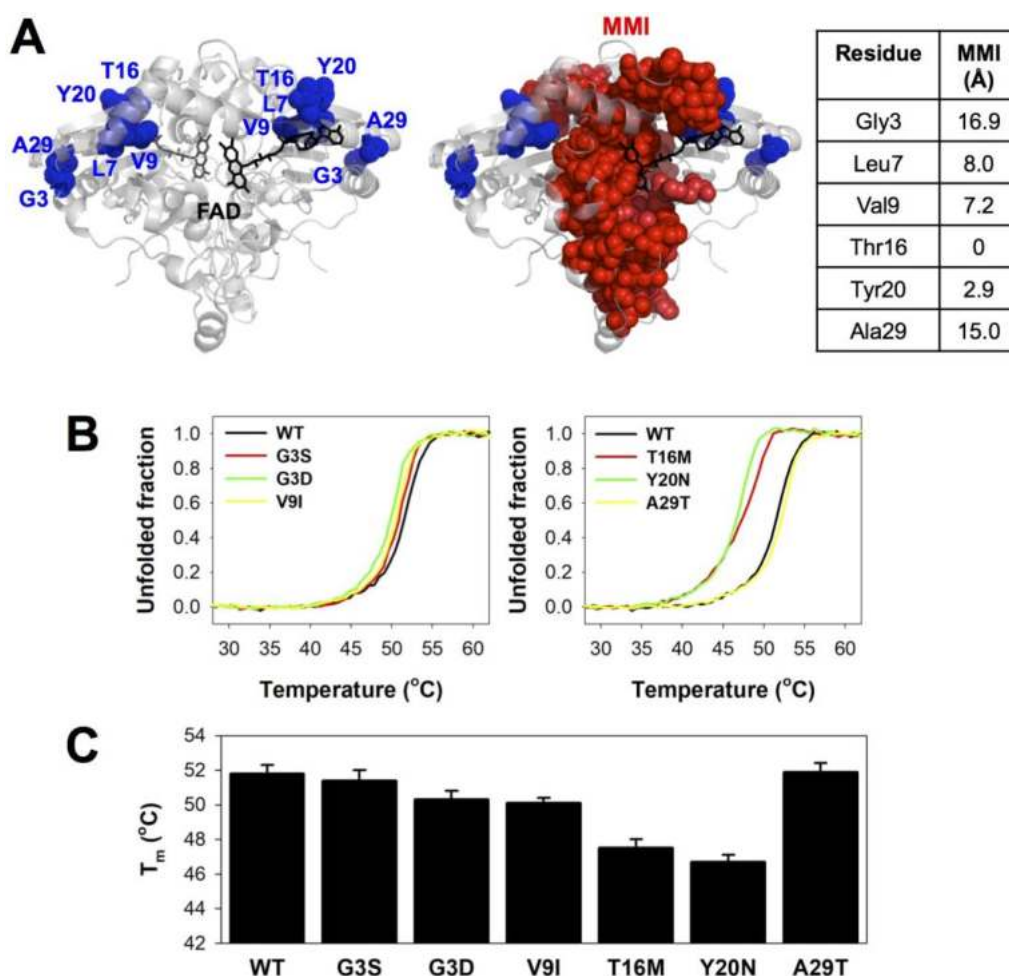
Figure 3. Thermal stability of NQO1 variants. (**A**) Structural location of the mutated residues, showing their proximity (as a minimal distance) to the monomer:monomer interface (MMI). The structure used for display has the PDB code 2F1O [72]. Residues belonging to the MMI were identified as described [90]. (**B**) Thermal denaturation profiles; (**C**) T$_m$ values (mean ± s.d. from six replicas, using proteins from two purifications).

It is important to note that the above-mentioned interpretation of mutational effects on NQO1 thermal stability simply implies that both the magnitude of the perturbation caused by the mutation and its distance-dependent propagation to the MMI are key features. Previous systematic mutational studies at the P187 and K240 residues of NQO1 generally supported this interpretation [39]. In the next sections, we will apply a similar approach to other features of NQO1, such as the mutational effects on local stability and binding of functional ligands.

3.3. The Local Stability of the Thermolysin Cleavage Site (TCS) Is Reduced by the Distant Mutations T16M and Y20N

Limited proteolysis of NQO1 is a sensitive method to probe the local stability at different regions of the protein [41,50]. In particular, proteolysis by thermolysin typically provides information on the local conformational stability of the region surrounding the primary cleavage site for this protease (i.e., the thermolysin cleavage site (TCS)) located in the NTD of $NQO1_{holo}$; it cleaves between S72-V73 [41]. For instance, the phosphomimetic mutation S82D locally destabilizes this region, accelerating 30-fold proteolysis by thermolysin of $NQO1_{holo}$ (i.e., it causes 2.0 kcal·mol^{-1} of local destabilization in terms of ΔG_{prot} as the mutational effect on the unfolding free energy between the native and cleavable states) [91]. It must be noted that all mutations studied in this work are located at least 15–20 Å from the TCS (Figure 4A).
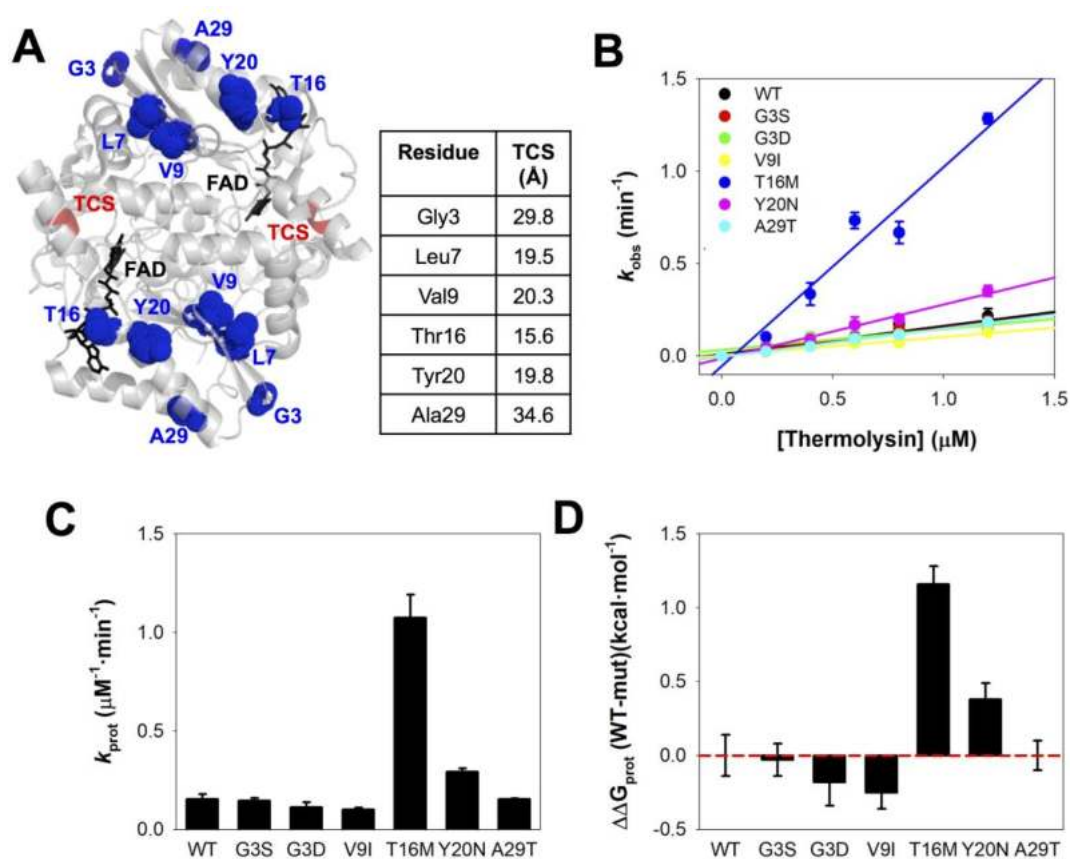
Figure 4. Local stability of the N-terminal domain (NTD) probed by proteolysis with thermolysin. (**A**) Location of mutated residues regarding the thermolysin cleavage site (TCS). The minimal distance between the mutated residue and the TCS backbone (Ser72-Val73) is also indicated (using the structure with PDB code 2F1O). (**B**) Linear dependence of apparent rate constants (k_{obs}) on protease concentration. The slopes provide the values of the second-order rate constants k_{prot}. (**C,D**) Values of k_{prot} for the NQO1 variants (**C**) and the effect of mutations on the local stability (**D**) as $\Delta\Delta G_{prot}$. Note that a positive value of $\Delta\Delta G_{prot}$ indicates a destabilizing effect.

Proteolytic patterns of all mutant NQO1$_{holo}$ proteins studied resembled that of WT NQO1 (Figure S1), supporting initial cleavage between S72-V73. Furthermore, proteolysis rate constants (k_{obs}) for all NQO1 variants showed a linear dependence on protease concentrations (Figure 4B and Figure S2), thus implying that changes on the second-order rate constant k_{prot} are related to effects on the thermodynamic stability between the native and cleavable state in the TCS (Figure 4C,D) [41]. We found that only the mutations T16M and Y20N caused local destabilization beyond the experimental error (by ~1.1 and 0.4 kcal·mol^{-1}, respectively; Figure 4D and Table S2). Thus, the structural perturbation introduced by these two mutations is *sensed* by the TCS, located 15–20Å away (Figure 4A).

3.4. The Mutations T16M, Y20N and A29T Perturb FAD Binding

The binding of FAD to NQO1 is very sensitive to structural perturbations caused by mutations in the NQO1$_{apo}$ and NQO1$_{holo}$ states [39–45,91]. Actually, when a single mutation reduces significantly FAD binding affinity (by at least 5-fold, i.e., about 1 kcal·mol^{-1} in binding free energy), this often decreases the FAD content of the protein upon purification from *E. coli* cell cultures [39,40,43,44,91]. Since bound FAD is mandatory for NQO1 catalytic function, these effects would affect the specific activity of the enzyme. Structurally, residues T16 and Y20 are close to the FAD molecule (≤5 Å), whereas V9 and A29 are at 12–13 Å and G3 is more than 25 Å away (Figure 5A).

The FAD content in purified NQO1 proteins was determined using near-UV/visible absorption spectroscopy. All the mutants showed high levels of FAD bound, with fractions in the range 0.8–0.9 (similar to those of the WT protein), with the only exception of T16M that showed reduced content (0.5 mol FAD/NQO1 monomer, Figure 5B,C). To evaluate quantitatively mutational effects on FAD binding affinity, we carried out titrations of apo-proteins with FAD monitored by fluorescence spectroscopy (Figure 5D,E, Figure S3 and Table S3). As expected from its lower content in FAD as-purified, the mutant T16M showed a 12-fold lower affinity that corresponds to a decrease in binding free energy of about 1.5 kcal·mol^{-1}. In addition, the mutants Y20N and A29T showed decreased binding affinity (3- and 6-fold lower than that of WT NQO1, respectively, that correspond to changes in binding free energy of 0.7 and 1.0 kcal·mol^{-1}, respectively) (Figure 5C–E and Table S3). It is interesting to note the effect of the mutation A29T, which involves a residue located at more than 13 Å away from the FAD binding site.

3.5. None of the Mutations Affect Dic Binding Affinity or Energetics

Dic strongly inhibits NQO1 with a K_d of 10–50 nM, acting as a competitive inhibitor of the NADP(H) coenzyme and the substrate [43,48,63,72,92]. Investigating the affinity and energetics of Dic binding to NQO1 mutants may provide insight into mutational effects on several protein traits: (i) Dic (and analogues thereof) inhibits the growth of cancer cell lines, presumably by preventing the high antioxidant activity provided by cancer-associated overexpression of NQO1 [63,93,94]. Consequently, alterations of Dic binding due to missense mutations may yield decreased cellular sensitivity towards this potential pharmacological treatment for cancer; (ii) Missense mutations can decrease the binding affinity for Dic by altering the conformation and energetics of the native state, shifting the conformational equilibrium towards non-competent states for binding. An excellent example of this behavior is observed in the common polymorphic variant P187S that binds Dic with about a 10-fold lower affinity than the WT protein and shows clearly different binding energetics associated with the folding of the CTD upon Dic binding [39,44,92]; (iii) Extensive mutational studies at the P187 and K240 sites have shown that long-range communication (over 15 Å from the binding site) of mutational effects acting on the conformation and stability of the CTD can be accurately characterized by a detailed analysis of Dic binding energetics [19,39]. As we show in Figure 6A, all the mutations investigated in this work are structurally located far from the Dic molecule (by at least 15 Å).
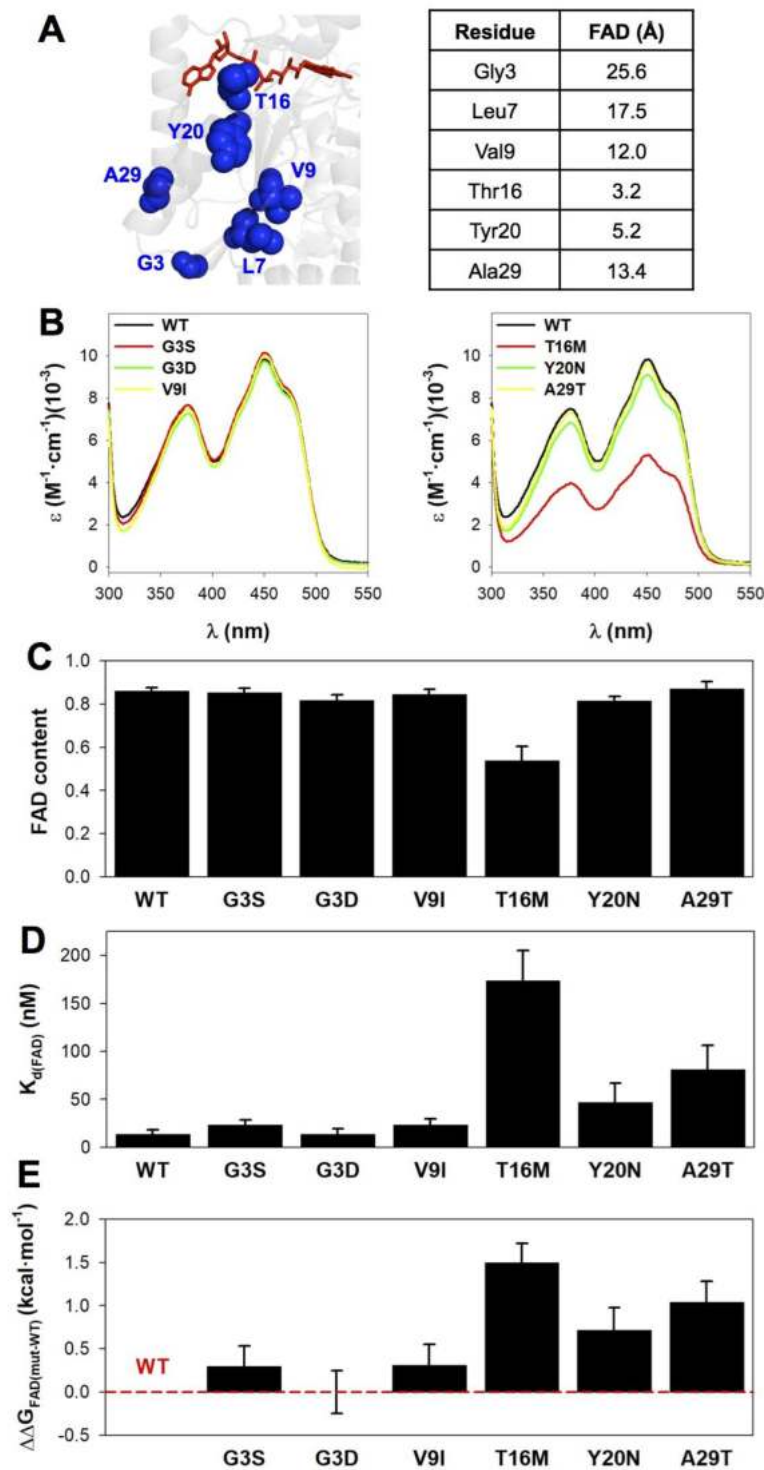
Figure 5. FAD content and binding affinity in purified NQO1 variants. (**A**) Location of mutated residues and minimal distances to the FAD molecule (using the structure with PDB code 2F1O). (**B**) Near-UV/visible absorption spectra of NQO1 variants. (**C**) FAD content (per NQO1 monomer) derived from absorption spectra (considering a $\varepsilon_{450} = 11,300$ M^{-1}·cm^{-1}). (**D**) FAD-binding affinity of NQO1$_{apo}$ proteins determined by fluorescence titrations. (**E**) Difference in binding free energies ($\Delta\Delta G_{FAD}$) between a given mutant and the WT protein. Errors in $\Delta\Delta G_{FAD}$ are those determined from linear propagation. Data in B–C are mean ± s.d. from at least three different purifications, and in D–E, from two different titrations with two different preparations of apo-proteins.
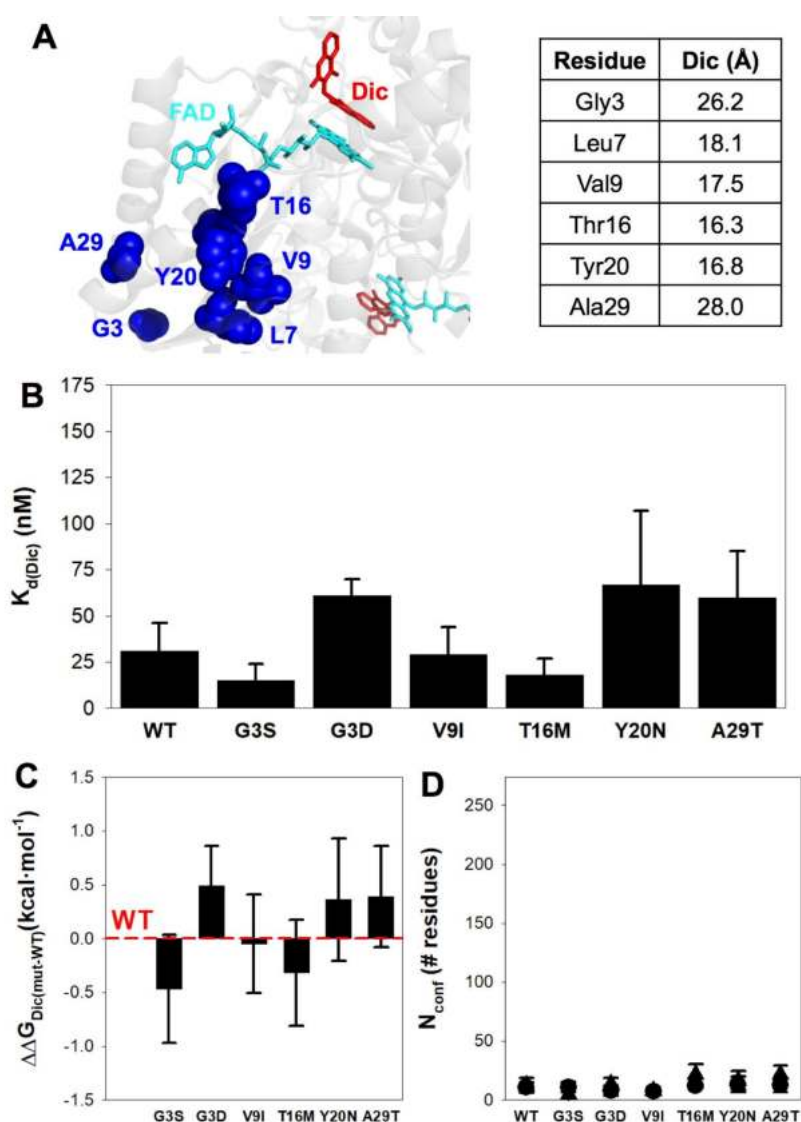
Figure 6. Affinity and structure–energetics analysis for Dic binding to NQO1 variants. (**A**) Location of mutated residues and minimal distances to Dic (using PDB code 2F1O). (**B**) Dissociation binding constants at 25 °C. Data are mean ± s.d. from at least three independent experiments for each variant. (**C**) Difference in binding free energy between a given mutant and the WT protein. Errors are those determined from linear propagation. (**D**) Magnitude of the conformational change (as a number of residues, N_{conf}) determined from experimental binding enthalpies (circles) and changes in heat capacity (triangles).

We have analyzed Dic-binding affinity and energetics by ITC (Figure 6 and Figure S4). All the NQO1 mutants bound Dic with an affinity comparable to that of NQO1 WT and close to the technical limit of the technique. The largest differences correspond to a mild 2-fold change, which implies changes in binding free energy ≤ 0.5 kcal·mol^{-1} (Figure 6B,C and Table S4). The apparent enthalpic and entropic contributions to binding were also very similar (within the experimental error, see Figure S5 and Table S4). In addition, two apparent thermodynamic binding parameters (ΔH and ΔC_p) can be used to evaluate the magnitude of the conformational change associated with Dic binding [39,44] (Figure S5 and Tables S4 and S5). Basically, these two parameters are related with the sum of an intrinsic binding contribution (that can be estimated from the crystal structure of NQO1$_{dic}$) and a second term originating from the conformational change induced upon ligand binding [39]. This second term is straighforwardly converted to the *number of residues involved in the conformational change* using

structure–energetic relationships for protein folding (thus yielding the parameter N_{conf}). Dic binding caused a minimal conformational change upon binding to all variants, with N_{conf} values (average of the two methods) of 10–20 residues (Figure 6D and Table S6).

3.6. Structural Analysis and Energy Calculations Provide Insight into Mutational Effects on Protein Stability and Function

To provide further molecular insight into the effects of this set of eight mutations on different NQO1 traits, we carried out in silico mutagenesis and structural and energetic analysis in both the holo-forms (NQO1$_{holo}$) and Dic bound forms (NQO1$_{dic}$). An overview of the most relevant structural consequences for each mutation are compiled in Figure 7.
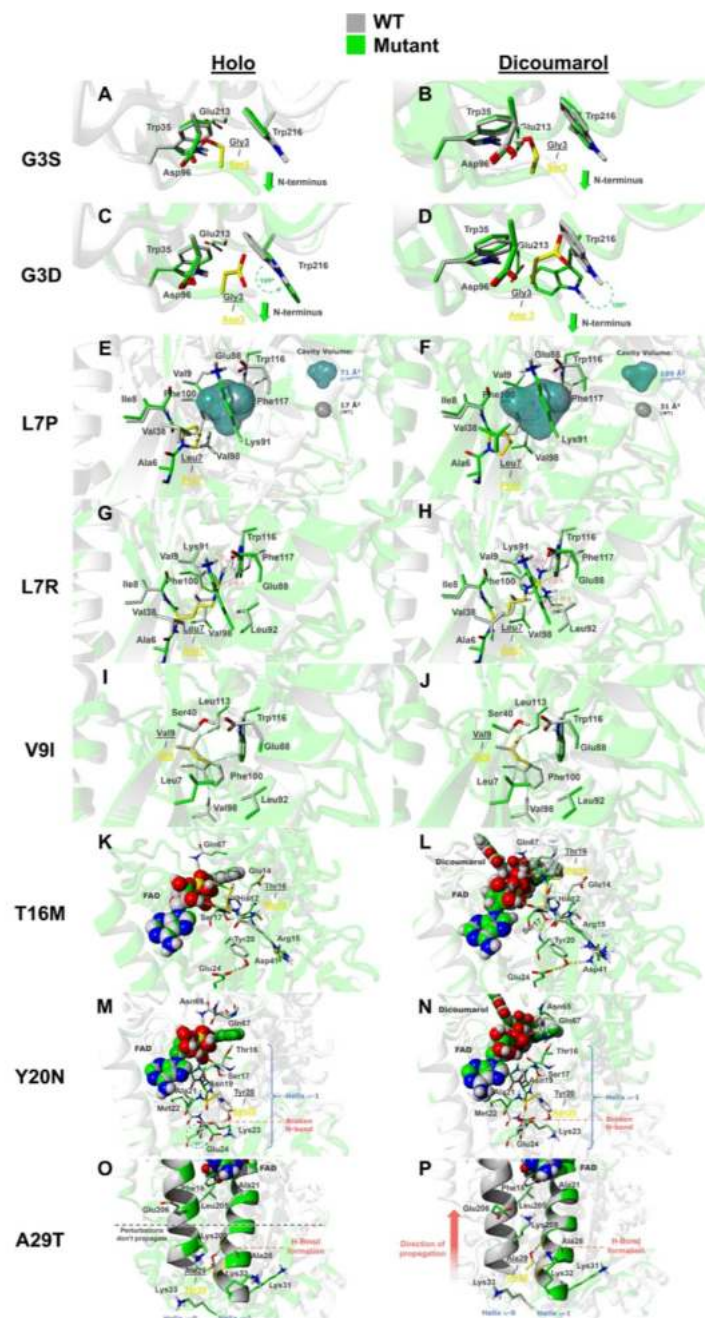
Figure 7. Overview of the mutational effects from in silico mutagenesis. Gray-colored amino acids represent the WT structure while green-colored amino acids represent the mutated models. Mutated

residues are shown as sticks colored in yellow for each case. (**A,B**) NQO1$_{holo}$ (**A**) and NQO1$_{dic}$ (**B**) states with the G3S mutation. Green arrows indicate displacement of the N-terminus on the opposite side of the mutation. (**C,D**) NQO1$_{holo}$ (**C**) and NQO1$_{dic}$ (**D**) states with the G3D mutation. Green dashed arrows indicate side chain rotation. (**E,F**) NQO1$_{holo}$ (**E**) and NQO1$_{dic}$ (**F**) with the L7P mutation. Blue and gray surfaces represent cavities in the protein core for the mutant and WT, respectively. (**G,H**) NQO1$_{holo}$ (**G**) and NQO1$_{dic}$ (**H**) with the L7R mutation. Clashes are shown as blue solid lines marked with inter-atomic distances. (**I,J**) NQO1$_{holo}$ (**I**) and NQO1$_{dic}$ (**J**) with the V9I mutation. Hydrophobic interactions are shown as solid lines between interacting atoms colored gray and blue for the interactions created by V9 (WT) and I9 (V9I mutant) residues, respectively. (**K,L**) NQO1$_{holo}$ (**K**) and NQO1$_{dic}$ (**L**) with the T16M mutation. Ligand atoms are displayed in ball representation. Hydrogen bonds are depicted as yellow dashed cylinders. (**M,N**) NQO1$_{holo}$ (**M**) and NQO1$_{dic}$ (**N**) with the Y20N mutation. (**O,P**) NQO1$_{holo}$ (**O**) and NQO1$_{dic}$ (**P**) with the A29T mutation. Helices α1 and α9 are shown in ribbon representation.

3.6.1. The Mutations G3S and G3D

Our experimental analyses have shown that the mutations G3S and G3D are essentially neutral. In the G3S mutant, modeling shows that S3 is surrounded by an hydrophobic patch formed by W35 and W216 and by two negatively charged amino acids, namely D96 and E213, in both NQO1$_{holo}$ and NQO1$_{dic}$ forms (Figure 7A,B). The mutation G3S promotes local changes in its neighboring amino acids but also displaces the N-terminus of the protein towards the opposite direction of its side-chain, thus helping accommodate the perturbation generated upon the introduction of a bulkier and polar residue. Overall energetic scoring by Rosetta suggested that this mutation has a mild–moderate destabilizing effect on NQO1$_{holo}$, whereas this effect is mild in the NQO1$_{dic}$ state (Table 2). The most relevant contributions to this destabilization in both the NQO1$_{holo}$ and NQO1$_{dic}$ states seem to arise from (Table S6): (i) unfavorable repulsive energy (Δfa_rep), whose increase accounts for an increment in the number of atoms that come into contact upon mutation, which is in agreement with the mutated residue bumping into other residues, such as W35 and W216, thus affecting the repulsive van der Waals energy; (ii) unfavorable isotropic solvation energy (Δfa_sol), which arises from burial of polar residues upon mutation (i.e., primarily that of the side chain of S3 that becomes embedded in the hydrophobic patch located around the mutation site).

Table 2. Mutational effects on conformational stability (as folding free energy between the mutant and WT proteins, ΔΔG, in kcal·mol^{-1}, in the NQO1$_{holo}$ and NQO1$_{dic}$ states, determined by Rosseta). Negative values indicate a destabilizing effect whereas positive values indicate a stabilizing effect.

	G3S	G3D	L7P	L7R	V9I	T16M	Y20N	A29T
NQO1$_{holo}$	−5.0	−11.6	−29.4	−14.1	1.9	−0.8	−8.1	−10.9
NQO1$_{dic}$	−1.5	−10.1	−25.5	−35.9	0.6	0.8	−5.4	0.4

Modeling of the mutation G3D shows larger structural and energetic distortions than those observed for the mutation G3S (Figure 7C,D), predicting moderate–large destabilizing effects in the NQO1$_{holo}$ and NQO1$_{dic}$ states (Table 2). This difference between the mutations G3S and G3D can be easily explained due to nature of the mutations, since G3D introduces a bulkier and charged residue that perturbs the surrounding hydrophobic patch, particularly residues W35 and W216—especially the latter one, which undergoes a 100° rotation around its χ$_1$ angle in the NQO1$_{dic}$ state and 169° around χ$_2$ in the NQO1$_{holo}$ state. This local conformational change generates a significant penalization due to the increase in the repulsive van der Waals energy and unfavorable isotropic solvation energy terms, presumably due to the insertion of a negatively charged residue into the hydrophobic patch formed by W35 and W216 (an effect seen in both NQO1$_{holo}$ and NQO1$_{dic}$ forms) (Figure 7A–D and Table S6). This may explain why the N-terminus is pushed a few Å away from the mutated site

(Figure 7C,D, green arrow shows direction of displacement) in the model of the G3D mutant. This could suggest that the N-terminus of the protein might help to reduce and prevent clashes by potentially accommodating some of the observed perturbations. This also possibly explains the minimal changes in thermal stability and ligand binding observed experimentally, although the mutation is perceived as *destabilizing* by Rosetta.

3.6.2. The Mutations L7P and L7R

Our experimental analyses have indicated that the mutations L7P and L7R are heavily destabilizing (Figure 2). Regarding the mutation L7P, we must note that proline is rarely found in the middle of β-sheets as its presence is disfavored because of the lack of accessible backbone N-H group to participate in hydrogen bonding, thus disrupting the β-sheet structure [95,96]. Energetic scoring by Rosetta analysis supports the tremendous destabilizing effect of the mutation L7P in both $NQO1_{holo}$ and $NQO1_{dic}$ states (Table 2). L7 is surrounded by a hydrophobic environment in the core of the protein, composed by A6, I8, V9, V38, L92, V98, F100 and W116 (Figure 7E,F). Most of the hydrophobic interactions established by L7 in this pocket are essentially abolished in the L7P mutant (in both $NQO1_{holo}$ and $NQO1_{dic}$ states), likely due to the cavity created by the mutation (Figure 7E,F). In the WT protein, there is only a small cavity with a volume of 17 and 31 $Å^3$ ($NQO1_{holo}$ and $NQO1_{dic}$ forms, respectively), whereas the L7P mutation increases the size of this cavity to a volume of 71 and 109 $Å^3$ ($NQO1_{holo}$ and $NQO1_{dic}$ forms, respectively). This cavity in the protein core may be, in part, responsible for the great destabilizing character of the mutation, as reflected in the energetic penalization in the attractive energy (fa_atr) term (Table S6). Interestingly, parametrizations of the conformational destabilization caused by cavities in model proteins have provided values in the range of 24–36 cal·mol^{-1}·$Å^3$ [97,98], thus allowing to calculate that just this cavity (without considering additional perturbations caused) would destabilize the native structure by 1.3–1.9 kcal·mol^{-1} ($NQO1_{holo}$) and 1.9–2.8 kcal·mol^{-1} ($NQO1_{dic}$). Additionally, in both $NQO1_{holo}$ and $NQO1_{dic}$ states, the mutation L7P breaks a backbone hydrogen bond between β1 and β3 parallel strands that destabilizes the β-sheet structure (Figure 8B,E), thus reducing the long-range hydrogen bonds energy term (hbond_lr_bb) (Table S6). The perturbation generated by P7 is accentuated by the distortion in dihedral Φ and Ψ angles caused by the geometry of the mutated residue and the prevention of hydrogen bond formation in the proximity of the mutated site (Figure 8C,F).

The mutation L7R is also highly destabilizing, as calculated by Rosetta (Table 2). This mutation introduces a severely perturbing, positively charged residue in the formerly described well-packed hydrophobic pocket (Figure 7G,H). Unlike L7P, L7R does not generate a cavity (Figure 7E,F) but instead causes steric clashes with neighbouring residues (Figure 7G,H and Table S7) in both $NQO1_{holo}$ and $NQO1_{dic}$ states. In this case, conformational destabilization mainly arises from unfavorable isotropic solvation energy (Δfa_sol) (Table S6), likely due to the burial of a charged residue in a hydrophobic environment.
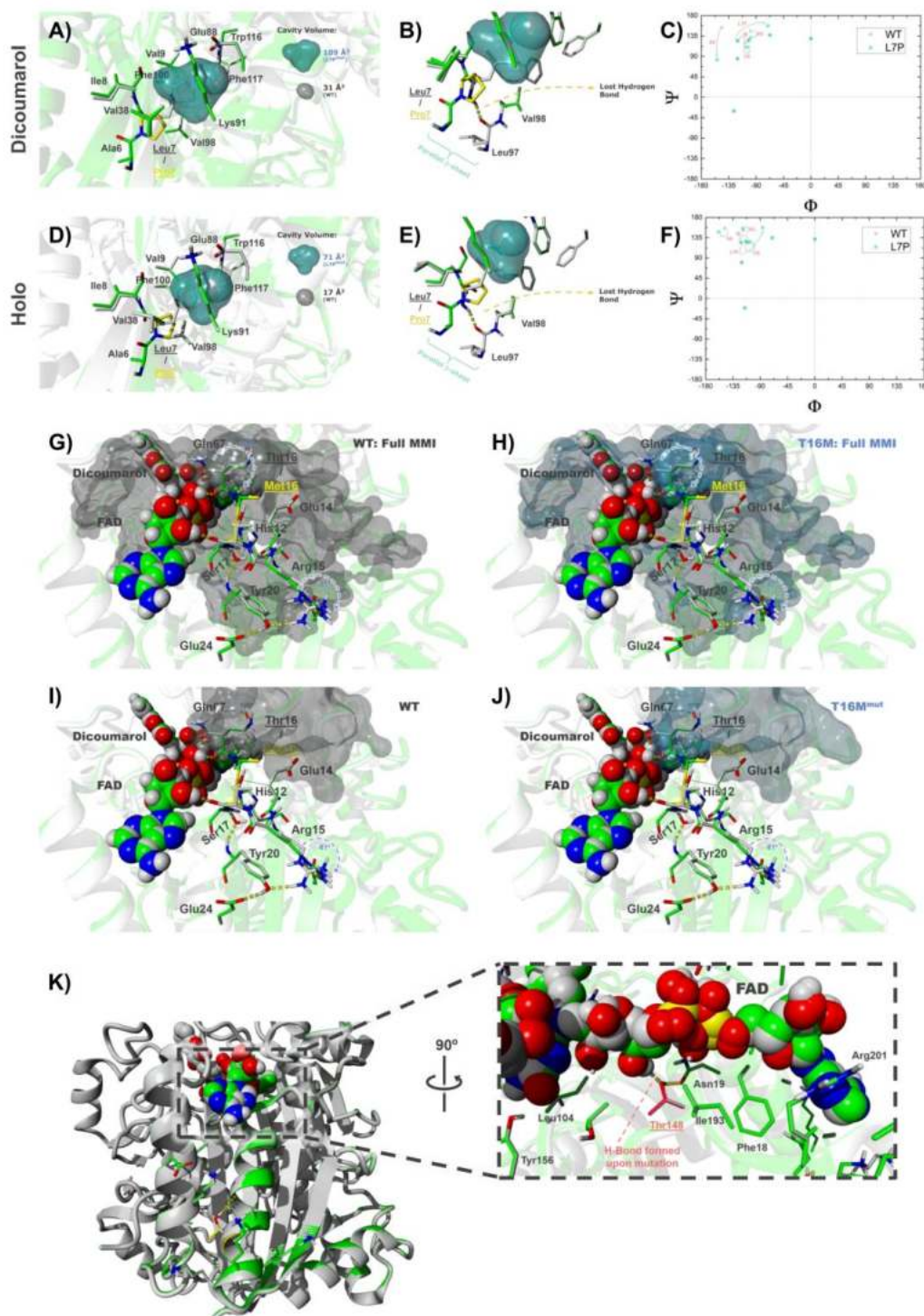
Figure 8. Detailed structural analysis of the mutations L7P (**A–F**), T16M (**G–J**) and A29T (**K**). (**A**) Overall view of the L7P site in the NQO1$_{dic}$ state (mutated residue in yellow). The mutation generates a cavity in the hydrophobic core of the protein (semi-transparent blue for L7P and gray for WT protein). (**B**) Close-up view on the hydrogen bond disrupted by L7P in the NQO1$_{dic}$ state. (**C**) Ramachandran plot of the first ten amino acids in the NQO1$_{dic}$ state. Arrows indicate the displacement in dihedral angles caused by L7P mutation that affects the β-sheet pattern near the mutated site. (**D**) Close-up of the mutated site (in yellow) in the NQO1$_{holo}$ state. The mutation L7P generates a cavity in the hydrophobic core of the protein represented in semi-transparent blue (L7P) and gray (WT). (**E**) Zoom-in

on the hydrogen bond formerly established by L7 and disrupted by the L7P mutation in NQO1$_{holo}$. (**F**) Ramachandran plot of the first ten amino acids in NQO1$_{holo}$ showing, with arrows, the displacement caused by L7P mutation that affects the β-sheet pattern near the mutated site. (**G–J**) Effects of the T16M mutation on the MMI of the NQO1$_{dic}$ state. The MMI surface is shown in gray for the WT and in bluish for the T16M mutant. The overall view is similar to that shown in Figure 7L. The complete MMI for the WT protein (**G**) vs. T16M mutant (**H**) is shown. Mutational effects are found in the top right corner. (**I,J**) Similar representations to those in panels (**G,H**) but only showing residues of the MMI close (<5 Å) to the mutated site. (**K**) Close-up view of the hydrogen bond formed between FAD and residue T148 in the A29T mutant in the NQO1$_{dic}$ state.

3.6.3. The Mutation V9I

Experimental characterization of the mutation V9I has shown nearly neutral effects. As analyzed by Rosetta, this mutation has marginal effects on conformational stability (Table 2). V9 (in NQO1 WT) and I9 (in the V9I mutant) are embedded in a similar hydrophobic environment to that of L7 (that is perturbed by the mutations L7P and L7R) (Figure 7I,J). Interestingly, the mutation V9I causes a favorable change in the van der Waals energy (fa_atr) term (Table S6) indicating increased van der Waals interactions with neighbouring hydrophobic residues (Figure 7I,J and Table S8) in both NQO1$_{holo}$ and NQO1$_{dic}$ states. Even though the introduction of a larger amino acid causes some structural strain (i.e., slightly displaces some neighbouring amino acids), this effect is essentially canceled out by the gain in favorable van der Waals interactions (Table S8).

3.6.4. The Mutation T16M

Our experimental analyses have shown that T16M has phenotypic consequences in several traits, particularly FAD binding, thermal stability and local stability of the TCS. Analysis by Rosetta shows that T16M mutation should have small effects on conformational stability (Table 2). Interestingly, decomposition of energetic contributions provided by Rosetta (Table S6) indicates different and opposing energetic contributions that nearly cancel out. A detailed structural analysis may provide further insight into these effects as well as in the phenotypic consequences of this mutation (Figure 7K,L).

The mutation T16M produces mild perturbations of the NQO1$_{holo}$ state (Figure 7K), causing small side-chain adjustments affecting residues R15, Y20, E24 and N65. Y20 and E24 residues are located in helix α1 and their rearrangement, together with the presence of M16, produces a small displacement in the FAD binding pose that could explain the effect of T16M on FAD binding.

Interestingly, we also observed noticeable changes in the modeled structures of NQO1$_{dic}$. The residue M16 (located in the first shell of interaction with the ligand FAD at a shortest distance < 5 Å) interacts through hydrophobic contacts with the residues nearby, such as the aliphatic portion of E14 side chain, and brings much closer the residue Q67. This, together with the perturbation in the backbone atoms of the residues adjacent to the mutation site, generates a moderate change in conformation on some residues close to the ligand binding site, ultimately altering its initial pose (i.e., that present in the WT). The residues affected by direct or indirect interaction with M16 are H12, E14, M16 (backbone) and F18, but the amino acids suffering the largest conformational changes on their side chains are Q67 and R15 (100° and 61° rotation across χ$_2$ and χ$_3$ angles, respectively), whose motions appear to distort important interactions for ligand binding (see Figure 7L and Video S1). The most significant change results in the reorientation of Q67, which now faces M16 side chain and, while it maintains its original hydrogen bond with FAD, forms another hydrogen bond with the carbonyl group present in F66 residue. The other large conformational change involves R15, which was engaged (in the WT NQO1$_{dic}$ state) in a hydrogen bond network with residues Y20 and E24 (located in helix α1). The mutation T16M disrupts the hydrogen bond with Y20 and makes R15 form two hydrogen bonds with D41 (achieving a virtually similar conformation to that shown in the WT NQO1$_{holo}$ state, see Video S2). It is interesting to speculate that since Dic is a competitive inhibitor of both the NAD(P)H and the substrate, these structural alterations may have consequences on the NQO1 catalytic cycle.

The residue M16 (as well as the wild-type residue T16) is part of the MMI. In fact, the changes described above can also be translated to some alterations of the MMI (Figure 8G–J) and may explain the effects of T16M in both thermal stability and FAD binding. The adjustments and side-chain reorientations described earlier are summarized in Video S1 (morphing conformations between the WT and T16M as NQO1$_{dic}$) and should be compared to Video S2 (morphing conformations between WT NQO1$_{dic}$ and NQO1$_{holo}$ states of the protein).

3.6.5. The Mutation Y20N

The mutant Y20N shows noticeable effects on thermal stability and FAD binding. Stability calculations using Rosetta provide moderate destabilizing effects (Table 2). Inspection of the modeled structures revealed certain structural alterations caused by this mutation (Figure 7M,N). Introduction of the smaller N20 residue prevents the formation of hydrogen bonds (in both NQO1$_{holo}$ and NQO1$_{dic}$ states) observed between Y20 and residues E24 and R15 (note that the latter residues also played important roles in the structural alterations caused by the mutation T16M). This alteration in hydrogen bonding causes slight backbone and side-chain movements on the residues nearby that are also part of the α1 helix, which also played a noteworthy role in the alterations experienced upon T16M mutation. These adjustments in the NQO1$_{dic}$ state cause slight changes in FAD binding pose (Video S3), comparable to those observed for the T16M mutation. In the holo form of the protein, a similar situation occurs when Y20 residue is mutated to N20. The broken hydrogen bond network destabilizes the surroundings of the mutated site, ultimately altering ligand binding pose (Video S4). In this case, E24 also rotates, by ~92°, its side chain around the χ_2 angle, presumably due to the inability to establish a hydrogen bond with residue N20. Overall, these results provide structural insight into the effect of Y20N on FAD binding affinity and also predict that this mutation might also affect the catalytic cycle of NQO1.

3.6.6. The Mutation A29T

The mutation A29T has essentially neutral effects, except for a moderate decrease in FAD binding affinity. Stability calculations using Rossetta suggest that this mutation substantially destabilizes the NQO1$_{holo}$ state (Table 2). Structural modeling shows that the T29 in this mutant establishes an intra-helix hydrogen bond between its side chain hydroxyl group and the carbonyl backbone group of amino acid A25 in both NQO1$_{holo}$ and NQO1$_{dic}$ states (Figure 7O,P), which might stabilize helix α1. In the NQO1$_{holo}$ form, this mutation causes side-chain conformational regroupings in the vicinity of the mutated site. However, in the NQO1$_{dic}$ state, these local adjustments propagate through helices α1 and α9 towards the FAD binding site and, ultimately, slightly modify its binding mode (See Videos S5 and S6) by forming a hydrogen bond with Oγ1 from residue T148 in the FAD binding site (Figure 8K). This interaction might contribute to the destabilizing effect of this mutation when the inhibitor is bound (as calculated by Rosetta) (Table 2).

3.7. The Role of Protein Local Dynamics and Stability on Mutational Effects

Protein structural dynamics are, likely, critical to understand many features of NQO1, such as the stabilizing effect of FAD binding towards proteasomal degradation of WT NQO1, ligand binding-mediated changes in the interaction with other proteins and the functional and stability alterations caused by missense mutations and polymorphisms [19,39,41,52,53,90,91]. Recently, we used hydrogen/deuterium exchange (HDX) monitored by mass spectrometry to provide a high-resolution map of the changes in protein local dynamics and stability of WT NQO1 in different ligation states (NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$) [90]. This study has allowed identification of a stable core (as non-exchanging segments) in NQO1$_{apo}$ that holds the NQO1 dimer, as well as the short- and long-range propagation of ligand binding effects (i.e., FAD and Dic). We have, thus, used this dynamic information in different ligation states to provide further molecular insight into the mutational effects in different protein traits experimentally characterized in this work.

Residues L7 and V9 belong to regions with very high structural stability in the NQO1$_{apo}$ state (i.e., these are located in the stable core) (Figure 9 and Figure S6 and Table S9). It is plausible that a significantly large structural perturbation at these sites would have dramatic effects on the ability of NQO1 to fold or the stability of the folded protein. Consequently, we could not obtain stable and soluble NQO1 protein with the largely-disrupting L7P or L7R mutations (Figure 2), whereas the more conservative V9I mutation mildly decreased dimer stability (Figure 3). Residues G3 and A29 are located in regions with moderate structural stability and their stability is hardly sensitive to ligand binding (Figure 9 and Figure S6, Tables S9 and S10). Accordingly, mutations G3S, G3D and A29T have mild effects on protein stability (Figure 3) and ligand binding, except for the six-fold decrease in FAD binding affinity in A29T (Figure 6). Residues T16 and Y20 are found in regions with low structural stability in NQO1$_{apo}$, although these regions manifest large structural stabilization by the stepwise binding of FAD and Dic (Figure 9 and Figure S6, Tables S9 and S10). Therefore, these analyses help to explain that mutations T16M and Y20N have moderate effects on protein stability (Figure 3) but exhibit important alterations in FAD binding (Figure 6), particularly the mutation T16M that affects a residue in close contact with the FAD molecule.
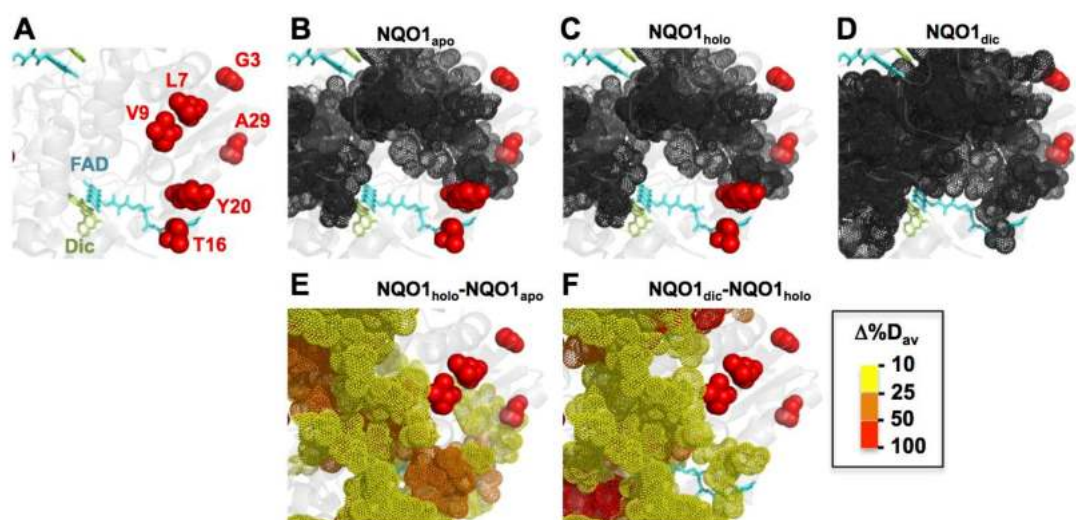
Figure 9. Location of mutated residues regarding local stability measurements by hydrogen/deuterium exchange (HDX). (**A**) Structural location of mutated residues regarding bound FAD and Dic. (**B**–**D**) Residues in dark grey/dot representation indicate those regions of the WT protein (NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$) that are highly stable (<20% HDX after 3 h). (**E**,**F**) Residues in dot representation (see color scale) indicate those whose stability (as $\Delta\%D_{av}$) was increased between the two states (NQO1$_{holo}$ vs. NQO1$_{apo}$ and NQO1$_{dic}$ vs. NQO1$_{holo}$). Plots are generated using primary data from [90].

Inspection of those regions of the protein that show high stability and/or strong ligand binding-dependent stability may provide further explananation for some of the functional alterations displayed by mutant NQO1 proteins. In Figure 9A–D, we show the highly stable regions (with minimal HDX after 3 h) in NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$ (see also Table S9). Residues L7 and V9 belong to this stable core in any of the three ligation states, whereas G3 does not in any of these ligation states (Figure 9A–D). The residue A29 is adjacent to this core in all three ligation states. In Figure 9E,F, we show those protein regions that undergo large stability changes upon binding FAD (Figure 9E) and Dic (Figure 9F). Regarding FAD binding, residues T16 and Y20 are buried in these regions that undergo large stability changes upon binding, thus supporting that their effects might be associated with alterations of the structural stabilization provided by FAD binding. A similar interpretation can be proposed to explain the effects of the mutation A29T, which affects a residue adjacent to those regions largely stabilized by FAD binding. Consequently, the lower affinity for FAD in the mutants T16M,

Y20N and A29T could be partially explained by their perturbing effects on regions stabilized upon FAD binding (Figure 9E). A similar interpretation could apply for the lack of effect of the mutations studied in this work on Dic binding, since overall, the residues affected by these mutations are located in regions that are not largely stabilized upon inhibitor binding (Figure 9F).

3.8. Correlations between Loss-of-Function Scores Derived from Experimental Analysis of Mutational Effects and Bioinformatic Tools

So far, we have shown that naturally-occurring mutations, found in either targeted (i.e., disease-associated) samples (COSMIC database) or whole-genome sequencing initiatives (gnomAD database), can affect, to different extents, diverse features in a multi-functional protein, such as NQO1. These results (compiled in the so-called N_t *set*) extend our previous knowledge on a variety of natural (*cancer-associated*, P187S and K240Q) as well as non-natural mutations generated at the P187 and K240 sites (named as *P187 and K240 sets*; see Table 2) [19,39]. To what extent can we predict the potential molecular and functional consequences (plausibly linked to their *potential pathogenicity*) of this set of 22 mutations using a combination of some of the most common bioinformatic algorithms (Table 4)? This task is challenging, since current algorithms usually provide a single metric to assess this potential pathogenicity, whereas, as we have shown here, mutational effects on protein functionality are intrinsically complex (i.e., the protein is multi-functional). In an attempt to assess our capacity to establish correlations between experimental analysis and bioinformatic predictive tools, we have determined a simple experimental score (ES) for this set of 22 mutations that takes into account, in a semi-quantitative manner, their effects on different functional features (Table 2). Although these are features characterized in vitro, we may expect that some of them can reflect loss-of-function phenotypes in vivo (and, thus, potential pathogenicity due to loss-of-function). For instance, severe folding/solubility problems, impaired thermal stability and reduced local stability are likely associated with folding and stability problems inside cells (e.g., the consequences of P187S on these features are clearly associated with its intracellular stability; [41,43,52,68]). In addition, reduced local stability and FAD binding affinity can be also associated with low intracellular activity and stability due to increased population of the inactive and degradation-prone apo-state (e.g., in the phosphomimetic mutant S82D; [91]). Reduced affinity for the inhibitor Dic may also reflect conformational alterations associated with enhanced intracellular degradation (the P187S polymorphism; [41]) and/or reduced intracellular sensitivity towards NQO1 inhibition, leading to reduced cancer growth [39,44,63,93].

ESs were meant to provide a semiquantitative ranking of pathogeniticy: an ES of 1 corresponded to a *severe* mutation, ES of 2 to a *mild* one and ES of 3 to a *neutral* mutation. Calculation of the ESs for this set of mutants showed that five mutations caused large defects on protein stability and/or solubility, and thus, these were considered as severe (i.e., ES of 1), including the mutations L7P and L7R (Table 2). The remaining 17 mutations showed an ES ranging from 1.4–3.0. Among the lowest ESs, we found the polymorphism P187S, the *gold standard* of loss-of-function in NQO1 [41,43,52,68]. Most of the mutations in the N_t and K240 sets showed mild to neutral ESs (ES from 2.2 to 3), whereas variants in the P187 set were more deleterious (ES from 1 to 2).

We used these ESs to ascertain whether popular bioinformatic tools would predict and rank potential pathogenic effects. To this end, we generated bioinformatic scores (BSs) to be semiquantitatively compared with the ESs (Tables 3 and 4). The results from these correlations are shown in Figure 10. When the full mutant set was analyzed, we observed a significant positive correlation between ESs and BSs (Figure 10). Visual inspection of this plot shows this correlation performs differently for different sets of mutations. The correlation for the N_t set is much stronger than that for the P187 set and, particularly, than that for the K240 set (Figure 10). The origin of this difference in performance is unclear. However, the very poor perfomance with the K240 set might have a straightforward explanation: this residue is highly conserved among mammalian NQO1 sequences, thus likely explaining that bioinformatic tools may identify mutations in this residue as potentially pathogenic. However, experimental analyses have shown that even highly disrupting

mutants (such as K240E and K240G) still show many functional aspects not departing much from those of WT NQO1 [19,39]. These results were somehow expected: current bioinformatic approaches may provide some good overall results for genotype–phenotype correlations but fail to predict accurately and individually this correlation for individual mutations (one of the holy grails of personalized medicine, [99]), particularly when mutations affect highly conserved residues.

Table 3. Classification of phenotypic consequences due to NQO1 mutations based on experimental data. The experimental scores (ES) are calculated as the average from individual scores determined for each phenotypic trait as follows: *Expression:* +++, 50–100% of WT levels; ++, 20–50% of WT levels; +, <20% of WT levels; *Thermal stability:* +++, within 2 °C of WT T_m; ++, within 2–5 °C of WT T_m; +, >5 °C lower than WT T_m; *Proteolysis:* +++, rate constant within three-fold vs. WT (0.65 kcal·mol^{-1}); ++, within 3 to 30-fold vs. WT (2 kcal·mol^{-1}); ++ >30-fold faster than WT (over 2 kcal·mol^{-1}); *FAD binding:* +++, within three-fold of WT K_d; ++, within 3–10 fold higher than WT K_d; +, >10-fold higher than WT K_d; *Dic binding:* +++, within three-fold of WT K_d; ++, within 3–10 fold higher than WT K_d; +, >10-fold higher than WT K_d.

Variant	Expression	Thermal Stability	Proteolysis	FAD Binding	Dic Binding	ES
G3S	+++	+++	+++	+++	+++	3
G3D	+++	+++	+++	+++	+++	3
L7P	+					1
L7R	+					1
V9I	+++	+++	+++	+++	+++	3
T16M	+++	++	++	+	+++	2.2
Y20N	+++	++	+++	++	+++	2.6
A29T	+++	+++	+++	++	+++	2.8
P187S	+++	+	+	+	+	1.4
P187E	+					1
P187R	+					1
P187L	+					1
P187A	+++	++	+	++	++	2
P187G	++	+	++	+++	++	2
P187T	+++	++	+	++	+	1.8
K240Q	+++	+++	++	++	++	2.4
K240I	+++	+++	++	+++	+++	2.8
K240E	+++	+++	+	++	++	2.2
K240T	+++	+++	+++	++	+++	2.8
K240H	+++	+++	++	++	+++	2.6
K240A	+++	+++	+	+++	+++	2.6
K240G	+++	+++	+	++	++	2.2

Table 4. Classification of phenotypic consequences of NQO1 mutations based on bioinformatic analysis. The bioinformatic scores (BSs) are calculated as the average from individual scores provided by different algorithms as follows: *Polyphen-2:* +++, Benign; ++, Possibly damaging; +, Probably damaging; *SIFT:* +++, tolerated; +, Not tolerated; *Mutation Taster:* +++, polymorphism; +, disease-causing; *Provean:* +++, neutral; +, deleterious; PON-P2: +++, neutral; ++, unknown; +, pathogenic. * N.Det.- not determined because two nucleotide have to change.

Variant	PolyPhen-2	SIFT	Mutation Taster	Provean	PON-P2	BS
G3S	+++	+++	+++	+++	+++	3
G3D	+++	+++	+++	+++	++	2.8
L7P	+	+	+	+	+	1
L7R	+	+	+	+	+	1
V9I	+++	++	+	+++	++	2.2
T16M	+	+	+	+++	+	1.4
Y20N	++	+++	+	+++	++	2.2
A29T	+++	+++	+	+++	+++	2.6
P187S	+++	+++	+++	+	++	2
P187E	++	+	N.Det. *	+	+	1.25
P187R	+	+	+	+	+	1
P187L	+	+	+	+	+	1
P187A	++	+++	+	+	++	1.8
P187G	++	+	N.Det. *	+	+	1.25
P187T	++	+++	+	+	+	1.6
K240Q	+	+	+	+++	++	1.6
K240I	+	+	+	+	++	1.2
K240E	+	+	+	+++	++	1.6
K240T	+	+	+	+++	++	1.6
K240H	+	+	N.Det. *	+++	++	1.75
K240A	+	+	N.Det. *	+++	++	1.75
K240G	+	+	N.Det. *	+	++	1.25
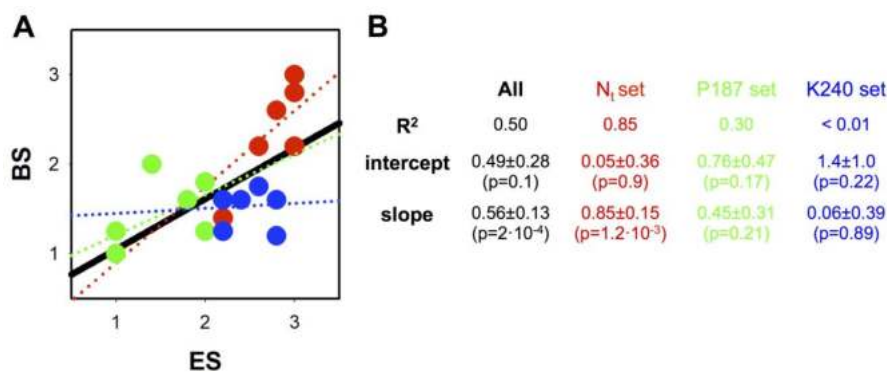
Figure 10. Correlation between experimental scores (ESs) and bioinformatic scores (BSs). (**A**) Plot of ES vs. BS for all mutants compiled in Tables 3 and 4. Red circles show mutants belonging to the *Nt set* (the eight mutants characterized in this work) and green and blue circles show mutants belonging to mutations at the P187 and K240 sites, respectively (*P187* and *K240 sets*, see Tables 3 and 4 [39]). The thick black line is the linear regression of all mutants, while dotted lines are linear fittings using one of the three sets (N$_t$, P187 or K240 sets; red, green and blue lines, respectively). (**B**) Best-fit parameters for linear regressions are shown in panel B.

4. Conclusions

From the experimental characterization of the effects of eight naturally-occurring mutations in the multifunctional NQO1 protein, combined with detailed structural analysis and bioinformatic predictions, we can draw several important conclusions and propose future research lines.

First, by investigating the consequences of missense mutations found in whole-genome sequencing studies (with different allelic frequencies; see Table 1) as well as those found in cancer cell lines, we have observed that the severity of the pathogenic effects does not correlate well with their presence in cancer samples (i.e., the mutations L7P and L7R cause similarly catastrophic effects) nor with frequency in untargeted whole-genome sequencing studies (e.g., L7R, a severe mutation, has a 6.5-fold higher allelic frequency than G3S, a neutral mutation). Interestingly, one of the most deleterious natural missense variants in NQO1 (P187S) is also the most common in the human population (with an allelic frequency over 0.3) and is the missense variant most robustly associated with predisposition to disease [47,65]. These results may apparently contradict recent large-scale mutational analysis in several disease-associated protein systems that showed a negative correlation between allelic frequency of missense mutations and their effects on structural and interacellular stability [5,9,10,26]. This might imply that the correlation between allelic frequency and potential pathogenicity is not universal, that NQO1 is just an odd exception or that incorporating multiple functional features (stability, activity, regulation, ...) into the picture may lead to more complex correlations. Obviously, understanding these and other possible scenarios deserves further research, including the analysis of larger sets of naturally-occuring missense mutations in diverse disease-associated proteins.

Second, structural–energetic–functional relationships of mutational effects are complex, and thus, an integrated view from biochemical and biophysical functional and stability studies, structural modeling and energetic considerations provides a suitable approach to characterize mutational effects in multifunctional proteins. Ideally, this should be also complemented with expression studies and stability analyses in eukaryotic cells. Obviously, all of this cannot be done at a whole-genome scale. However, at a smaller scale, these studies could provide insight into key concepts to be implemented in current *in silico* approaches for large-scale genotype–phenotype predictions. For instance, it is particularly relevant to consider the locally destabilizing effects of mutations and how these effects can propagate to distant sites in protein structures, thus contributing to pleiotropic (i.e., mulfunctional) effects of missense mutations [19,22,27–29,39,100]. It is also of interest that the threshold to cause intracellular destabilization with pathogenic consequences seems to be quite low, in the order of 2–3 kcal·mol^{-1} (often determined as the effects on overall conformational stability from experimental or computational methods) [9,10,26]. We should also consider local stability effects, not only for mutational effects on intracellular stability but also for functional consequences. Although P187S is known to cause devastating effects on intracellular stability and function, these effects are associated with changes in local stability in the order of 2–3 kcal·mol^{-1} [41], and effects within this range are observed for other naturally-occurring variants (such as T16M, Y20N, A29T and K240Q) (Table 3). Since propagation of mutations effects to distant sites in protein structure is likely a universal behavior [31,101–103], we would anticipate that the pleitropic effects described for missense mutations in NQO1 could apply to many other disease-associated protein systems [19].

Third, it is also of interest that scores based on quantitative analysis of mutational effects and those built using diverse current bioinformatic tools provided a positive correlation for a set of 22 mutations on NQO1. However, as we have shown, this predictive power is still disappointing when genotype–phenotype correlations are attempted for single mutations and also performs quite poorly for certain mutated sites (e.g., the K240 site). Thus, our work further supports the notion that detailed characterization of mutational effects by extensive functional and stability analysis will help to improve our capacity for accurate and large-scale genotype–phenotype correlations.

variants with Dic at 25 °C. Figure S5. Thermodynamic dissection of Dic binding to NQO1 variants. Figure S6. HDX analysis of structural regions in which NQO1 mutations are localized. Table S1. Effect on conformational stability due to NQO1 mutations determined using structure-based methods. Table S2. Second-order rate constants (k_{prot}) for proteolysis by thermolysin and mutational effects on local stability ($\Delta\Delta G_{prot}$). Table S3. FAD binding affinity of NQO1 variants. Table S4. Thermodynamic parameters for Dic binding to NQO1 variants. Table S5. Magnitude of the conformational change (N_{conf}) associated with Dic binding from analysis of apparent binding enthalpies (ΔH) and heat capacities (ΔC_p). Table S6. Energetic dissection of mutational effects as determined by Rosetta. Table S7. Analysis of steric clashes present in the $NQO1_{holo}$ and $NQO1_{dic}$ states caused by the mutation L7P. Table S8. Hydrophobic interactions established between V9 (WT) and I9 (mutant V9I) in the $NQO1_{holo}$ and $NQO1_{dic}$ states. Table S9. Regions with < 20% HDX after 3 h. Table S10. Regions with $\Delta\%D_{av} \geq 10\%$. Video S1. Morphing conformations between WT $NQO1_{dic}$ and T16M mutated model. Video S2. Morphing conformations between WT $NQO1_{dic}$ and WT $NQO1_{holo}$ states. Video S3. Morphing conformations between WT $NQO1_{dic}$ and Y20N mutated model. Video S4. Morphing conformations between WT $NQO1_{holo}$ and Y20N mutated model. Video S5. Morphing conformations between WT $NQO1_{dic}$ and A29T mutated model. Video S6. Morphing conformations between WT $NQO1_{holo}$ and A29T mutated model.

Author Contributions: Conceptualization, A.L.P.; methodology, M.C.-M., E.S. and A.L.P.; software, M.C.-M. and A.L.P.; validation, M.C.-M., E.S. and A.L.P; formal analysis, J.L.P.-G., M.C.-M., I.S.-R., E.S. and A.L.P.; investigation, J.L.P.-G., M.C.-M., I.S.-R. and A.L.P.; resources, E.S. and A.L.P.; data curation, M.C.-M., E.S. and A.L.P.; writing—original draft preparation, M.C.-M., E.S. and A.L.P.; writing—review and editing, J.L.P.-G., M.C.-M., I.S.-R., E.S. and A.L.P.; visualization, A.L.P.; supervision, A.L.P.; project administration, A.L.P.; funding acquisition, E.S. and A.L.P. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shendure, J.; Akey, J.M. The origins, determinants, and consequences of human mutations. *Science* **2015**, *349*, 1478–1483. [CrossRef]
2. Lek, M.; Karczewski, K.J.; Minikel, E.V.; Samocha, K.E.; Banks, E.; Fennell, T.; O'Donnell-Luria, A.H.; Ware, J.S.; Hill, A.J.; Cummings, B.B.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285–291. [CrossRef] [PubMed]
3. The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature* **2020**, *578*, 82–93. [CrossRef] [PubMed]
4. Kiryluk, K.; Goldstein, D.B.; Rowe, J.W.; Gharavi, A.G.; Wapner, R.; Chung, W.K. Precision Medicine in Internal Medicine. *Ann. Intern. Med.* **2019**, *170*, 635–642. [CrossRef]
5. Stein, A.; Fowler, D.M.; Hartmann-Petersen, R.; Lindorff-Larsen, K. Biophysical and Mechanistic Models for Disease-Causing Protein Variants. *Trends Biochem. Sci.* **2019**, *44*, 475–488. [CrossRef]
6. Manolio, T.A.; Fowler, D.M.; Starita, L.M.; Haendel, M.A.; MacArthur, D.G.; Biesecker, L.G.; Worthey, E.; Chisholm, R.L.; Green, E.D.; Jacob, H.J.; et al. Bedside Back to Bench: Building Bridges between Basic and Clinical Genomic Research. *Cell* **2017**, *169*, 6–12. [CrossRef]
7. Casanueva, M.O.; Burga, A.; Lehner, B. Fitness trade-offs and environmentally induced mutation buffering in isogenic C. elegans. *Science* **2012**, *335*, 82–85. [CrossRef]
8. Pey, A.L.; Stricher, F.; Serrano, L.; Martinez, A. Predicted effects of missense mutations on native-state stability account for phenotypic outcome in phenylketonuria, a paradigm of misfolding diseases. *Am. J. Hum. Genet.* **2007**, *81*, 1006–1024. [CrossRef]
9. Nielsen, S.V.; Stein, A.; Dinitzen, A.B.; Papaleo, E.; Tatham, M.H.; Poulsen, E.G.; Kassem, M.M.; Rasmussen, L.J.; Lindorff-Larsen, K.; Hartmann-Petersen, R. Predicting the impact of Lynch syndrome-causing missense mutations from structural calculations. *PLoS Genet.* **2017**, *13*, e1006739. [CrossRef]
10. Abildgaard, A.B.; Stein, A.; Nielsen, S.V.; Schultz-Knudsen, K.; Papaleo, E.; Shrikhande, A.; Hoffmann, E.R.; Bernstein, I.; Gerdes, A.M.; Takahashi, M.; et al. Computational and cellular studies reveal structural destabilization and degradation of MLH1 variants in Lynch syndrome. *eLife* **2019**, *8*, e49138. [CrossRef] [PubMed]

11. Blouin, J.M.; Bernardo-Seisdedos, G.; Sasso, E.; Esteve, J.; Ged, C.; Lalanne, M.; Sanz-Parra, A.; Urquiza, P.; de Verneuil, H.; Millet, O.; et al. Missense UROS mutations causing congenital erythropoietic porphyria reduce UROS homeostasis that can be rescued by proteasome inhibition. *Hum. Mol. Genet.* **2017**, *26*, 1565–1576. [CrossRef]

12. Boycott, K.M.; Rath, A.; Chong, J.X.; Hartley, T.; Alkuraya, F.S.; Baynam, G.; Brookes, A.J.; Brudno, M.; Carracedo, A.; den Dunnen, J.T.; et al. International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. *Am. J. Hum. Genet.* **2017**, *100*, 695–705. [CrossRef]

13. Mesa-Torres, N.; Fabelo-Rosa, I.; Riverol, D.; Yunta, C.; Albert, A.; Salido, E.; Pey, A.L. The role of protein denaturation energetics and molecular chaperones in the aggregation and mistargeting of mutants causing primary hyperoxaluria type I. *PLoS ONE* **2013**, *8*, e71963. [CrossRef]

14. Pey, A.L.; Mesa-Torres, N.; Chiarelli, L.R.; Valentini, G. Structural and Energetic Basis of Protein Kinetic Destabilization in Human Phosphoglycerate Kinase 1 Deficiency. *Biochemistry* **2013**, *52*, 1160–1170. [CrossRef]

15. Pey, A.L.; Maggi, M.; Valentini, G. Insights into human phosphoglycerate kinase 1 deficiency as a conformational disease from biochemical, biophysical, and in vitro expression analyses. *J. Inherit. Metab. Dis.* **2014**, *37*, 909–916. [CrossRef]

16. Pey, A.L.; Padin-Gonzalez, E.; Mesa-Torres, N.; Timson, D.J. The metastability of human UDP-galactose 4′-epimerase (GALE) is increased by variants associated with type III galactosemia but decreased by substrate and cofactor binding. *Arch. Biochem. Biophys.* **2014**, *562*, 103–114. [CrossRef]

17. Majtan, T.; Pey, A.L.; Gimenez-Mascarell, P.; Martinez-Cruz, L.A.; Szabo, C.; Kozich, V.; Kraus, J.P. Potential Pharmacological Chaperones for Cystathionine Beta-Synthase-Deficient Homocystinuria. *Handb. Exp. Pharmacol.* **2018**, *245*, 345–383.

18. Fernandez-Higuero, J.A.; Betancor-Fernandez, I.; Mesa-Torres, N.; Muga, A.; Salido, E.; Pey, A.L. Structural and functional insights on the roles of molecular chaperones in the mistargeting and aggregation phenotypes associated with primary hyperoxaluria type I. *Adv. Protein Chem. Struct. Biol.* **2019**, *114*, 119–152.

19. Medina-Carmona, E.; Betancor-Fernández, I.; Santos, J.; Mesa-Torres, N.; Grottelli, S.; Batlle, C.; Naganathan, A.N.; Oppici, O.; Cellini, B.; Ventura, S.; et al. Insight into the specificity and severity of pathogenic mechanisms associated with missense mutations through experimental and structural perturbation analyses. *Hum. Mol. Genet.* **2019**, *28*, 1–15. [CrossRef]

20. Macias, I.; Lain, A.; Bernardo-Seisdedos, G.; Gil, D.; Gonzalez, E.; Falcon-Perez, J.M.; Millet, O. Hereditary tyrosinemia type I-associated mutations in fumarylacetoacetate hydrolase reduce the enzyme stability and increase its aggregation rate. *J. Biol. Chem.* **2019**, *294*, 13051–13060. [CrossRef] [PubMed]

21. Fossbakk, A.; Kleppe, R.; Knappskog, P.M.; Martinez, A.; Haavik, J. Functional studies of tyrosine hydroxylase missense variants reveal distinct patterns of molecular defects in Dopa-responsive dystonia. *Hum. Mutat.* **2014**, *35*, 880–890. [CrossRef]

22. McCorvie, T.J.; Gleason, T.J.; Fridovich-Keil, J.L.; Timson, D.J. Misfolding of galactose 1-phosphate uridylyltransferase can result in type I galactosemia. *Biochim. Biophys. Acta* **2013**, *1832*, 1279–1293. [CrossRef]

23. McCorvie, T.J.; Kopec, J.; Pey, A.L.; Fitzpatrick, F.; Patel, D.; Chalk, R.; Streetha, L.; Yue, W.W. Molecular basis of classic galactosemia from the structure of human galactose 1-phosphate uridylyltransferase. *Hum. Mol. Genet.* **2016**, *25*, 2234–2244. [CrossRef] [PubMed]

24. Mesa-Torres, N.; Betancor-Fernández, I.; Oppici, E.; Cellini, B.; Salido, E.; Pey, A.L. Evolutionary Divergent Suppressor Mutations in Conformational Diseases. *Genes* **2018**, *9*, 352. [CrossRef]

25. Clausen, L.; Abildgaard, A.B.; Gersing, S.K.; Stein, A.; Lindorff-Larsen, K.; Hartmann-Petersen, R. Protein stability and degradation in health and disease. *Adv. Protein Chem. Struct. Biol.* **2019**, *114*, 61–83. [PubMed]

26. Scheller, R.; Stein, A.; Nielsen, S.V.; Marin, F.I.; Gerdes, A.M.; Marco, M.D.; Papaleo, E.; Lindorff-Larsen, K.; Hartmann-Petersen, R. Towards mechanistic models for genotype-phenotype correlations in phenylketonuria using protein stability calculations. *Hum. Mutat.* **2019**, *40*, 444–457. [CrossRef]

27. Erlandsen, H.; Pey, A.L.; Gamez, A.; Perez, B.; Desviat, L.R.; Aguado, C.; Koch, R.; Surendran, S.; Tyring, S.; Matalon, R.; et al. Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16903–16908. [CrossRef]

28. Pey, A.L.; Majtan, T.; Sanchez-Ruiz, J.M.; Kraus, J.P. Human cystathionine beta-synthase (CBS) contains two classes of binding sites for S-adenosylmethionine (SAM): Complex regulation of CBS activity and stability by SAM. *Biochem. J.* **2013**, *449*, 109–121. [CrossRef]

29. Fortian, A.; Castano, D.; Ortega, G.; Lain, A.; Pons, M.; Millet, O. Uroporphyrinogen III synthase mutations related to congenital erythropoietic porphyria identify a key helix for protein stability. *Biochemistry* **2009**, *48*, 454–461. [CrossRef] [PubMed]

30. Schmiesing, J.; Lohmoller, B.; Schweizer, M.; Tidow, H.; Gersting, S.W.; Muntau, A.C.; Braulke, T.; Muhlhausen, C. Disease-causing mutations affecting surface residues of mitochondrial glutaryl-CoA dehydrogenase impair stability, heteromeric complex formation and mitochondria architecture. *Hum. Mol. Genet.* **2017**, *26*, 538–551. [CrossRef]

31. Naganathan, A.N. Modulation of allosteric coupling by mutations: From protein dynamics and packing to altered native ensembles and function. *Curr. Opin. Struct. Biol.* **2018**, *54*, 1–9. [CrossRef] [PubMed]

32. Adzhubei, I.A.; Schmidt, S.; Peshkin, L.; Ramensky, V.E.; Gerasimova, A.; Bork, P.; Kondrashov, A.S.; Sunyaev, S.R. A method and server for predicting damaging missense mutations. *Nat. Methods* **2010**, *7*, 248–249. [CrossRef]

33. Guerois, R.; Nielsen, J.E.; Serrano, L. Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *J. Mol. Biol.* **2002**, *320*, 369–387. [CrossRef]

34. Sanchez, I.E.; Tejero, J.; Gomez-Moreno, C.; Medina, M.; Serrano, L. Point mutations in protein globular domains: Contributions from function, stability and misfolding. *J. Mol. Biol.* **2006**, *363*, 422–432. [CrossRef]

35. Tokuriki, N.; Stricher, F.; Schymkowitz, J.; Serrano, L.; Tawfik, D.S. The stability effects of protein mutations appear to be universally distributed. *J. Mol. Biol.* **2007**, *369*, 1318–1332. [CrossRef] [PubMed]

36. Nisthal, A.; Wang, C.Y.; Ary, M.L.; Mayo, S.L. Protein stability engineering insights revealed by domain-wide comprehensive mutagenesis. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 16367–16377. [CrossRef]

37. Khan, S.; Vihinen, M. Performance of protein stability predictors. *Hum. Mutat.* **2010**, *31*, 675–684. [CrossRef]

38. Roscoe, B.P.; Thayer, K.M.; Zeldovich, K.B.; Fushman, D.; Bolon, D.N. Analyses of the effects of all ubiquitin point mutants on yeast growth rate. *J. Mol. Biol.* **2013**, *425*, 1363–1377. [CrossRef]

39. Pey, A.L. Biophysical and functional perturbation analyses at cancer-associated P187 and K240 sites of the multifunctional NADP(H):quinone oxidoreductase 1. *Int. J. Biol. Macromol.* **2018**, *118*, 1912–1923. [CrossRef] [PubMed]

40. Pey, A.L.; Megarity, C.F.; Timson, D.J. FAD binding overcomes defects in activity and stability displayed by cancer-associated variants of human NQO1. *Biochim. Biophys. Acta* **2014**, *1842*, 2163–2173. [CrossRef]

41. Medina-Carmona, E.; Palomino-Morales, R.J.; Fuchs, J.E.; Padín-Gonzalez, E.; Mesa-Torres, N.; Salido, E.; Timson, D.J.; Pey, A.L. Conformational dynamics is key to understanding loss-of-function of NQO1 cancer-associated polymorphisms and its correction by pharmacological ligands. *Sci. Rep.* **2016**, *6*, 20331. [CrossRef] [PubMed]

42. Claveria-Gimeno, R.; Velazquez-Campoy, A.; Pey, A.L. Thermodynamics of cooperative binding of FAD to human NQO1: Implications to understanding cofactor-dependent function and stability of the flavoproteome. *Arch Biochem. Biophys.* **2017**, *636*, 17–27. [CrossRef]

43. Medina-Carmona, E.; Fuchs, J.E.; Gavira, J.A.; Mesa-Torres, N.; Neira, J.L.; Salido, E.; Palomino-Morales, R.; Burgos, M.; Timson, D.J.; Pey, A.L. Enhanced vulnerability of human proteins towards disease-associated inactivation through divergent evolution. *Hum. Mol. Genet.* **2017**, *26*, 3531–3544. [CrossRef]

44. Medina-Carmona, E.; Neira, J.L.; Salido, E.; Fuchs, J.E.; Palomino-Morales, R.; Timson, D.J.; Pey, A.L. Site-to-site interdomain communication may mediate different loss-of-function mechanisms in a cancer-associated NQO1 polymorphism. *Sci. Rep.* **2017**, *7*, 44352. [CrossRef]

45. Munoz, I.G.; Morel, B.; Medina-Carmona, E.; Pey, A.L. A mechanism for cancer-associated inactivation of NQO1 due to P187S and its reactivation by the consensus mutation H80R. *FEBS Lett.* **2017**, *591*, 2826–2835. [CrossRef]

46. Li, R.; Bianchet, M.A.; Talalay, P.; Amzel, L.M. The three-dimensional structure of NAD(P)H:quinone reductase, a flavoprotein involved in cancer chemoprotection and chemotherapy: Mechanism of the two-electron reduction. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8846–8850. [CrossRef]

47. Beaver, S.K.; Mesa-Torres, N.; Pey, A.L.; Timson, D.J. NQO1: A target for the treatment of cancer and neurological diseases, and a model to understand loss of function disease mechanisms. *Biochim. Biophys. Acta Proteins Proteom.* **2019**, *1867*, 663–676. [CrossRef]

48.   Pey, A.L.; Megarity, C.F.; Medina-Carmona, E.; Timson, D.J. Natural Small Molecules as Stabilizers and Activators of Cancer-Associated NQO1 Polymorphisms. *Curr. Drug Targets* **2016**, *17*, 1506–1514. [CrossRef] [PubMed]

49.   Lienhart, W.D.; Gudipati, V.; Uhl, M.K.; Binter, A.; Pulido, S.A.; Saf, R.; Zangger, K.; Gruber, K.; Macheroux, P. Collapse of the native structure caused by a single amino acid exchange in human NAD(P)H:quinone oxidoreductase(1.). *FEBS J.* **2014**, *281*, 4691–4704. [CrossRef] [PubMed]

50.   Chen, S.; Deng, P.S.; Bailey, J.M.; Swiderek, K.M. A two-domain structure for the two subunits of NAD(P)H:quinone acceptor oxidoreductase. *Protein Sci.* **1994**, *3*, 51–57. [CrossRef]

51.   Faig, M.; Bianchet, M.A.; Talalay, P.; Chen, S.; Winski, S.; Ross, D.; Amzel, L.M. Structures of recombinant human and mouse NAD(P)H:quinone oxidoreductases: Species comparison and structural changes with substrate binding and release. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3177–3182. [CrossRef]

52.   Martinez-Limon, A.; Alriquet, M.; Lang, W.H.; Calloni, G.; Wittig, I.; Vabulas, R.M. Recognition of enzymes lacking bound cofactor by protein quality control. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12156–12161. [CrossRef]

53.   Salido, E.; Timson, D.J.; Betancor-Fernández, I.; Palomino-Morales, R.; Pey, A.L. Targeting HIF-1alpha Function in Cancer through the Chaperone Action of NQO1. *Preprints* **2020**, *2020030285*.

54.   Anusevicius, Z.; Sarlauskas, J.; Cenas, N. Two-electron reduction of quinones by rat liver NAD(P)H:quinone oxidoreductase: Quantitative structure-activity relationships. *Arch. Biochem. Biophys.* **2002**, *404*, 254–262. [CrossRef]

55.   Landi, L.; Fiorentini, D.; Galli, M.C.; Segura-Aguilar, J.; Beyer, R.E. DT-Diaphorase maintains the reduced state of ubiquinones in lipid vesicles thereby promoting their antioxidant function. *Free Radic. Biol. Med.* **1997**, *22*, 329–335. [CrossRef]

56.   Siegel, D.; Bolton, E.M.; Burr, J.A.; Liebler, D.C.; Ross, D. The reduction of alpha-tocopherolquinone by human NAD(P)H: Quinone oxidoreductase: The role of alpha-tocopherolhydroquinone as a cellular antioxidant. *Mol. Pharmacol.* **1997**, *52*, 300–305. [CrossRef]

57.   Siegel, D.; Gustafson, D.L.; Dehn, D.L.; Han, J.Y.; Boonchoong, P.; Berliner, L.J.; Ross, D. NAD(P)H:quinone oxidoreductase 1: Role as a superoxide scavenger. *Mol. Pharmacol.* **2004**, *65*, 1238–1247. [CrossRef]

58.   Ross, D.; Siegel, D. NQO1 in protection against oxidative stress. *Curr. Opin. Toxicol.* **2018**, *7*, 67–72. [CrossRef]

59.   Anoz-Carbonell, E.; Timson, D.J.; Pey, A.L.; Medina, M. The Catalytic Cycle of the Antioxidant and Cancer-Associated Human NQO1 Enzyme: Hydride Transfer, Conformational Dynamics and Functional Cooperativity. *Antioxidants* **2020**, *9*, 772. [CrossRef]

60.   Timson, D.J. Dicoumarol: A Drug which Hits at Least Two Very Different Targets in Vitamin K Metabolism. *Curr. Drug Targets* **2017**, *18*, 500–510. [CrossRef] [PubMed]

61.   Oh, E.T.; Kim, J.W.; Kim, J.M.; Kim, S.J.; Lee, J.S.; Hong, S.S.; Goodwin, J.; Ruthenborg, R.J.; Jung, M.G.; Lee, H.J.; et al. NQO1 inhibits proteasome-mediated degradation of HIF-1alpha. *Nat. Commun.* **2016**, *7*, 13593. [CrossRef]

62.   Asher, G.; Tsvetkov, P.; Kahana, C.; Shaul, Y. A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73. *Genes Dev.* **2005**, *19*, 316–321. [CrossRef] [PubMed]

63.   Nolan, K.A.; Zhao, H.; Faulder, P.F.; Frenkel, A.D.; Timson, D.J.; Siegel, D.; Ross, D.; Burke, T.R., Jr.; Stratford, I.J.; Bryce, R.A. Coumarin-based inhibitors of human NAD(P)H:quinone oxidoreductase-1. Identification, structure-activity, off-target effects and in vitro human pancreatic cancer toxicity. *J. Med. Chem.* **2007**, *50*, 6316–6325. [CrossRef]

64.   Cullen, J.J.; Hinkhouse, M.M.; Grady, M.; Gaut, A.W.; Liu, J.; Zhang, Y.P.; Weydert, C.J.; Domann, F.E.; Oberley, L.W. Dicumarol inhibition of NADPH:quinone oxidoreductase induces growth inhibition of pancreatic cancer via a superoxide-mediated mechanism. *Cancer Res.* **2003**, *63*, 5513–5520.

65.   Lajin, B.; Alachkar, A. The NQO1 polymorphism C609T (Pro187Ser) and cancer susceptibility: A comprehensive meta-analysis. *Br. J. Cancer* **2013**, *109*, 1325–1337. [CrossRef]

66.   Traver, R.D.; Horikoshi, T.; Danenberg, K.D.; Stadlbauer, T.H.; Danenberg, P.V.; Ross, D.; Gibson, N.W. NAD(P)H:quinone oxidoreductase gene expression in human colon carcinoma cells: Characterization of a mutation which modulates DT-diaphorase activity and mitomycin sensitivity. *Cancer Res.* **1992**, *52*, 797–802.

67.   Traver, R.D.; Siegel, D.; Beall, H.D.; Phillips, R.M.; Gibson, N.W.; Franklin, W.A.; Ross, D. Characterization of a polymorphism in NAD(P)H: Quinone oxidoreductase (DT-diaphorase). *Br. J. Cancer* **1997**, *75*, 69–75. [CrossRef]

68. Siegel, D.; Anwar, A.; Winski, S.L.; Kepa, J.K.; Zolman, K.L.; Ross, D. Rapid polyubiquitination and proteasomal degradation of a mutant form of NAD(P)H:quinone oxidoreductase 1. *Mol. Pharmacol.* **2001**, *59*, 263–268. [CrossRef]

69. Eguchi-Ishimae, M.; Eguchi, M.; Ishii, E.; Knight, D.; Sadakane, Y.; Isoyama, K.; Yabe, H.; Mizutani, S.; Greaves, M. The association of a distinctive allele of NAD(P)H:quinone oxidoreductase with pediatric acute lymphoblastic leukemias with MLL fusion genes in Japan. *Haematologica* **2005**, *90*, 1511–1515. [PubMed]

70. Lienhart, W.D.; Strandback, E.; Gudipati, V.; Koch, K.; Binter, A.; Uhl, M.K.; Rantasa, D.M.; Bourgeois, B.; Madl, T.; Zangger, K.; et al. Catalytic competence, structure and stability of the cancer-associated R139W variant of the human NAD(P)H:quinone oxidoreductase 1 (NQO1). *FEBS J.* **2017**, *284*, 1233–1245. [CrossRef]

71. Pan, S.S.; Forrest, G.L.; Akman, S.A.; Hu, L.T. NAD(P)H:quinone oxidoreductase expression and mitomycin C resistance developed by human colon cancer HCT 116 cells. *Cancer Res.* **1995**, *55*, 330–335. [PubMed]

72. Asher, G.; Dym, O.; Tsvetkov, P.; Adler, J.; Shaul, Y. The crystal structure of NAD(P)H quinone oxidoreductase 1 in complex with its potent inhibitor dicoumarol. *Biochemistry* **2006**, *45*, 6372–6378. [CrossRef]

73. Fuchs, J.E.; Muñoz, I.G.; Timson, D.J.; Pey, A.L. Experimental and computational evidence on conformational fluctuations as a source of catalytic defects in genetic diseases. *RSC Adv.* **2016**, *6*, 58604. [CrossRef]

74. Robertson, A.D.; Murphy, K.P. Protein Structure and the Energetics of Protein Stability. *Chem. Rev.* **1997**, *97*, 1251–1268. [CrossRef] [PubMed]

75. Leaver-Fay, A.; Tyka, M.; Lewis, S.M.; Lange, O.F.; Thompson, J.; Jacak, R.; Kaufman, K.; Renfrew, P.D.; Smith, C.A.; Sheffler, W.; et al. ROSETTA3: An object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol.* **2011**, *487*, 545–574.

76. Alford, R.F.; Leaver-Fay, A.; Jeliazkov, J.R.; O'Meara, M.J.; DiMaio, F.P.; Park, H.; Shapovalov, M.V.; Renfrew, P.D.; Mulligan, V.K.; Kappel, K.; et al. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J. Chem. Theory Comput.* **2017**, *13*, 3031–3048. [CrossRef]

77. Kellogg, E.H.; Leaver-Fay, A.; Baker, D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins* **2011**, *79*, 830–838. [CrossRef] [PubMed]

78. Krieger, E.; Vriend, G. YASARA View—Molecular graphics for all devices—From smartphones to workstations. *Bioinformatics* **2014**, *30*, 2981–2982. [CrossRef]

79. Pandurangan, A.P.; Ochoa-Montaño, B.; Ascher, D.B.; Blundell, T.L. SDM: A server for predicting effects of mutations on protein stability. *Nucleic Acids Res.* **2017**, *45*, W229–W235. [CrossRef]

80. Bianchet, M.A.; Faig, M.; Amzel, L.M. Structure and mechanism of NAD[P]H:quinone acceptor oxidoreductases (NQO). *Methods Enzymol.* **2004**, *382*, 144–174.

81. Dehouck, Y.; Kwasigroch, J.M.; Gilis, D.; Rooman, M. PoPMuSiC 2.1: A web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinform.* **2011**, *12*, 151. [CrossRef]

82. Rodrigues, C.H.; Pires, D.E.; Ascher, D.B. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.* **2018**, *46*, W350–W355. [CrossRef]

83. Laimer, J.; Hiebl-Flach, J.; Lengauer, D.; Lackner, P. MAESTROweb: A web server for structure-based protein stability prediction. *Bioinformatics* **2016**, *32*, 1414–1416. [CrossRef]

84. Laimer, J.; Hofer, H.; Fritz, M.; Wegenkittl, S.; Lackner, P. MAESTRO–multi agent stability prediction upon point mutations. *BMC Bioinform.* **2015**, *16*, 116. [CrossRef]

85. Parthiban, V.; Gromiha, M.M.; Schomburg, D. CUPSAT: Prediction of protein stability upon point mutations. *Nucleic Acids Res.* **2006**, *34*, W239–W242. [CrossRef]

86. Sim, N.L.; Kumar, P.; Hu, J.; Henikoff, S.; Schneider, G.; Ng, P.C. SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* **2012**, *40*, W452–W457. [CrossRef]

87. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745–2747. [CrossRef] [PubMed]

88. Niroula, A.; Urolagin, S.; Vihinen, M. PON-P2: Prediction method for fast and reliable identification of harmful variants. *PLoS ONE* **2015**, *10*, e0117380. [CrossRef]

89. Schwarz, J.M.; Rodelsperger, C.; Schuelke, M.; Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* **2010**, *7*, 575–576. [CrossRef]

90. Vankova, P.; Salido, E.; Timson, D.J.; Man, P.; Pey, A.L. A dynamic core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer. *Biomolecules* **2019**, *9*, 728. [CrossRef]

91. Medina-Carmona, E.; Rizzuti, B.; Martin-Escolano, R.; Pacheco-Garcia, J.L.; Mesa-Torres, N.; Neira, J.L.; Guzzi, R.; Pey, A.L. Phosphorylation compromises FAD binding and intracellular stability of wild-type and cancer-associated NQO1: Insights into flavo-proteome stability. *Int. J. Biol. Macromol.* **2019**, *125*, 1275–1288. [CrossRef]

92. Megarity, C.F.; Timson, D.J. Cancer-associated variants of human NQO1: Impacts on inhibitor binding and cooperativity. *Biosci. Rep.* **2019**, *39*, BSR20191874. [CrossRef]

93. Scott, K.A.; Barnes, J.; Whitehead, R.C.; Stratford, I.J.; Nolan, K.A. Inhibitors of NQO1: Identification of compounds more potent than dicoumarol without associated off-target effects. *Biochem. Pharmacol.* **2011**, *81*, 355–363. [CrossRef] [PubMed]

94. Betancor-Fernandez, I.; Timson, D.J.; Salido, E.; Pey, A.L. Natural (and Unnatural) Small Molecules as Pharmacological Chaperones and Inhibitors in Cancer. *Handb. Exp. Pharmacol.* **2018**, *45*, 345–383.

95. Jacob, J.; Duclohier, H.; Cafiso, D.S. The role of proline and glycine in determining the backbone flexibility of a channel-forming peptide. *Biophys. J.* **1999**, *76*, 1367–1376. [CrossRef]

96. Morgan, A.A.; Rubenstein, E. Proline: The distribution, frequency, positioning, and common functional roles of proline and polyproline sequences in the human proteome. *PLoS ONE* **2013**, *8*, e53785. [CrossRef] [PubMed]

97. Eriksson, A.E.; Baase, W.A.; Zhang, X.J.; Heinz, D.W.; Blaber, M.; Baldwin, E.P.; Matthews, B.W. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* **1992**, *255*, 178–183. [CrossRef]

98. Xue, M.; Wakamoto, T.; Kejlberg, C.; Yoshimura, Y.; Nielsen, T.A.; Risor, M.W.; Sanggaard, K.W.; Kitahara, R.; Mulder, F.A.A. How internal cavities destabilize a protein. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 21031–21036. [CrossRef]

99. Pey, A.L. Towards Accurate Genotype–Phenotype Correlations in the CYP2D6 Gene. *J. Pers. Med.* **2020**, *10*, 158. [CrossRef]

100. Gersting, S.W.; Kemter, K.F.; Staudigl, M.; Messing, D.D.; Danecka, M.K.; Lagler, F.B.; Sommerhoff, C.P.; Roscher, A.A.; Muntau, A.C. Loss of function in phenylketonuria is caused by impaired molecular motions and conformational instability. *Am. J. Hum. Genet.* **2008**, *83*, 5–17. [CrossRef]

101. Rajasekaran, N.; Naganathan, A.N. A Self-Consistent Structural Perturbation Approach for Determining the Magnitude and Extent of Allosteric Coupling in Proteins. *Biochem. J.* **2017**, *474*, 2379–2388. [CrossRef]

102. Rajasekaran, N.; Sekhar, A.; Naganathan, A.N. A Universal Pattern in the Percolation and Dissipation of Protein Structural Perturbations. *J. Phys. Chem. Lett.* **2017**, *8*, 4779–4784. [CrossRef]

103. Rajasekaran, N.; Suresh, S.; Gopi, S.; Raman, K.; Naganathan, A.N. A General Mechanism for the Propagation of Mutational Effects in Proteins. *Biochemistry* **2017**, *56*, 294–305. [CrossRef]

MATERIAL SUPLEMENTARIO DE LA PUBLICACIÓN 1

Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein.

Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein.

Juan Luis Pacheco-García, Mario Cano-Muñoz, Isabel Sánchez-Ramos, Eduardo Salido and Angel L. Pey.

Supplementary material

Figure S1. Representative SDS-PAGE for proteolysis kinetics by thermolysin (at 1.2 µM protease). Time 0 samples contained no protease. The band corresponding to thermolysin (TH) is indicated.
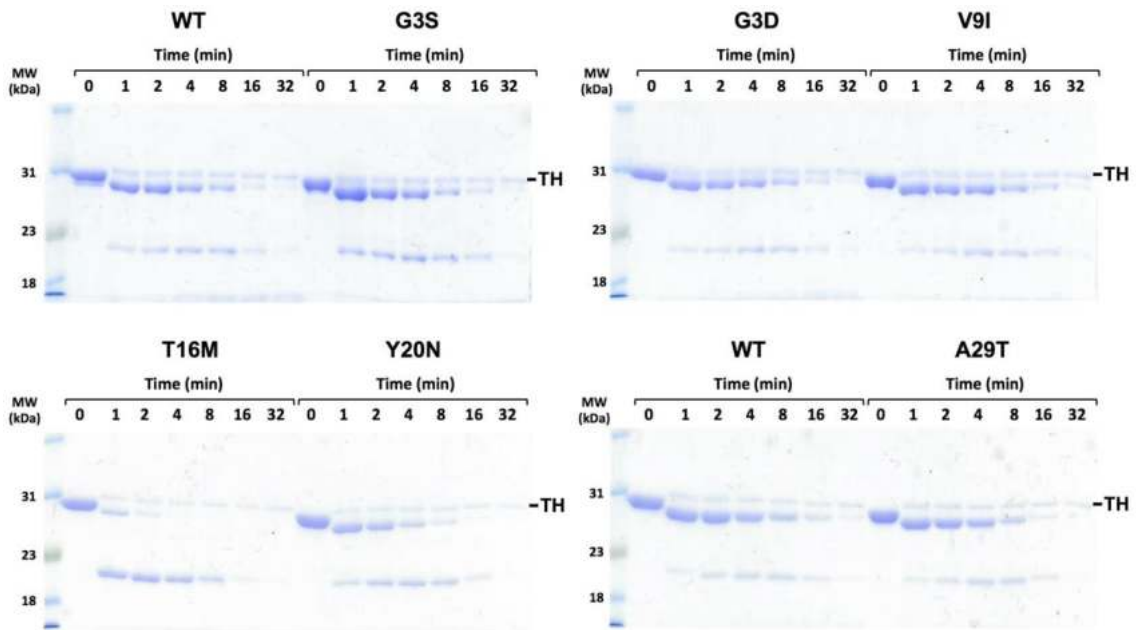
Figure S2. Kinetics of proteolysis of NQO1 proteins with thermolysin. NQO1 proteins were incubated with different protease concentrations and the time-course of cleavage was monitored by SDS-PAGE (see Figure S1). Lines represent best-fits to a single exponential to provide the values of k_{obs}.
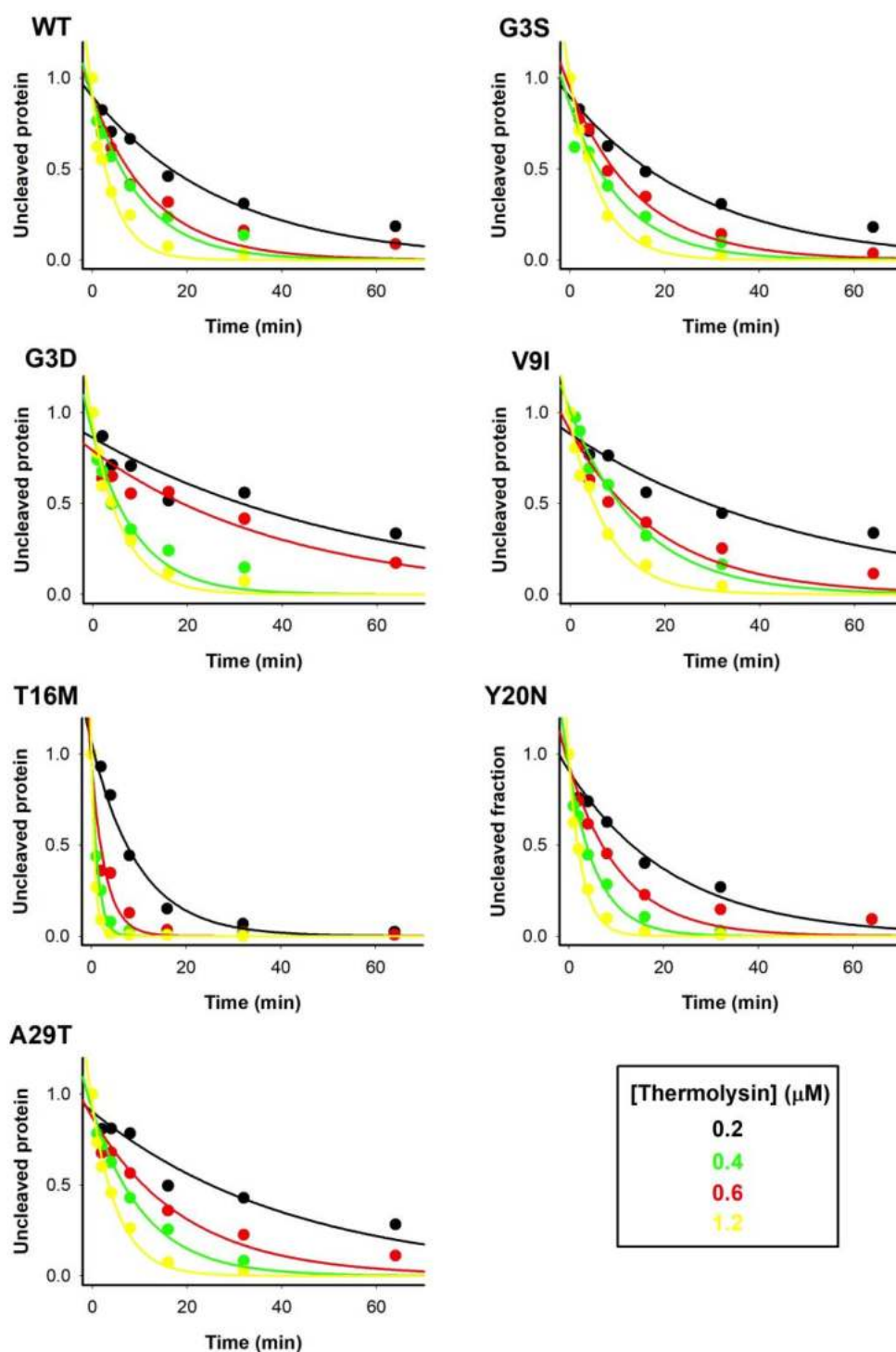
Figure S3. Titrations of apo-proteins with FAD. For each variant, data are from two titrations using two different preparations of apo-protein. Lines are fits to a single-type of binding sites model. Measurents were carried out in 1 x 0.3 cm quartz cuvettes, with excitation at 280 nm and emission at 350 nm. Slits were 5 nm. NQO1 concentration was 250 nM in monomer.
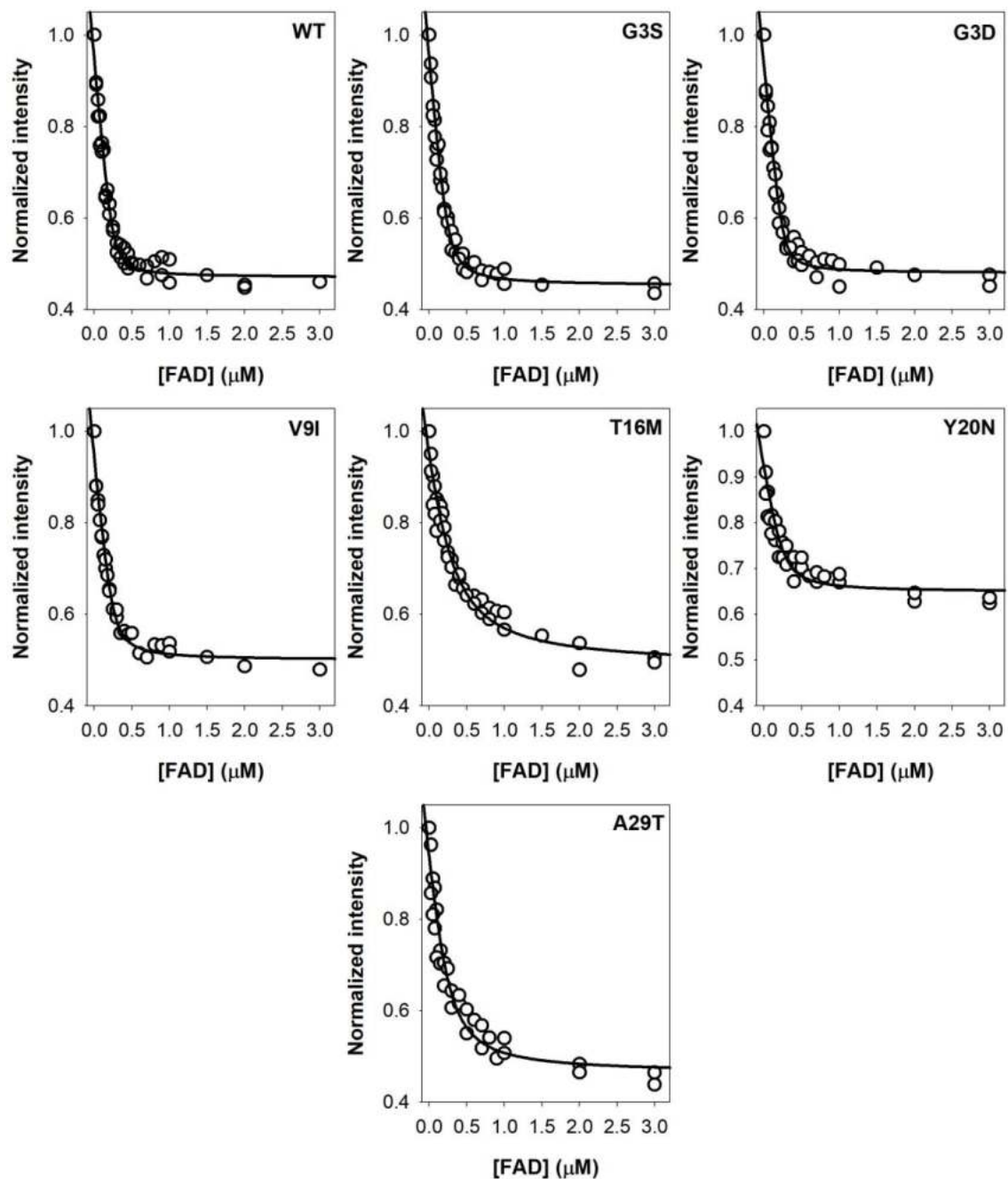
Figure S4. Titrations of NQO1 variants with Dic at 25ºC. Left panels show representative titrations. Experiments were carried out using 11-12 µM protein monomer and 120 µM Dic. 100 µM FAD was added to the protein and Dic solutions. Titrations consisted in one initial injection of 0.5 µL followed by 20-22 injections of 1.75 µL, spaced by 100 s. Right panels show integrated data as well as fittings to a single type of independent binding sites model. Dilution heats were determined as a fitting parameter and subtracted.
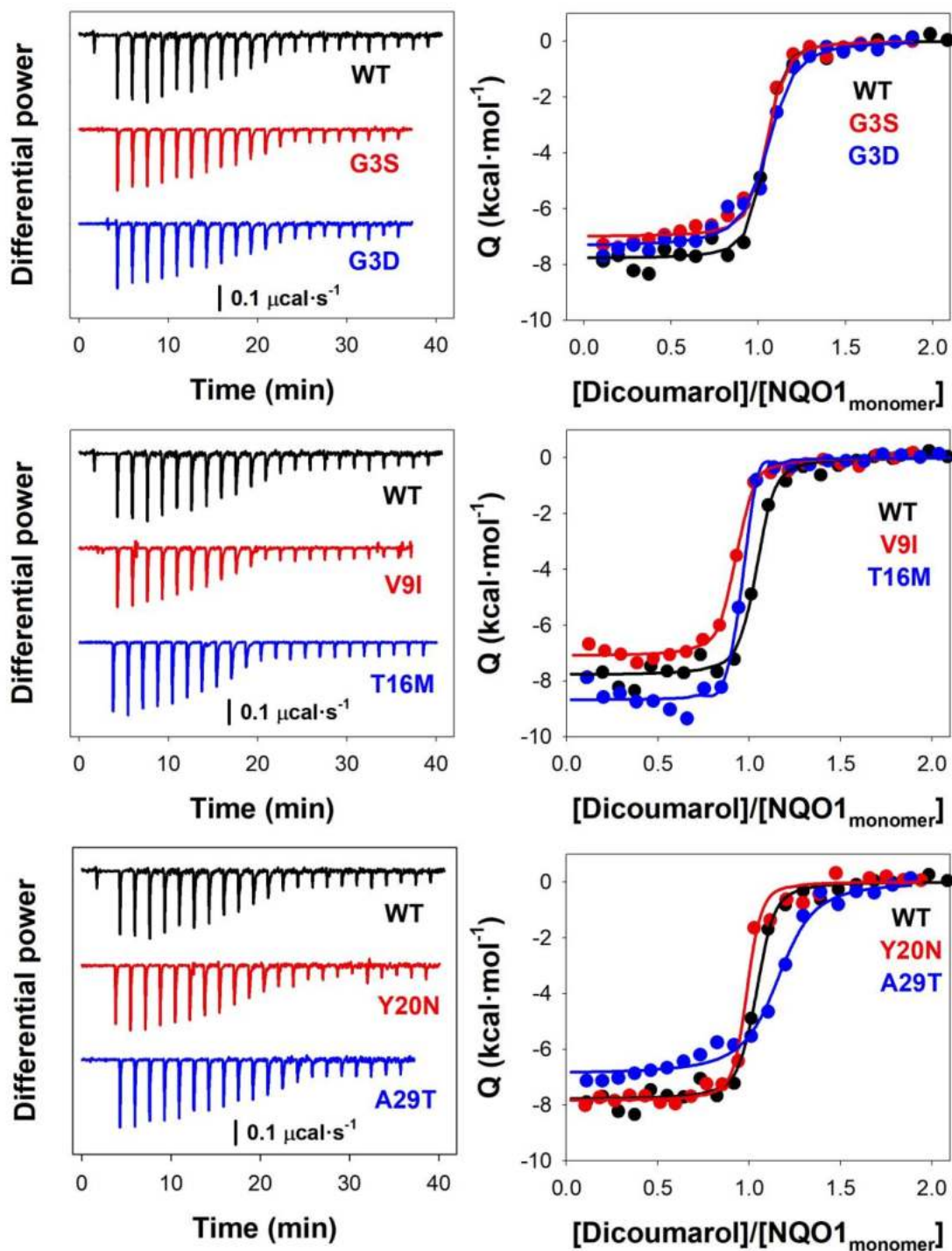
Figure S5. Thermodynamic dissection of Dic binding to NQO1 variants. A) Enthalpic and entropic contributions to binding (at 25°C; data are mean ±s.d. from at least three independent experiments). B) Temperature dependence of binding enthalpies. Data at 25°C are mean±s.d. from at least three independent experiments, whereas at other temperatures are from a single experiment and errors are those from fittings. C) Binding heat capacity changes (ΔC_p) determined from linear fittings of ΔH vs. Temperature (panel B). Errors are those from fittings.
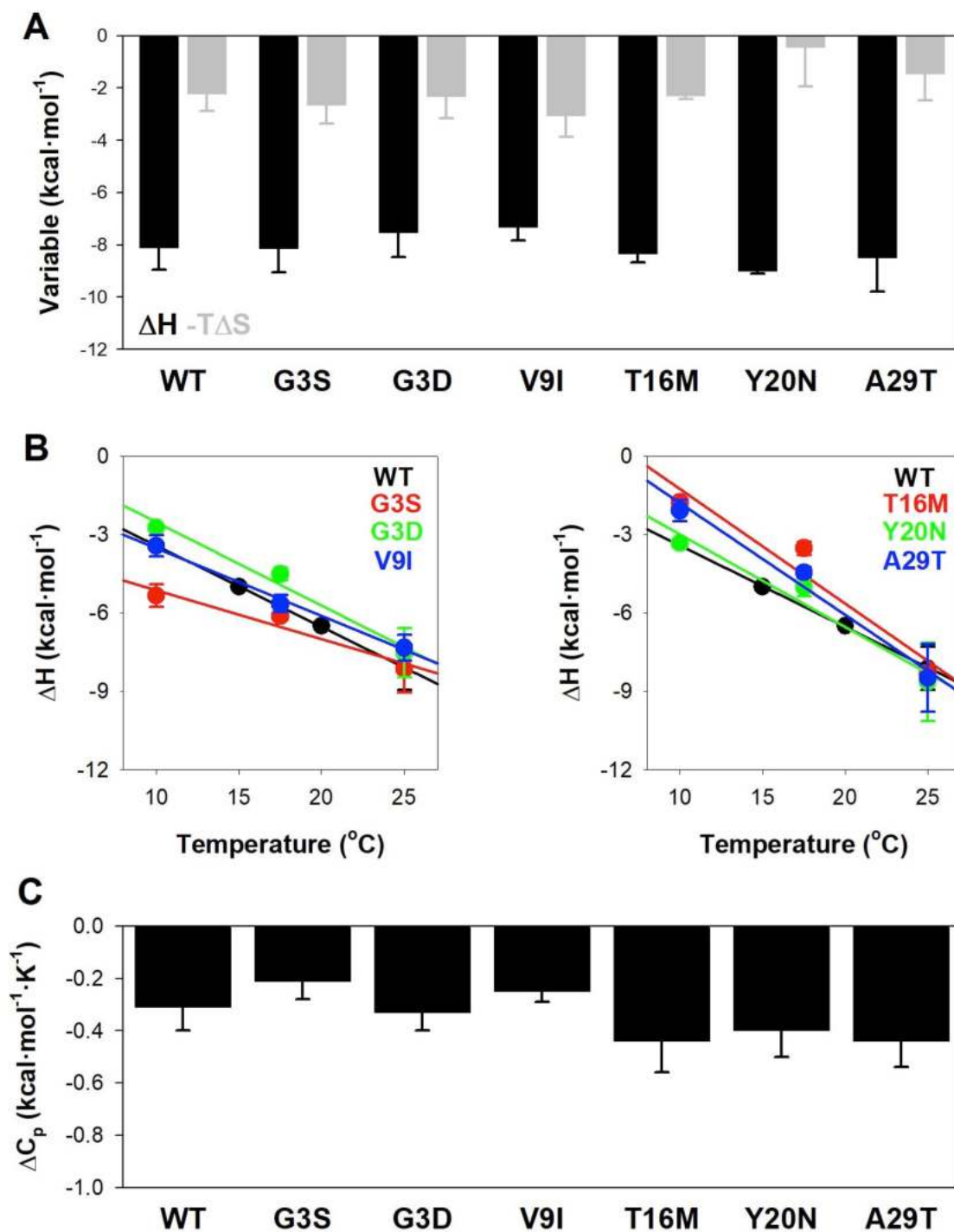
Figure S6. HDX analysis of structural regions in which NQO1 mutations are localized. A) Local stability parameters (derived from HDXMS studies on $NQO1_{apo}$, $NQO1_{holo}$ and $NQO1_{dic}$): A_{burst} (% of D incorporated in the dead time of HDXMS measurements), A_{slow} and k_{slow} are the % of D incorporated in a slower kinetic phase as well as the associated first-order rate constant. B) HDX kinetics for NQO1 segments in which the missense mutations are introduced. Plots are generated using data from [1].
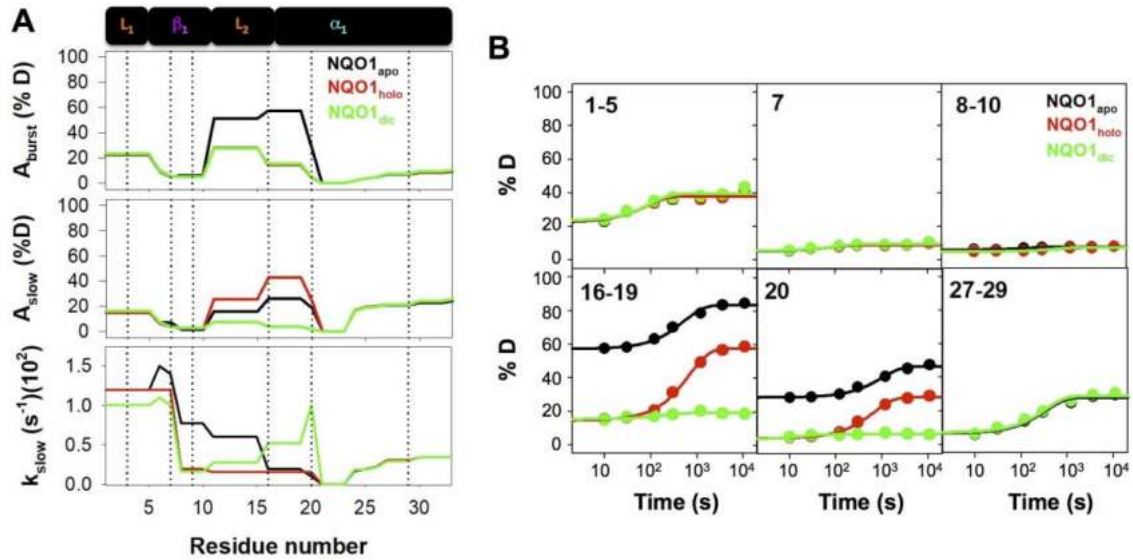
Table S1. Effect on conformational stability due to NQO1 mutations determined using structure-based methods. A negative value indicates a destabilizing effect.

Variant	Effect on stability (kcal·mol^{-1})					
	SDM	PopMusic	DynaMut	MAESTRO	CUPSAT	Average
G3S	-2.52±1.15	-0.88±0.09	-0.25±0.68	-0.16±0.11	-3.59±1.06	-1.48±0.95
G3D	-2.41±1.02	-1.03±0.08	-0.31±1.04	0.08±0.18	-3.33±1.49	-1.40±0.94
L7P	-4.01±0.01	-4.02±0.10	-1.88±0.24	-1.45±0.04	-10.3±2.0	-4.33±1.18
L7R	-3.08±0.10	-2.20±0.12	-1.56±0.35	-0.86±0.01	-2.88±0.61	-2.11±0.82
V9I	-0.57±0.01	-0.7±0.07	0.63±0.32	0.19±0.03	-1.65±0.31	-0.42±0.55
T16M	-0.09±0.01	-0.61±0.03	-0.11±0.06	-0.37±0.03	0.50±0.55	-0.14±0.21
Y20N	-1.31±0.58	-1.42±0.11	-1.47±0.31	-0.03±0.03	-1.46±0.52	-1.13±0.62
A29T	-2.81±0.08	-1.25±0.05	0.28±0.74	0.21±0.21	-2.85±0.80	-1.28±1.27
P187S	-1.19±0.05	-2.38±0.06	-0.52±0.28	-1.11±0.06	-4.58±0.52	-1.96±0.67
P187E	-0.99±0.03	-3.26±0.03	-0.36±0.42	-0.34±0.03	1.75±1.06	-0.24±1.55
P187R	-1.32±0.79	-2.30±0.08	0.31±0.25	-0.48±0.01	-1.36±0.43	-1.03±0.97
P187L	0.38±0.14	-0.87±0.07	1.49±0.40	0.32±0.01	-6.10±0.77	-0.96±0.83
P187A	-0.27±0.01	-2.22±0.04	-0.83±0.27	-0.38±0.05	-7.04±0.48	-2.15±0.78
P187G	-1.48±0.01	-2.43±0.05	-1.48±0.31	-0.65±0.01	-4.08±0.43	-2.02±0.63
P187T	-0.80±0.08	-1.68±0.03	-0.37±0.13	-0.43±0.01	-2.62±0.30	-1.18±0.52
K240Q	0.38±0.16	-0.60±0.06	-0.93±0.51	-0.22±0.01	0.78±0.32	-0.12±0.49
K240I	0.68±0.07	0.18±0.04	0.13±0.11	-0.19±0.02	0.54±0.49	0.27±0.31
K240E	0.09±0.09	-0.76±0.03	-0.72±0.25	-0.27±0.02	1.28±0.32	-0.08±0.35
K240T	0.09±0.24	0.56±0.04	-0.80±0.25	-0.29±0.02	0.61±0.32	0.03±0.50
K240H	0.47±0.06	-0.21±0.02	0.27±0.27	-0.17±0.01	0.62±0.39	0.20±0.29
K240A	0.62±0.03	-0.70±0.03	-0.41±0.25	-0.76±0.02	0.48±0.23	-0.15±0.55
K240G	0.61±0.06	-0.71±0.03	-0.90±0.11	-0.88±0.02	0.18±0.28	-0.34±0.63

Table S2. Second-order rate constants (k_{prot}) for proteolysis by thermolysin and mutational effects on local stability ($\Delta\Delta G_{prot}$).

Variant	k_{prot} (µM·min^{-1})	$\Delta\Delta G_{prot}$ (kcal·mol^{-1})
WT	0.151 ± 0.020	0.00 ± 0.11
G3S	0.146 ± 0.013	-0.02 ± 0.09
G3D	0.111 ± 0.032	-0.18 ± 0.19
V9I	0.101 ± 0.009	-0.24 ± 0.09
T16M	1.075 ± 0.115	1.16 ± 0.10
Y20N	0.292 ± 0.017	0.39 ± 0.09
A29T	0.153 ± 0.005	0.01 ± 0.08

Table S3. FAD binding affinity of NQO1 variants. Results are from fittings shown in Figure S3.

	$K_{d(FAD)}$ (nM)	$\Delta G_{(FAD)}$ (kcal·mol^{-1})	$\Delta\Delta G_{(FAD)}$ (kcal·mol^{-1})
WT	13.9±4.3	-10.72±0.18	---
G3S	23.2±5.4	-10.41±0.14	0.31±0.23
G3D	13.9±5.2	-10.72±0.22	0.00±0.25
V9I	23.3±6.5	-10.41±0.17	0.31±0.24
T16M	174±31	-9.22±0.11	1.50±0.22
Y20N	46.8±20.0	-10.00±0.25	0.72±0.26
A29T	80.8±25.7	-9.68±0.19	1.04±0.24

Table S4. Thermodynamic parameters for dicoumarol binding to NQO1 variants.
Data are mean±s.d. from 3-4 experiments at 25°C. ΔC_p values are determined from the linear dependence of binding enthalpies on temperature (Figure S5B).

	N	K_d (nM)	ΔG (kcal·mol^{-1})	ΔH (kcal·mol^{-1})	$-T\Delta S$ (kcal·mol^{-1})	ΔC_p (kcal·mol^{-1}·K^{-1})
WT	1.00±0.03	31±15	-10.33 ±0.35	-8.12±0.83	-2.21±0.67	-0.31±0.09
G3S	0.99±0.03	15±9	-10.80 ±0.36	-8.14±0.90	-2.66±0.71	-0.21±0.07
G3D	1.01±0.09	61±9	-9.84 ±0.09	-7.53±0.94	-2.31±0.84	-0.33±0.07
V9I	0.83±0.09	29±15	-10.38 ±0.30	-7.33±0.50	-3.06±0.80	-0.25±0.04
T16M	0.90±0.03	18±9	-10.65 ±0.34	-8.35±0.32	-2.29±0.12	-0.44±0.12
Y20N	0.98±0.09	67±40	-9.97±0.45	-8.64±1.49	-1.33±1.65	-0.37±0.11
A29T	0.97±0.11	60±25	-9.94±0.32	-8.49±1.30	-1.45±1.02	-0.44±0.10

Table S5. Magnitude of the conformational change (N_{conf}) associated with Dic binding from analysis of apparent binding enthalpies (ΔH) and heat capacities (ΔC_p). Expressed as number of residues.

	N_{conf} (ΔH)	N_{conf} (ΔC_p)
WT	10.8±3.9	12.3±6.5
G3S	10.9±4.2	5.1±5.1
G3D	8.0±4.4	13.8±5.1
V9I	7.1±2.3	8.0±2.9
T16M	11.9±1.5	21.7±8.7
Y20N	13.2±6.9	16.7±7.9
A29T	12.5±6.0	21.7±7.2

Table S6: **Energetic dissection of mutational effects as determined by Rosetta**. The energy of the protein is calculated using an energy function as a linear combination of several individual terms (provided as the energy terms in REF15). Values are shown in such a way that ΔΔG values > 0 are stabilizing and values < 0 are destabilizing for a given mutant vs. the WT protein.

Individual energy terms that sum up the ΔΔG value of mutation calculated for each variant.

Structure Form	Mutation	fa_atr:	fa_rep:	fa_sol:	fa_intra_rep:	fa_intra_sol:	lk_ball:	fa_elec:	hbond_sr_bb:	hbond_lr_b b:	hbond_bb_sc:	hbond_sc:	dslf_fa13:	omega:	fa_dun:	p_aa_pp:	yhh_planarity:	ref:	rama_prepro:	cart_bonded:	ΔΔG (kcal/mol)
Holo	G3S	2.45	-1.56	-3.23	-0.02	-0.10	-0.02	-0.96	0.00	-0.04	0.07	0.00	0.00	-0.08	-1.11	-0.55	0.00	1.09	0.13	-1.07	-5.02
	G3D	-2.85	-1.05	-7.82	-0.03	-0.38	1.29	-0.16	0.00	0.00	-0.05	0.01	0.00	-0.08	-2.57	-0.69	0.00	2.94	0.47	-0.68	-11.64
	L7P	-5.64	-6.93	-2.04	0.01	0.38	-0.42	-2.14	0.00	-1.84	-0.03	-0.06	0.00	-2.94	0.80	-0.22	0.00	3.31	-0.63	-10.98	-29.36
	L7R	7.26	-3.09	-15.83	-0.06	-0.82	0.23	0.44	0.00	0.03	0.05	0.25	0.00	-0.01	-1.91	-0.23	0.00	1.76	-0.05	-2.08	-14.09
	V9I	3.86	-1.68	-0.52	-0.02	-0.06	0.12	-0.03	0.00	0.04	0.02	0.00	0.00	-0.03	-0.20	0.05	0.00	0.34	0.09	-0.11	1.86
	T16M	7.39	-0.93	-1.67	-0.04	-0.31	-0.39	-0.21	0.06	0.00	-0.47	0.24	0.00	-0.23	-3.54	-0.72	0.02	-0.50	0.24	0.23	-0.82
	Y20N	-6.41	0.89	-1.80	0.02	-0.22	1.79	-0.42	-0.03	0.00	-0.32	-1.75	0.00	0.16	-1.30	-0.91	0.14	1.92	-0.18	0.36	-8.06
	A29T	4.74	-4.36	-7.36	-0.09	-0.07	0.37	1.67	-0.29	0.02	-0.01	0.08	0.00	-0.02	-0.29	-0.43	0.00	0.17	-0.68	-4.41	-10.94
Dicoumarol bound	G3S	3.96	-1.00	-4.50	-0.01	-0.09	0.35	0.21	0.00	-0.03	-0.33	0.00	0.00	-0.12	-0.12	-0.25	0.00	1.09	-0.18	-0.46	-1.48
	G3D	0.86	-2.44	-5.73	-0.03	-0.43	0.67	-1.54	0.00	0.11	-2.69	0.25	0.00	-0.29	-1.05	-0.62	0.00	2.94	0.67	-0.80	-10.10
	L7P	-2.64	-6.20	-4.31	0.11	-1.07	-0.64	-1.08	0.00	-1.79	0.58	2.08	0.00	-1.54	-0.17	0.33	0.00	3.31	-0.41	-12.02	-25.47
	L7R	12.52	-3.00	-24.64	0.07	-0.72	0.85	4.73	0.00	-0.09	0.15	3.64	0.00	-0.20	-1.74	-0.13	0.00	1.76	0.12	-29.26	-35.94
	V9I	2.89	-1.78	-0.30	0.08	0.57	0.68	-0.15	0.00	0.11	-1.76	0.14	0.00	0.00	-0.36	0.13	0.00	0.34	0.16	-0.10	0.64
	T16M	12.85	-2.80	-4.46	-0.08	-0.10	-0.15	1.10	0.01	0.00	0.99	-0.05	0.00	-0.07	-4.21	-0.78	0.00	-0.50	0.39	-1.34	0.82
	Y20N	-7.25	0.85	-1.59	0.12	0.98	1.46	-0.35	-0.12	0.00	-0.69	0.08	0.00	0.14	-0.61	-0.58	0.17	1.92	-0.14	0.21	-5.39
	A29T	5.91	-1.50	-6.30	-0.09	-0.05	-0.55	4.15	0.38	-0.11	-0.98	2.27	0.00	-0.03	-0.04	0.09	0.00	0.17	-0.25	-2.67	0.41

Table S7. Analysis of steric clashes present in the NQO1holo and NQO1dic states caused by the mutation L7P. No clashes were found in L7 for the models of NQO1 WT.

State	Residue 1	Atom pair 1	Residue 2	Atom pair 2	Distance (Å)
NQO1holo	R7	NH1	V9	CG2	2.87
	R7	NH2	E88	OE1	2.65
	R7	CZ	W116	HZ2	2.16
NQO1dic	R7	NH2	E88	OE1	2.15
	R7	NH1	L92	HG	2.29
	R7	CZ	W116	HZ2	2.37
	R7	NH1	K91	HB	2.51

Table S8. Hydrophobic interactions established between V9 (WT) and I9 (mutant V9I) in the NQO1holo and NQO1dic states

State	Variant	Residue 1	Atom pair 1	Residue 2	Atom pair 2	Distance (Å)
NQO1holo	WT	V9	CG1	L113	CD1	3.52
		V9	CG2	L7	CG	3.83
		V9	CB	F100	CD2	3.85
	V9I	I9	CG2	L113	CD1	3.55
		I9	CG1	L7	CG	3.80
		I9	CB	F100	CD2	4.17
		I9	CD1	L7	CD2	3.84
		I9	CD1	V98	CG1	3.83
		I9	CD1	W116	CZ2	3.67
		I9	CD1	F100	CD2	3.72
NQO1holo	WT	V9	CG1	L113	CD1	3.56
		V9	CG2	L7	CG	3.82
	V9I	I9	CG2	L113	CD1	3.64
		I9	CG1	L7	CG	3.84
		I9	CD1	W116	CZ2	3.56
		I9	CD1	W116	CZ3	3.87
		I9	CD1	F100	CD2	3.77
		I9	CD1	V98	CG1	3.78

Table S9. Regions with < 20% HDX after 3h. Data from [1]. These regions are considered as *highly stable*.

State	Segments
NQO1apo	6-10, 21-24, 97-102, 114-119, 142-145, 166-185, 206-211.
NQO1holo	6-10, 21-24, 97-102, 114-119, 142-145, 166-185, 206-211.
NQO1dic	6-10, 16-24, 88-90, 97-124, 142-145, 166-185, 206-211.

Table S10. Regions with Δ%D$_{av}$ ≥ 10%. This parameter [1] is calculated by comparing the HDX kinetics of NQO1$_{holo}$ vs. NQO1$_{apo}$ and NQO1$_{dic}$ vs. NQO1$_{holo}$. Thus, these represent the stabilization induced on WT NQO1 upon binding FAD and dicoumarol, respectively.

States	Δ%D$_{av}$	Segments
NQO1$_{holo}$ vs. NQO1$_{apo}$	10-25	20, 46-68, 75-76, 133, 149-155, 157-165, 191-201, 206, 211, 262-271.
	25-50	11-19, 69-74, 103-113, 202-205.
	> 50	---
NQO1$_{dic}$ vs. NQO1$_{holo}$	10-25	16-20, 46-54, 56-68, 75-76, 107-113, 120-121, 149-155, 229-238.
	25-50	69-74, 103-106, 128-140.
	> 50	122-127.

References

1.   Vankova, P.; Salido, E.; Timson, D.J.; Man, P.; Pey, A.L. A dynamic core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer. *Biomolecules* 2019, *9,* 728.

PUBLICACIÓN 2

Counterintuitive structural and functional effects due to naturally occurring mutations targeting the active site of the disease-associated NQO1 enzyme.

Counterintuitive structural and functional effects due to naturally occurring mutations targeting the active site of the disease-associated NQO1 enzyme*

Juan Luis Pacheco-García[1], Ernesto Anoz-Carbonell[2], Dmitry S. Loginov[3], Daniel Kavan[3], Eduardo Salido[4], Petr Man[3], Milagros Medina[2] and Angel L. Pey[5] (ID)

1 Departamento de Química Física, Universidad de Granada, Spain
2 Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias, Instituto de Biocomputación y Física de Sistemas Complejos (BIFI) (GBsC-CSIC Joint Unit), Universidad de Zaragoza, Spain
3 Institute of Microbiology – BioCeV, Academy of Sciences of the Czech Republic, Vestec, Czech Republic
4 Center for Rare Diseases (CIBERER), Hospital Universitario de Canarias, Universidad de la Laguna, Tenerife, Spain
5 Departamento de Química Física, Unidad de Excelencia en Química Aplicada a Biomedicina y Medioambiente e Instituto de Biotecnología, Universidad de Granada, Spain

Our knowledge on the genetic diversity of the human genome is exponentially growing. However, our capacity to establish genotype–phenotype correlations on a large scale requires a combination of detailed experimental and computational work. This is a remarkable task in human proteins which are typically multifunctional and structurally complex. In addition, mutations often prevent the determination of mutant high-resolution structures by X-ray crystallography. We have characterized here the effects of five mutations in the active site of the disease-associated NQO1 protein, which are found either in cancer cell lines or in massive exome sequencing analysis in human population. Using a combination of H/D exchange, rapid-flow enzyme kinetics, binding energetics and conformational stability, we show that mutations in both sets may cause counterintuitive functional effects that are explained well by their effects on local stability regarding different functional features. Importantly, mutations predicted to be highly deleterious (even those affecting the same protein residue) may cause mild to catastrophic effects on protein function. These functional effects are not well explained by current predictive bioinformatic tools and evolutionary models that account for site conservation and physicochemical changes upon mutation. Our study also reinforces the notion that naturally occurring mutations not identified as disease-associated can be highly deleterious. Our approach, combining protein biophysics and structural biology tools, is readily accessible to broadly increase our understanding of genotype–phenotype correlations and to improve predictive computational tools aimed at distinguishing disease-prone against neutral missense variants in the human genome.

Abbreviations
CD, circular dichroism; CTD, C-terminal domain; DBS, dicoumarol binding site; DCPIP, 2,6-dichlorophenolindophenol; Dic, dicoumarol; DLS, dynamic light scattering; FAD, flavin-adenine dinucleotide; FBS, FAD binding site; HDX, hydrogen/deuterium exchange; HDX-MS, hydrogen/deuterium exchange monitored by mass-spectrometry; HT, hydride-transfer; K_d, dissociation constant; k_{HT}, limiting k_{obs} at [NADH] → ∞; k_{obs}, observed rate constant; LC–MS/MS, liquid chromatography mass-spectrometry/mass-spectrometry; MMI, monomer:monomer interface; NQO1, NAD(P)H:quinone oxidoreductase 1; NQO1$_{dic}$, NQO1$_{holo}$ with Dic bound; NQO1$_{holo}$, NQO1 with FAD bound; NQO1$_{red}$, NQO1 with FADH$_2$ bound; NTD, N-terminal domain; SDS/PAGE, polyacrylamide gel electrophoresis in the presence of sodium dodecylsulphate; SEC, size-exclusion chromatography; WT, wild-type.

Introduction

The huge advances in DNA sequencing technologies have uncovered a tremendous genetic diversity in human population. The next step is to be able to distinguish between pathogenic and neutral variants [1–3]. To do so, it is necessary to integrate experimentation to feed and improve computational tools able to carry out accurate and large-scale prediction of mutational effects and genotype–phenotype correlations [4].

The NAD(P)H:quinone oxidoreductase 1 (NQO1) protein is associated with common diseases such as cancer, Alzheimer's, and Parkinson's disease [5]. NQO1 catalyses the FAD-dependent reduction of a large set of quinone substrates, including redox maintenance of vitamins, detoxification of xenobiotics, activation of cancer pro-drugs and regulation of the NADH/NAD$^+$ ratio [6,7]. A schematic representation of the NQO1 catalytic cycle is displayed in Fig. 1A. NQO1 is a dimeric and two-domain enzyme: the N-terminal domain (NTD, residues 1–225) contains the FAD binding site (FBS) and most of the active site residues, while the C-terminal domain (CTD, residues 225–274) completes the active site and the monomer:-monomer interface (MMI) [8–12]. Dicoumarol (Dic) is a tight competitive inhibitor of NADH and the substrate [13], for which high-resolution structural information for its binding to NQO1 is available by X-ray crystallography [14]. We have recently shown that ligand binding (FAD and Dic) and mutational effects propagate to long distances in the native state ensemble of NQO1, potentially affecting different functional features in counterintuitive fashions [12,15–20]. Therefore, NQO1 represents a biomedically relevant and challenging system to compare the performance of computational and experimental methods to explain and predict genotype–phenotype on a large scale for a multi-functional protein.

The catalytic competence of NQO1 may be altered upon mutation and affect different functional features: FAD binding, the oxido-reduction reaction pathways, enzyme cooperativity and quantum (tunnelling) effects in the hydride-transfer (HT) reaction [12,19]. As mentioned above, mutational and ligand binding effects may alter the stability of functional sites across the entire protein structure (in some cases over 30 Å) [12,21–23]. In this work, we have studied the effects of five naturally occurring mutations (p.W106R, p.W106C, p.F107C, p.M155I and p.H162N) in the active site of NQO1 affecting four different residues (Fig. 1A). Inspection of a structural model of NQO1$_{dic}$ (with FAD and Dic bound) indicates that all mutations may affect the interactions with the cofactor, NAD(P)H and/or the substrate and thus, the

capacity of the enzyme to achieve catalytically competent states (Fig. 1B,C). W106 is close to the FAD flavin ring (2.8 Å for flavin ring N5 to backbone N of W106) and packs against a hydrophobic pocket from the adjacent monomer (Fig. 1B,C), while F107 is also close to the FAD (2.8 Å for flavin ring O4 to backbone N of F107). Thus, mutations p.W106R and p.W106C would affect the stability of the active site (either by introducing a positive charge and a cavity, respectively, in a hydrophobic environment), while mutations at both W106 and F107 might also influence the flavin redox potential. All the mutated residues are strictly conserved in mammalian NQO1 sequences (Fig. 1C) with different degrees of conservation in the overall sequence (down to 78.5% in sequence identity vs. human NQO1). Interestingly, NQO1 sequences from reptiles (with a slightly lower identity, 66.7–70.3%) show a consistent divergence at the H162 site (actually to the active-site mutation p.H162N; Sequence IDs: XP_033014806.1; XP_044283688.1, XP_005310627.1 and XP_044843791.1). Three of these rare mutations in the human NQO1, [p.W106C (c.318G>C), p.M155I (c.465G>A) and p.H162N (c.484C>A)] are found in one or two cancer somatic cell lines in the COSMIC database (accessed by 29 June 2022; https://cancer.sanger.ac.uk/cosmic/gene/analysis?ln = NQO1#variants), whereas the two remaining rare mutations [p.W106R (c.316T>C) and p.F107C (c.320T>G)] appear in the human population with no clear association with disease (with allelic frequencies of 2.8×10^{-5} and 1.8×10^{-5}, respectively; according to The Genome Aggregation Database, gnomAD, accessed by 29 June 2022; https://gnomad.broadinstitute.org/gene/ENSG00000181019?dataset = gnomad_r2_1,). Based on different bioinformatic predictive tools, all mutations are expected to be largely deleterious except p.H162N (Table 1). Since the mutated sites are highly conserved, this set of naturally occurring mutations in the active site of NQO1 should affect function mostly due to the changes in charge, polarity and hydrophobicity [24].

Results and Discussion

All active-site variants retain the fold of the wild-type protein but some affect flavin-adenine dinucleotide content

The WT and active-site mutants were purified by IMAC to a high purity (Fig. 2A). None of the mutations had large effects on the expression levels as judged from NQO1 protein yields (between 50% and 70% of WT levels; Fig. 2A). Inspection of the near UV–visible absorption spectra indicated that all variants contained FAD bound, but the active-site
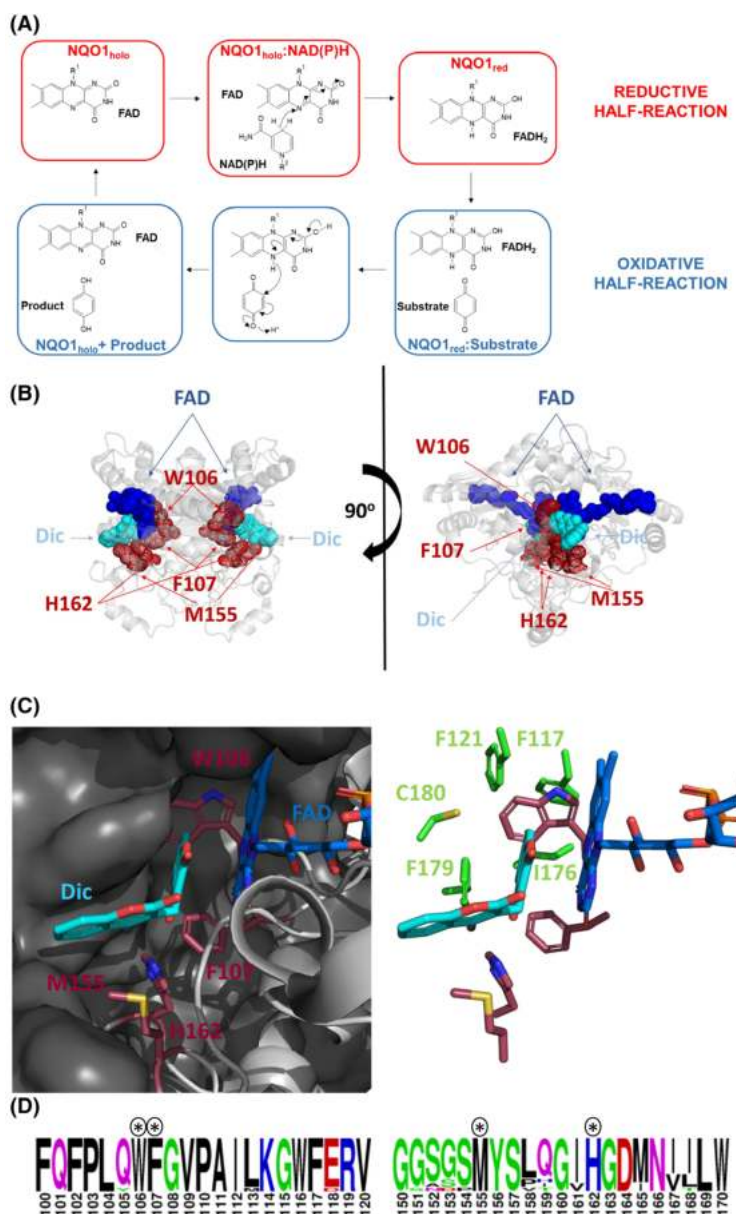
Fig. 1. Structural location and sequence conservation of the residues W106, F107, M155 and H162 in the active site of NQO1. (A) Simplified catalytic cycle of NQO1. The reductive half-reaction (in red) begins with FAD bound to NQO1 (NQO1$_{holo}$) that binds the NAD(P)H coenzyme, and likely through a direct HT reduces the FAD to FADH$_2$ (NQO1$_{red}$). In the oxidative half-reaction, the enzyme binds the substrate that becomes reduced upon oxidation of the FADH$_2$, and upon release of the product regenerates the NQO1$_{holo}$ state. Adapted from Ref. [13]. (B) The display, based on PDB 2F1O [14], shows the mutated residues (in red) and the FAD (in blue) and Dic (in cyan) molecules in human NQO1, to highlight the potential importance of mutations in the active site performance. The two views are from the same display just rotated by 90°. (B) Close view of the NQO1 active site. Mutated residues are shown in CPK coloured sticks (in red). The NQO1 monomer binding the shown FAD within the dimer is in the light grey cartoon, while the neighbouring monomer is in the dark grey surface. The right panel shows a similar display but highlights the residues (in green sticks) of the adjacent monomer that forms the hydrophobic pocket allocating W106. (C) Conservation of these four active site residues among mammalian protein sequences (highlighted with an asterisk). Segments include residues 100–120 and 150–170 of the human sequence. Thirteen sequences were used including different orders. Primates (96.7–100% identity with hNQO1; IDs: NP_000894, NP_001125152.1, XP_030650486.1 and NP_001247927.1); Artiodactyla, (89.4–92.3% identity; IDs: XP_006203814.1, XP_010977513.1, NP_001029707.1 and NP_001153085.1); others (carnivora, Chiroptera, Perissodactyla, Didelphimorphia and Rodentia; 78.1–87.5% identity; IDs:XP_038523128.1, XP_006909876.1, XP_005608432.1, XP_001378228.2 and XP_005345766.1). Sequences were retrieved using BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) and the human sequence as query. Panel C was created using WebLogo (https://weblogo.berkeley.edu/logo.cgi).

Table 1. Prediction of the pathogenicity of active-site mutations in NQO1 experimentally characterized in this work. Original scores (OS) were provided by different tools for each variant (in parenthesis) and were normalized (normalized score NS, from 0, neutral, to 1, highly deleterious) using the procedures indicated in the footnotes. The average NS for each variant is reported as NS_{pred} in the *consensus* prediction (and classified as mild, moderate or severe).

Prediction → Variant ↓	POLYPHEN-2 (OS)[a]	SIFT (OS)[b]	CADD (OS)[c]	MUT.ASSESSOR (OS)[a]	REVEL (OS)[a]	METAL R (OS)[a]	CONSENSUS (NS_{pred})
p.W106C (COSV57731335)	Prob.Dam. (1.00)	Deleterious (0.00)	Like. Delet. (31)	High (0.965)	Like. Disease (0.635)	Tolerated (0.198)	Severe (0.69)
p.W106R (rs769655604)	Prob.Dam. (1.00)	Deleterious (0.00)	Like. Delet. (32)	High (0.983)	Like. Disease (0.778)	Tolerated (0.203)	Severe (0.72)
p.F107C (rs760765580)	Prob.Dam. (1.00)	Deleterious (0.01)	Like. Benign (28)	Medium (0.909)	Like. Disease (0.522)	Tolerated (0.135)	Severe (0.64)
p.M155I (COSV57732645)	Prob.Dam. (0.966)	Deleterious (0.04)	Like. Benign (25)	Medium (0.574)	Like. Benign (0)	Tolerated (0.061)	Moderate (0.48)
p.H162N (COSV57731604)	Benign (0.015)	Tolerated (0.85)	Like. Benign (16)	Neutral (0.24)	Like. Benign (0)	Tolerated (0.015)	Mild (0.07)

[a]OS were used as NS from 0 (Benign, Neutral, Likely Benign or Tolerated) to 1 (Probably Damaging, High, Likely Disease or Not tolerated); [b]OS were normalized as binary (Tolerated, 0; Deleterious, 1); [c]OS were linearly normalized from 0 (Benign, Like. Benign, OS 0) to 1 (Deleterious, Like. Delet. original score 99).

mutants reduced this content compared with that of the WT protein (except p.F107C) (Fig. 2B,C). Alterations in the microenvironment of the FAD bound to some of the active-site variants are supported by differences in the near-UV CD spectra of samples saturated with FAD, particularly for the variant p.W106R (Fig. 2B). Additional analyses by circular dichroism (CD) and fluorescence spectroscopies and dynamic light scattering (DLS) supported only mild alterations in the overall conformation of NQO1 upon mutation (Fig. 2D–F). While the secondary structure was not significantly affected (Fig. 2D), the mutation p.W106R increased the fluorescence intensity by about 30% without affecting the shape of the spectra (with an average SCM for all variants of 354.03 ± 0.16 nm). Since Tyr residues are not fluorescent in this range, we suggest that this might originate from either local changes in the microenvironment of the remaining five Trp residues or in the quenching induced by FAD binding. Analysis of the hydrodynamic behaviour revealed some interesting changes, particularly the increased width of the size distribution and the average hydrodynamic radius of p.M155I and p.H162N. We must note that these results are consistent with an increase in hydrodynamic volume of about 15% and 30%, for the p.M155I and p.H162N mutants, respectively. This expanded conformation caused by the p.H162N mutation (in the holo-state) is similar in magnitude to that observed in the WT holo-protein upon withdrawal of FAD [22,25] and may imply an increased population of partially folded states in the native ensemble of holo-p.H162N.

The mutants p.W106R and p.W106C

W106 is involved in the MMI, the FBS and the Dic binding site (DBS). W106 is buried in the structure of WT NQO1 determined by X-ray crystallography ($9.9 \pm 0.6\%$ of accessible surface area, ASA, without considering ligands bound and using the six monomers of PDB 2F1O [14] and the software GETAREA, http://curie.utmb.edu/getarea.html, [26]). The mutations p.W106R and p.W106C are expected to cause large structural perturbations, the former due to the introduction of a positive charge and the latter by creating a cavity in a hydrophobic and buried region of NQO1 (see Ref. [19] for the characterization of some non-natural cavity-making mutations in NQO1).

We first compared the FAD binding affinity of p.W106R and p.W106C with that of WT NQO1 (Fig. 3A,C) by fluorescence titrations. These titrations indicated a much lower affinity in the mutant p.W106R, while the mutant p.W106C showed a similar K_d value for FAD binding than that of the WT protein. Titrations using CD spectroscopy (Fig. 3B) allowed to estimate of a K_d for FAD binding of 7400 ± 4000 nM (about 500-fold higher than that of WT NQO1). To the best of our knowledge, this mutant shows the lowest affinity for FAD ever reported for a single NQO1 missense variant, decreasing by 3.7 kcal·mol^{-1} the apparent binding free energy.

We have determined the changes in the structural stability of the protein caused by the p.W106R and p.W106C mutations using HDX-MS (Figs S1, S2 and S4). We first looked at the effects on the entire protein. The effects are remarkably different for both
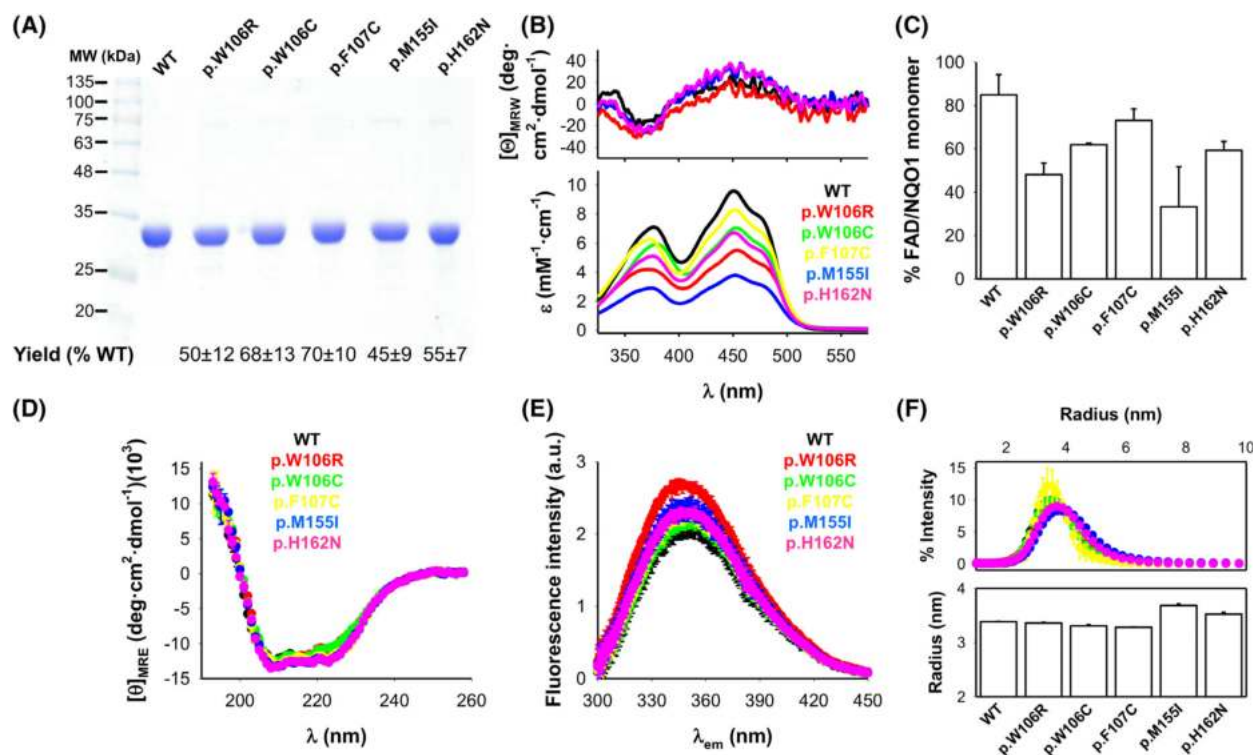
Fig. 2. Purification, conformation and bound FAD to NQO1 WT and the active-site mutants. (A) SDS/PAGE (12% acrylamide) of purified NQO1 variants. In each lane, 5 µg of protein was loaded; NQO1 protein yields (as mg NQO1 protein·L^{-1} cultures) are also shown as average ± SD from three different protein expression experiments. The amount of NQO1 was determined from UV–visible absorption measurements. (B, C) FAD bound to purified NQO1 proteins by UV–visible absorption and CD spectra. Lower panel B shows the average absorption spectra for each NQO1 variant and the upper panel shows CD spectra for selected variants. Panel C shows the average ± SD of the % of NQO1 monomer with bound FAD. (D) Far-UV CD spectra (average ± SD) for NQO1 variants with a fivefold excess of FAD; (E) emission fluorescence spectra (exc. 280 nm) of NQO1 variants in the presence of fivefold excess of FAD. (F) Hydrodynamic behaviour (by DLS) of NQO1 variants in the presence of a fivefold excess of FAD. The upper panel shows the distribution of the apparent radius corresponding to the NQO1 dimer and the lower panel shows the average hydrodynamic radius. Errors correspond to the SD from three different purifications (panel C) or replicate (panels D–F). Errors in panel B are not displayed for sake of display, but representative relative values (and SD) are shown in panel C.

mutants when we analyse them in the holo- and dic-states (Figs 4A–C, 5A–C and 6A; Figs S5A–C, S6A–C, S7A–C, S8A–C, S9A–C and S10A–C). In the NQO1$_{holo}$ state, the p.W106R mutation caused some local and mild stabilizing/destabilizing effects, whereas p.W106C mildly destabilized the entire structure (Fig. 6A). In the NQO1$_{dic}$ state, destabilization by p.W106R was much more extensive and stronger, affecting almost the whole NTD (40% of the residues of the entire protein; Figs 5B and 6A), whereas p.W106C caused virtually no effects (Fig. 6A). Since NTD is critical for enzymatic function [10,25], it is likely that mutation p.W106R will affect other functional features.

We then analysed the effects of p.W106R on the stability of different functional sites: the MMI, involved in the conformational stability of the protein and communication between monomers during the

catalytic cycle [12,13,20,27], the FBS, essential in the redox reaction and in close contact with W106 and the DBS, where the Dic inhibitor likely occupies partly the NADH/substrate binding site and/or represents a transition state analogue and interacts with W106 [13]. The results obtained paralleled those found for the entire protein. The mutation p.W106R mild–moderately affected the stability of the MMI, FBS and DBS in the NQO1$_{holo}$ state, and these effects were much stronger upon Dic binding (Fig. 6B–D). The mutation p.W106C mildly affected the MMI, FBS or DBS in the NQO1$_{holo}$ state and these effects were essentially abolished in the NQO1$_{dic}$ state (Fig. 6B–D). Thus, the effects of p.W106R and p.W106C are consistent with the large effect of the former mutation on FAD binding affinity and also predict much stronger effects for it on enzyme kinetics than those of p.W106C.
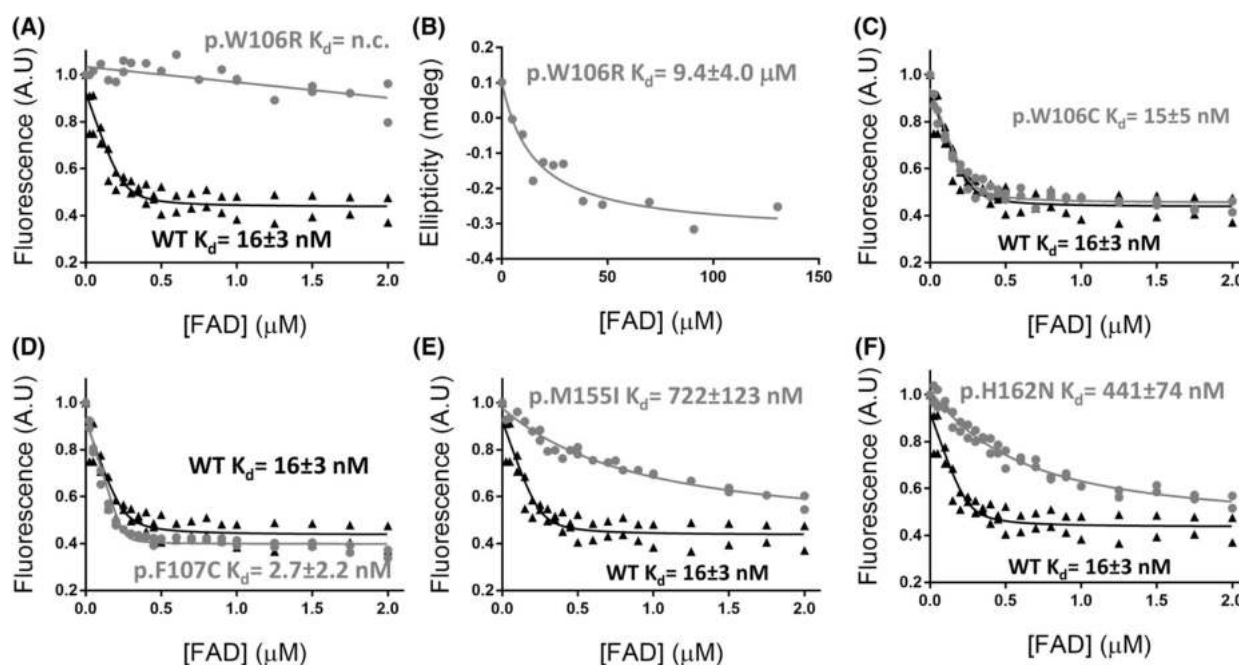
Fig. 3. FAD binding affinity of WT NQO1 and the active site mutants. Titrations of apo-proteins with FAD determined by fluorescence (A, C–F) or circular dichroism (B) spectroscopies. Titrations were carried out in 20 mM K-phosphate, pH 7.4. The temperature was 25 °C. Data came from at least two different titrations, except for panel B (single titration). Lines are best fits for a single type of independent binding sites. N.c., not converged.

The catalytic cycle of NQO1 can be divided into two steps: the reductive and oxidative half-reactions (Fig. 1A), being the former the rate-limiting in the WT enzyme [13]. In addition, pre-steady-state enzyme kinetic analysis of WT NQO1 has revealed the existence of two different pathways, termed *fast* and *slow*, for the reduction of the two FAD molecules within the NQO1$_{holo}$ dimer by NADH, as well as for oxidation of the FADH$_2$ in NQO1$_{red}$ by DCPIP [12,13,19] (Tables 2 and 3; Figs S11A, S12A, S13A, S14A, S15A and S16A). When we compared the kinetics of p.W106R and p.W106C mutants for these two-half reactions, their behaviours were clearly different (Tables 2 and 3; Figs S11B,C, S12B,C and S13B,C). The mutant p.W106R largely perturbed two aspects in the catalytic cycle: (a) The reductive half-reaction of the FAD was about 6000-fold slower for the mutant p.W106R; (b) there is an alteration in the reaction equilibrium and full reduction of the FAD cofactor cannot be achieved. Thus, DCPIP-mediated reoxidation of FAD cannot be measured and this prevents the completion of the catalytic cycle (i.e. reduction of a proper substrate) in the p.W106R mutant. In addition, the spectral properties of species B and C were different in p.W106R than to those in WT NQO1 (with somewhat higher absorbance in the mutant, as

reflected in the lower %ΔA_{450}; Fig. S12A,B). Interestingly, whereas WT enzyme kinetics showed a hyperbolic dependence on [NADH] (suggesting a transition from EX2 to EX1 mechanism as the NADH is increased sufficiently), no dependence was observed for p.W106R (Fig. 7A,B). When we analysed the mutant p.W106C, much more subtle changes were observed. The p.W106C mutant showed similar spectral properties for the reductive and oxidative half-reactions than those of the WT protein (Figs S11A,C, S12A,C, S13A, C, S14A,B, S15A,B and S16A,B). At stoichiometric concentrations of NQO1$_{holo}$:NADH and NQO1$_{red}$: DCPIP, the reductive half-reaction was slightly accelerated in p.W106C for both *fast* and *slow* pathways (up to 50%), whereas both steps were slowed down in the oxidative half-reaction (two- to sixfold) (Table 2). The NADH-dependence of k_{obs} for the reductive half-reaction in p.W106C also showed subtle changes: the k_{HT} for both *fast* and *slow* pathways was increased by 40–80%, whereas the $K_{d(NADH)}$ was increased in a similar way, leading to small or no changes in the catalytic efficiency ($k_{HT}/K_{d(NADH)}$) in both steps compared with the WT protein (Table 3 and Fig. 7C). Therefore, we must only note the effect of p.W106C on the kinetics of the oxidative half-reduction, that could be associated with the mild structural
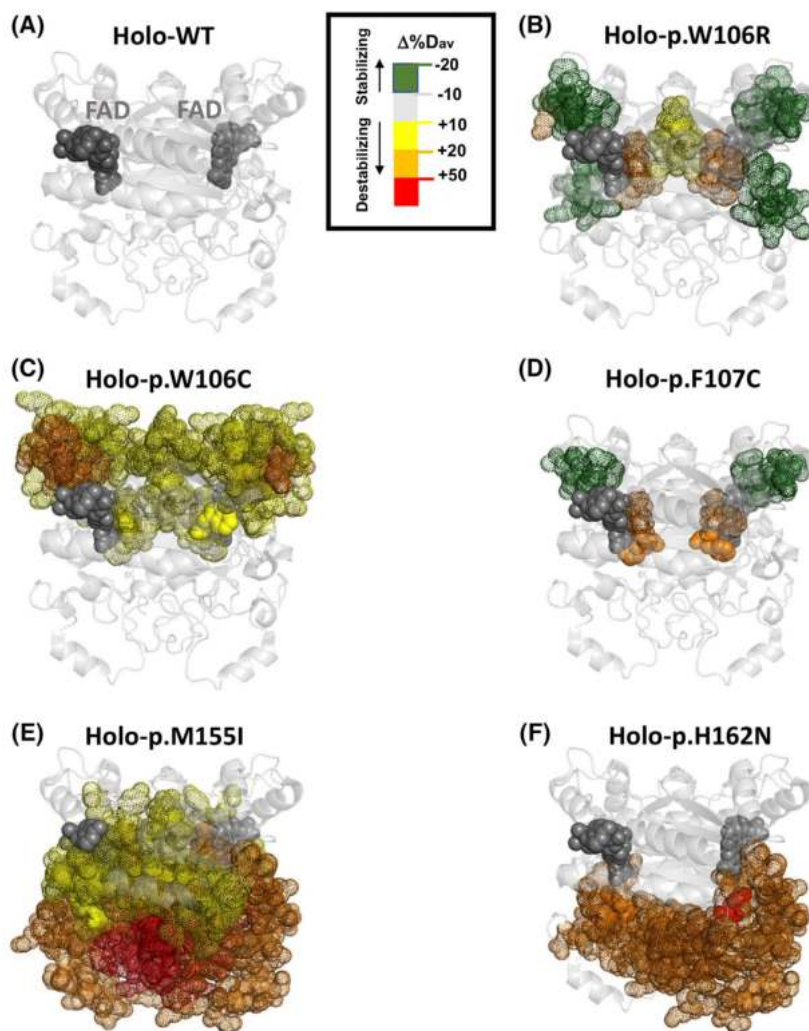
Fig. 4. Changes in local stability of NQO1$_{holo}$ variants determined by HDX-MS. Panel A shows the location of FAD (in dark grey spheres) in the WT structure. Panels B–F show the mutational effects on local stability. Colouring of B–F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [14].

destabilization of the FBS and DBS in the NQO1$_{holo}$ and NQO1$_{dic}$ states of p.W106C (Fig. 6C,D). Regarding the cooperative kinetic behaviour, negative cooperativity was slightly lower in the p.W106C mutant (~ 6) than for WT NQO1 (~ 11) as measured by an operational cooperative index (the quotient of catalytic efficiencies of the *fast/slow* pathways). This lower cooperativity might be explained by altered allosteric communication between active sites due to changes in the stability of the MMI (Fig. 6B; please see also [12,13]).

Altogether, these observations support that p.W106R cannot achieve a competent binding state for catalysis between the nicotinamide of NADH and the isoalloxazine of FAD and this limits the overall kinetics (Appendix S1). In addition, the low binding affinity of p.W106R for FAD (Fig. 3A,B) and the large destabilization of the FBS and DBS (Fig. 6C,D) may lead to the destabilization of NQO1$_{holo}$,

particularly in the reduced state, and further contribute to the slow kinetics observed for the reductive half-reaction [28,29]. A key role of altered active site electrostatics in the catalytic impairment and structural destabilization of the active site by p.W106R is supported by calculations on the active site electrostatic surface potentials (Fig. S17) (note the R106 residue, with a strong basic character and a pK_a in the range 13–14 will be always protonated). The effects of p.W106C in the active site stability and performance were much milder. These results nicely illustrate how two different non-conservative mutations at the same residue of the active site can lead to drastically different effects on protein function.

The mutant p.F107C

F107 is involved in the MMI, FBS and DBS. F107 is buried in the structure (7.6 ± 3.5% of ASA, using the
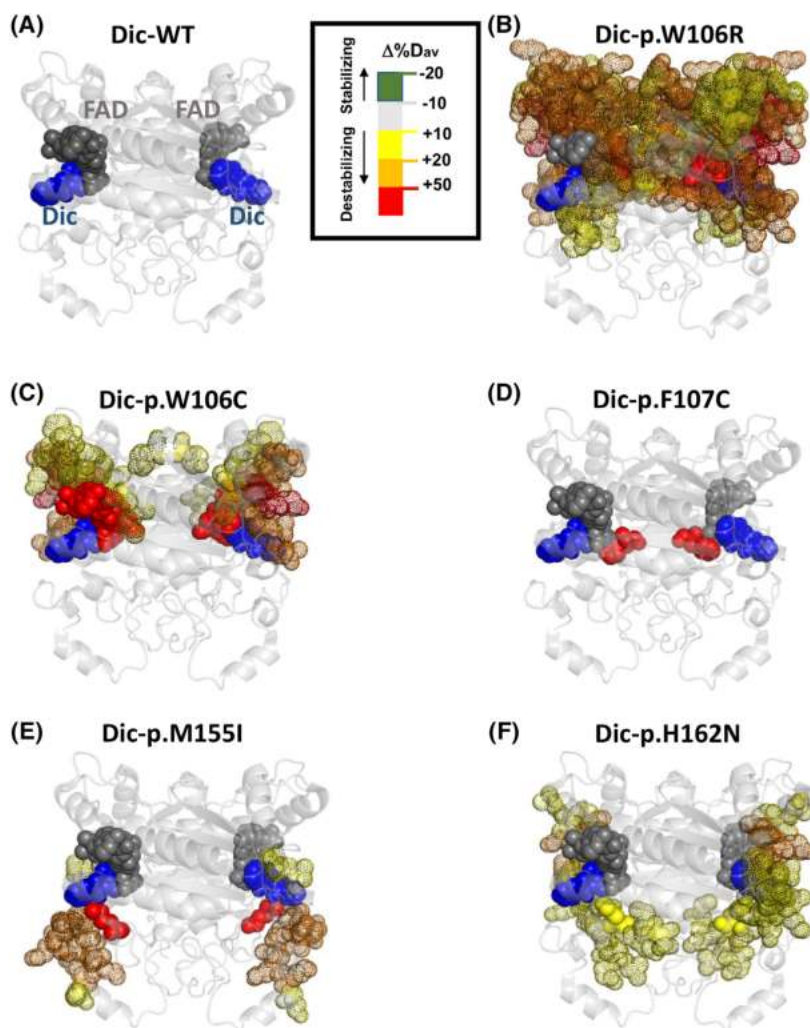
Fig. 5. Changes in local stability of NQO1$_{dic}$ variants determined by HDX-MS. Panel A shows the location of FAD (in dark grey spheres) and Dic (in blue spheres) in the WT structure. Panels B–F show the mutational effects on local stability. Colouring of B–F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [14].

same procedure described for W106). The mutation p.F107C would cause large structural perturbation creating a cavity and introducing a more polar residue [30,31]. The mutation p.F107C did not affect FAD binding affinity (Fig. 3D; please note that the affinity of WT is close to the limit of detection of the technique, and thus we cannot state that the affinity of p.F106C is higher). These results are striking since F107 is located in the FBS and is clearly a non-conservative mutation. HDX-MS analysis revealed a slight destabilization of the FBS and DBS in the NQO1$_{holo}$ state, with no effect on the NQO1$_{dic}$ state (Figs 4–6; Figs S1, S2, S5D, S6D, S7D, S8D, S9D, S10D and S17). These stability measurements are consistent with little or no effects on FAD binding and predict mild or no effects on enzyme kinetics.

However, enzyme kinetic analysis revealed puzzling results for p.F107C. Despite causing little perturbation on the stability of the active site (possibly the most

noticeable the loss of Van der Waals interactions with the Dic inhibitor), the kinetic analysis revealed some alterations due to p.F107C (Fig. 7D; Figs S11D, S12D, S13D, S14C, S15C and S16C; Tables 2 and 3). At a 1 : 1 ratio of NQO1$_{holo}$ and NADH, this mutation reduced by twofold the k_{obs} for the *fast* pathway for the reductive half-reaction, with no effect on the constant for the *slow* pathway. Nonetheless, in contrast to the WT protein, p.F107C FAD reduction is dominated by the *fast* pathway, yielding essentially identical spectra for the B and C species, and suggesting particular large alterations in the reaction equilibrium of the *slow* reaction pathway (Fig. S12D). Additionally, under similar conditions, the oxidative half-reaction was hardly affected (Table 2; Figs S14D, S15D and S16D). Analysis of the NADH-dependence of the FAD reduction kinetics provided further support to these notions as well as some explanations for these effects. For the *fast* pathway, we observed a
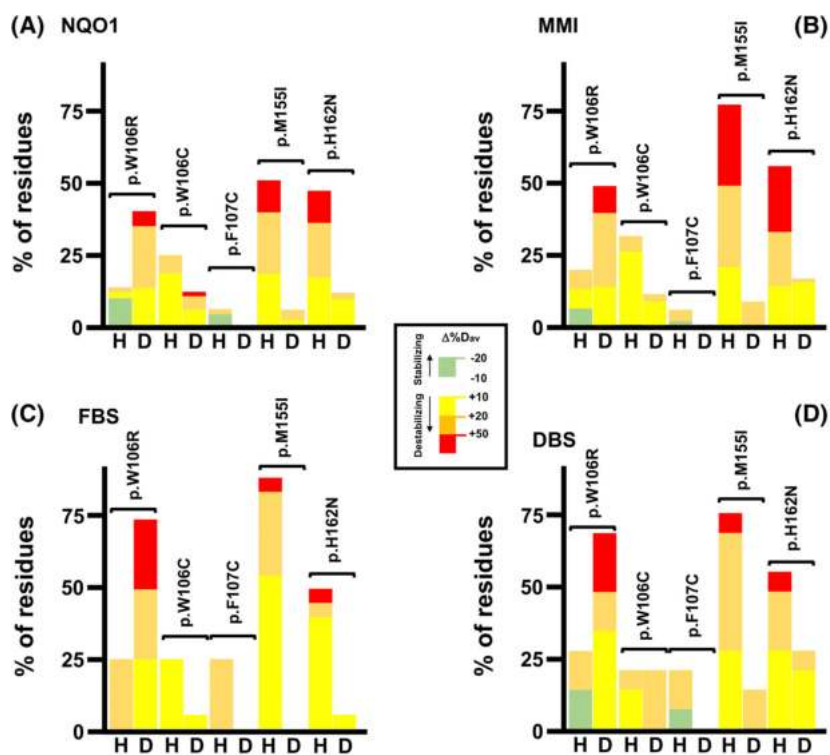
Fig. 6. Mutational effects on the local stability of different functional sites determined by HDX-MS. Effects on the entire NQO1 structure (A), the MMI (B), the FBS (C) and the DBS (D). Amino acids belonging to different functional sites were determined as described [20]. Data are expressed as the % of the residues belonging to each category (A, entire protein, 274 residues; B, MMI, 76 residues; C, FBS, 21 residues; D, DBS, 15 residues; according to Ref. [20]) affected by a given mutation and ligation state (NQO1$_{holo}$ and NQO1$_{dic}$ states are referred to as H and D in the x-axis). The colour code shows the sign and magnitude of mutational effects (as $\Delta\%$ D_{av}) in a given ligation state (H or D).

Table 2. Summary of observed rate constants (k_{obs}) for the reductive and oxidative half-reactions involving NQO1. Measurements were carried out in 20 mM HEPES-KOH, pH 7.4 at 6 °C. Evolution of the reaction was followed in the 400–1000 nm wavelength range using stopped-flow equipment with a photodiode array detector ($n > 3$, mean ± SD). 7.5 μM of NQO1$_{holo}$ protein was mixed with 7.5 μM NADH (reductive half-reaction) or 7.5 μM of NQO1$_{red}$ was mixed with 7.5 μM DCPIP (oxidative half-reaction). Primary data and fittings are shown in Figs S11, S13, S14 and S16.

	Reductive half-reaction		Oxidative half-reaction	
	k_{obs} (s^{-1})		k_{obs} (s^{-1})	
NQO1 variant	Fast	Slow	Fast	Slow
WT	78 ± 1	7.2 ± 0.3	> 500	160 ± 14
p.W106R	1.3 ± 0.2 × 10^{-2}	3.3 ± 0.7 × 10^{-3}	N.d.[a]	N.d.[a]
p.W106C	83 ± 3	12.0 ± 1.4	286 ± 34	24 ± 8
p.F107C	41 ± 2	8.8 ± 0.1	> 500	134 ± 20
p.M155I	18.2 ± 3.7	8.3 ± 4.0 × 10^{-3}	N.d.[a]	N.d.[a]
p.H162N	95 ± 3	11.9 ± 1.5	≫ 500	25 ± 2

[a]Full reduction of NQO1$_{holo}$ cannot be achieved in these variants to evaluate its reoxidation (N.d., not det.).

3.2-fold decrease in k_{HT} and a 1.8-fold increase in affinity for FAD, resulting in a 45% decrease in catalytic efficiency (Table 3). Interestingly, the *slow* pathway for FAD reduction was performed efficiently but, as observed for the same pathway in p.W106R, it was NADH-independent thus suggesting that the mutation is changing the rate-limiting step (analogous to the EX2 to EX1 shift in mechanism, see Appendix S1). The cooperative index for this mutant was ~ 10, similar to that of WT, consistent with an efficient communication between active sites during the reaction (note that the MMI is hardly affected by the mutation p.F107C; Fig. 6B). However, conclusions related to these effects on the negative cooperativity are not clear since the reduction kinetics is strongly dominated by the amplitude and kinetics of the *fast* step. Although these catalytic alterations are not as dramatic as those of p.W106R, our results indicate that in some cases,

Table 3. Enzyme kinetic parameters for the reductive half-reaction of NQO1$_{holo}$ variants with NADH. Primary data and fittings are shown in Fig. 7.

	Fast			Slow		
	k_{HT} (s^{-1})	$K_{d(NADH)}$ (μM)	$k_{HT}/K_{d(NADH)}$ (s^{-1}·μM^{-1})	k_{HT} (s^{-1})	$K_{d(NADH)}$	$k_{HT}/K_{d(NADH)}$ (s^{-1}·μM^{-1})
WT	281 ± 12	15.2 ± 1.9	18.5 ± 2.4	14.3 ± 1.5	8.2 ± 3.3	1.7 ± 0.7
p.W106R	$12.5 \pm 0.1 \times 10^{-3a}$	N.a.^{a}	N.a.^{a}	$2.6 \pm 0.4 \times 10^{-3a}$	N.a.^{a}	N.a.^{a}
p.W106C	397 ± 17	29 ± 3	13.9 ± 1.5	26 ± 1	10.9 ± 1.5	2.4 ± 0.4
p.F107C	87 ± 2	8.3 ± 0.7	10.5 ± 0.9	8.8 ± 0.1^a	N.a.^{a}	N.a.^{a}
p.M155I	337 ± 61	64 ± 21	5.2 ± 2.0	0.104 ± 0.028	43 ± 24	$2.4 \pm 1.5 \times 10^{-3}$
p.H162N	415 ± 31	27 ± 5	15.3 ± 6.4	N.a.^{b}	N.a.^{b}	2.1 ± 0.9^b

^{a}k_{obs} is independent of NADH concentration, therefore data do not allow us to determine $K_{d(NADH)}$ (N.a., not applicable) and the given value corresponds to a limiting k_{HT} calculated as the average of k_{obs} at different [NADH]; ^{b}k_{obs} is linearly dependent on [NADH]; therefore, data do not allow to determine neither $K_{d(NADH)}$ nor k_{HT} (N.a., not applicable) and the given value corresponds to the second-order rate constant of the process calculated as the slope of the k_{obs} vs. [NADH] plot.

even mild perturbations of the structural stability of the active site may cause noticeable alterations in enzyme catalysis and kinetic mechanism. An alternative hypothesis is that this mutation is affecting dynamics relevant to the reaction mechanism. This is not well probed by HDX-MS (note that dynamics cannot be properly addressed since the EX1 regime is marginally detected by this technique; Fig. S3) and possibly the relevant dynamics are faster (in the μs to ms scale) than those sampled by HDX-MS.

The mutant p.M155I

M155 belongs to the MMI and the DBS and is next to H156 which is part of the FBS and the DBS. M155 is buried in the structure ($12.9 \pm 3.9\%$ of ASA, using the same procedure described for p.W106R). This mutation causes an increase in hydrophobicity [30] without a change in size [31].

Despite the mutation p.M155I does not appear to be highly disruptive, its effects on the local stability of the holo-state are strong, affecting about 50% of residues in NQO1, the MMI, the FBS and the DBS. It does destabilize both the NTD and CTD. Upon Dic binding, these effects are significantly reduced (Figs 4–6; Figs S1, S2, S5E, S6E, S7E, S8E, S9E, S10E and S17). It is interesting to note that this mutation reduced the FAD binding affinity by ~ 50-fold (Fig. 3E), consistent with a strong destabilization of the FBS in the NQO1$_{holo}$ state (Fig. 6C). This is one of the lowest affinities for FAD reported for a single missense variant of NQO1 ([12,19] and this work), reducing the apparent binding free energy by 2.3 kcal·mol^{-1}. The strong destabilization on the MMI, FBS and DBS in the NQO1$_{holo}$ state may also have implications in the catalytic performance of this mutant.

To further characterize the functional consequences of p.M155I, we carried out kinetic analysis for the reductive and oxidative half-reactions at stoichiometric protein to reductive/oxidative ligand (NADH/DCPIP) (Table 2; Figs S11E, S12E and S13E). For the reductive half-reaction, the two-step mechanism was observed but with large effects on k_{obs}, which were reduced by 4-fold and 900-fold in the *fast* and *slow* steps. In addition, the spectral properties of the intermediate species B and C showed markedly higher absorption intensities than those of these species in WT NQO1 (Fig. S12A,E), supporting severe alterations of equilibria in the overall reaction mechanism. Oxidation of FAD by DCPIP could not be analysed due to the inability to achieve the full NQO1$_{red}$ state upon reduction. The NADH-concentration dependence of k_{obs} for the reductive half-reaction (Fig. 7E and Table 3) confirmed substantial alterations in the kinetic properties. For the *fast* step, p.M155I reduced the affinity for NADH by fourfold, resulting in a threefold decrease in catalytic efficiency (Table 3). These effects were dramatic for the *slow* step that showed 140-fold lower k_{HT}, fivefold higher $K_{d(NADH)}$ and resulted in a 700-fold decrease in catalytic efficiency compared with WT NQO1 (Table 3). Accordingly, the cooperative index for p.M155I was about 600-fold higher than that of WT NQO1. This large negative cooperativity can be regarded as an extreme case of half-of-sites reactivity in p.M155I (i.e. only one active site is operative per dimer in a relevant time-scale, up to 1 min). Therefore, the strong effects on FAD and NADH binding, catalytic efficiency and negative cooperativity seem to primarily stem from large destabilization of the active site (FBS and DBS) and the MMI in the NQO1$_{holo}$ state (Fig. 6C–B).
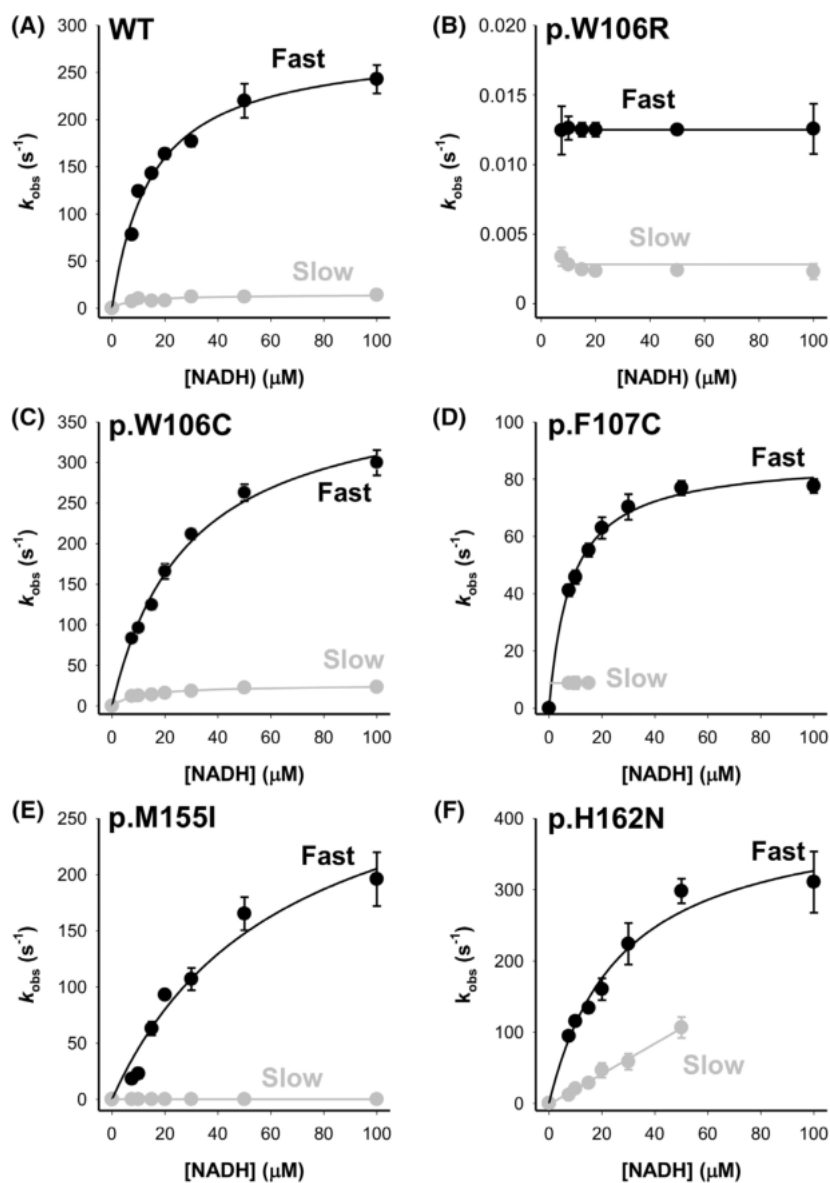
Fig. 7. Pre-steady state enzyme kinetic analyses for the reductive half-reaction of NQO1 WT and active-site mutants. Panels A–F show the dependence of the observed rate constants on NADH concentration for the different NQO1 variants. Data for the *fast* and *slow* steps are shown in black and grey, respectively. Lines are best fits to a hyperbolic function or to a straight line.

The mutant p.H162N

H162 belongs to the MMI, FBS and DBS and it is buried in the structure (6.6 ± 0.4% of ASA, using the same procedure described for p.W106R). The mutation p.H162N causes little change in hydrophobicity [30] or size [32]. Therefore, we may initially consider this mutation as quite conservative.

The mutation p.H162N strongly reduces the FAD binding affinity (by 30-fold vs. WT NQO1; Fig. 3F) equivalent to an apparent binding free energy penalty of ~ 2.0 kcal·mol^{-1}. The effects of the mutation p.H162N on protein stability are quite similar to those observed for p.M155I (Figs 4–6; Figs S1, S2,

S5F, S6F, S7F, S8F, S9F, S10F and S18). The NQO1$_{holo}$ state was largely destabilized globally and at the MMI, FBS and DBS (about 50% of the residues are destabilized; Fig. 6), and these effects were much weaker in the NQO1$_{dic}$ state (particularly in the FBS). As proposed for the mutation p.M155I, the low stability of the FBS in the NQO1$_{holo}$ state may cause the low FAD binding affinity, and reduced stability of the MMI, FBS and DBS in the holo-state may affect catalytic performance and functional cooperativity.

Rate constants for the reductive half-reaction (at stoichiometric NQO1$_{holo}$:NADH) showed little or no effects due to p.H162N, while in the oxidative half-

reduction, the *fast* step was accelerated (beyond the temporal resolution of the technique) and the *slow* step was slowed down by sixfold (Table 2). The spectral changes associated with the two steps in the reductive (Fig. S12A,F) and oxidative (Fig. S15A,D) processes supported that the mutation p.H162N does not affect the overall reaction mechanism. When evaluating the NADH dependence of k_{obs}, k_{HT} and $K_{d(NADH)}$ for the *fast* step were slightly affected (both increased by 50–80%) leading to nearly no change in catalytic efficiency (Fig. 7F and Table 3). Interestingly, for the *slow* step, the k_{obs} values despite being larger than at equivalent NADH concentrations in the WT case, linearly increased on substrate concentration with no evidence of saturation (i.e. similar to an EX2 scenario; Appendix S1). This may result from different effects of this mutation in the *slow* pathway for FAD reduction: a large increase in the $K_{d(NADH)}$ and/or k_{off}, or a large decrease in the k_{on}.

Genotype–phenotype correlations using multifeatured experimental characterization of active-site mutants and bioinformatic predictions

The results obtained from our detailed experimental characterization of the NQO1 active-site mutants are summarized in Table 4. There, mutational effects on 13 features regarding protein functionality and stability are semiquantitatively ranked. Overall, we may conclude that the mutations p.W106R, p.M155I and p.H162N are largely detrimental to NQO1 function and stability, whereas the effects of p.W106C and

p.F107C are milder. Interestingly, mutational effects predicted by different widely used bioinformatic tools did not agree well with the experimental outcome, predicting (within some degree of disagreement between tools) than all mutations except p.H162N should be largely detrimental. Therefore, experimental and computational prediction of mutational seemed to correlate poorly.

To analyse more quantitatively this apparently poor correlation, we determined scores for the experimental results (13 features compiled in Table 4) and those provided by computational predictions (using six different tools, Table 1). All 19 scores were normalized in such a way that a score of 0 indicated WT-like behaviour and a score of 1 indicated a largely deleterious effect. Thus, we evaluated average scores for experimental analyses (NS_{exp}) and predictions (NS_{pred}) (Tables 1 and 4). These scores were aimed to capture either the average effects (with the same weight for each of the features; note that expression of all variants was quite successful and thus, none seemed to largely affect NQO1 foldability) observed experimentally on NQO1 functional and stability and the average performance of the predictions. When we simply plotted the two scores (Fig. 8), we observed a very weak correlation between the two scores, as expected. This simple exercise also confirmed that the functional effects of the mutations p.W106C and p.F107C (two largely non-conservative mutations) are much milder than those predicted by *in silico* tools, while the opposite behaviour was found for p.H162N (in principle a more conservative mutation). Overall, these results are

Table 4. Summary of the effects of active site mutations. For each protein variant and feature, these were clustered for semiquantitative comparison in four categories: (++++) mildly improved vs. WT; (+++) WT-like; (++) mildly moderately impaired vs. WT; (+) largely impaired vs. WT. NS_{exp} is the average of experimental normalized scores, considering ++++ as −0.5, +++, as 0, ++ as 0.5 and +as 1.

	NQO1 variant					
	WT	p.W106R	p.W106C	p.F107C	p.M155I	p.H162N
Protein conformation	+++	+++	+++	+++	++	++
FAD bound	+++	++	++	+++	++	++
FAD affinity	+++	+	+++	++++	+	+
Global stability by HDX-Holo	+++	+++	++	+++	+	+
Global stability by HDX-Dic	+++	+	++	+++	+++	++
FBS stability by HDX	+++	+	++	++	+	+
DBS stability by HDX	+++	+	++	+++	+	+
MMI stability by HDX	+++	+	++	+++	+	+
Fast FAD reduction	+++	+	+++	++	+++	+++
Slow FAD reduction	+++	+	+++	++	++	++
FAD oxidation	+++	+	++	+++	+	+++
Reaction mechanism	+++	++	+++	++	+	+++
Cooperativity	+++	+	+++	++	+++	+++
NS_{exp}	0	0.77	0.27	0.15	0.65	0.54
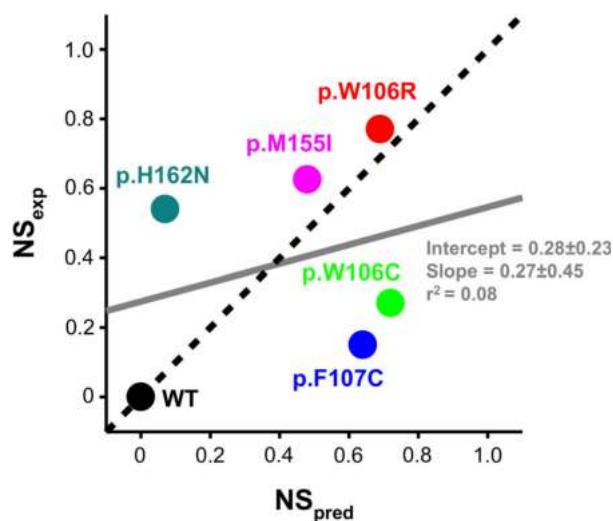
Fig. 8. Lack of correlation between mutational effects from experiments (NS_{exp}) and predictive tools (NS_{pred}) for the active-site mutants. Predicted scores (NS_{pred}) are the average from normalized scores using six different predictive tools (Table 1). Experimental scores (NS_{exp}) are the average of normalized scores derived from the characterization of 13 functional and stability features (Table 4). The dashed black line represents a perfect linear correlation between scores, whereas the solid grey line shows the actual (poor) linear correlation.

somewhat counterintuitive, since active site residues evolve slowly and are strongly constrained in terms of physico-chemical properties of the residue such as polarity, charge and hydrophobicity [24,33].

Conclusions

State-of-the-art DNA sequencing technologies allow to compare the genomes of hundreds of thousands of individuals and have revealed an astonishingly large genetic diversity in the human genome. In this context, classical tools to establish genotype–phenotype correlations and predict pathogenicity of missense variants *in silico* are still not sufficiently accurate, particularly when individual mutations are analysed (actually, the output of many of these tools is binary, for example, pathogenic or damaging vs. benign or neutral). In this work, we carried out comprehensive structure–function studies of five naturally occurring missense mutations affecting residues belonging to the active site of the cancer-associated NQO1 enzyme and predicted to be pathogenic in most of the cases (Table 1). Experimental characterization of mutational effects shows widely different and counterintuitive consequences (Table 4). For instance, two non-conservative mutations affecting the same residue (p.W106R and p.W106C) have

largely different consequences (the former is catastrophic while the latter is much milder). In addition, quite conservative mutations (p.M155I and H162N) display much larger effects than some non-conservative changes (p.W106C). Remarkably, mutations with large functional consequences are shown to target the stability of different functional states (p.W106R, the $NQO1_{dic}$ state vs. p.M155I and p.H162N, the $NQO1_{holo}$ state). Although in most cases, structure–function correlations can be nicely drawn, some mutations appear to affect enzyme function with little structural impact (p.F107C). Our work highlights the necessity to incorporate additional information to predictive tools, such as structural effects on different ligation states (i.e. the flavin, substrate) and several protein functional features as well as some notion of structural plasticity (as the ability of the protein conformational landscape to adapt to different types of mutations in different or similar structural locations).

Materials and methods

Protein expression and purification

Mutations were introduced by site-directed mutagenesis in the wild-type (WT) NQO1 cDNA cloned into the pET-15b vector (pET-15b-NQO1) by GenScript (Leiden, the Netherlands). Codons were optimized for expression in *Escherichia coli* and mutagenesis was confirmed by sequencing the entire cDNA. The plasmids were transformed in *E. coli* BL21(DE3) cells (Agilent Technologies, Santa Clara, CA, USA) for protein expression. These constructs contain a hexa-his N-terminal tag for purification.

For protein purifications, a preculture (40 mL) was prepared from a single clone for each variant and grown for 16 h at 37 °C in LBA (Luria-Bertani medium with 0.1 mg·mL^{-1} ampicillin) and diluted into 2.4–4.8 L of LBA. After 3 h with shaking (200 r.p.m.) at 37 °C, NQO1 expression was induced by the addition of 0.5 mM IPTG for 6 h at 25 °C. Cells were harvested by centrifugation at 8000 *g* and frozen overnight at −80 °C. NQO1 proteins were purified using immobilized nickel affinity chromatography (IMAC) columns (Cytiva, Marlborough, MA, USA) as described [13]. Isolated dimeric fractions of NQO1 variants were exchanged to HEPES-KOH buffer 50 mM pH 7.4 using PD-10 columns (Cytiva). The UV–visible spectra of purified NQO1 proteins were registered in a Cary 50 spectrophotometer (Agilent Technologies, Waldbronn, Germany) and used to quantify the protein content as described in [11]. For the samples used in pre-steady state kinetic analyses, NQO1 proteins were incubated with 1 mM FAD and the excess of FAD was removed using PD-10 columns, obtaining a saturation fraction (FAD:NQO1

monomer) higher than 90% based on UV–visible spectra. Apo-proteins were obtained by treatment with 2 M urea and 2 M KBr as described [20], obtaining samples with < 2% saturation fraction of FAD based on UV–visible spectra. Samples were stored at −80 °C upon flash freezing in liquid N_2. Protein purity and integrity were checked by polyacrylamide gel electrophoresis in the presence of sodium dodecylsulphate (SDS/PAGE). Each NQO1 variant was expressed at least three times.

Protein conformation

After centrifugation for 20 min at 21 000 *g* at 4 °C, the FAD bound to NQO1 proteins as purified was determined from the UV–visible spectrum using $\varepsilon_{(280)} = 47\,900$ $M^{-1} \cdot cm^{-1}$ for NQO1, $\varepsilon_{(450)} = 22\,000$ $M^{-1} \cdot cm^{-1}$ and $\varepsilon_{(280)} = 11\,300$ $M^{-1} \cdot cm^{-1}$ for FAD [11]. Spectra were collected in a Cary 50 spectrophotometer using 0.3 cm path-length quartz cuvettes and ~ 30 μM of NQO1 in the protein subunit. Data were reported as mean ± SD from three different purifications.

Far-UV CD spectroscopy was performed at 25 °C in K-phosphate 20 mM at pH 7.4 using 5 μM protein (in NQO1 monomer) with 25 μM FAD. Spectra were collected at 25 °C in a Jasco J-710 spectropolarimeter (Tokyo, Japan) in the 195–260 nm range, at 100 nm·min^{-1}, using 1 nm bandwidth, 1 s response time and 1 mm path length cuvette. Each spectrum was the average of six scans, and each sample was prepared in triplicate. Mean residue ellipticities ([Θ]$_{MRW}$) were calculated using Eqn (1):

$$[\Theta]_{MRW} = \frac{MRW \cdot \Theta_{obs}}{10 \cdot l \cdot c}, \tag{1}$$

where MRW was equal to the molecular weight of the NQO1 monomer (31 691 g·mol^{-1}) divided by $(N-1)$, being $N = 280$ the number of residues in the monomer, Θ_{obs} was the ellipticity (in mdeg), l was the path length (in cm) and c the concentration of protein in mg·mL^{-1}. Spectra were reported as mean ± SD from three replicates.

Near-UV CD spectroscopy was performed at 25 °C in K-phosphate 20 mM at pH 7.4 using 20 μM protein as purified (in NQO1 monomer) with 50 μM FAD as described above for far-UV CD measurements. Measurements were carried out in the 250–600 nm range, at 100 nm·min^{-1}, using 1 nm bandwidth, 1 s response time and a 5-mm path length cuvette. Each spectrum was the average of 10 scans, and the appropriate blank in the absence of protein was acquired and subtracted.

Fluorescence spectra were acquired in a Varian Cary Eclipse spectrofluorometer (Agilent Technologies) using 1 cm path length cuvettes and 1 μM (in monomer) of protein in the presence of 5 μM FAD in 20 mM K-phosphate pH 7.4. The excitation wavelength was 280 nm and the emission fluorescence was collected between 300 and 450 nm. Excitation and emission slits were 5 nm. All spectra were acquired at 25 °C at a 120 nm·min^{-1} scan rate and 10 scans were registered and averaged. Blanks without protein were routinely measured and subtracted. The spectral center of mass (SCM) was determined using Eqn (2):

$$SCM = \frac{\sum\limits_{330\,nm}^{380\,nm} I_\lambda \cdot \lambda}{\sum\limits_{330\,nm}^{380\,nm} I_\lambda}. \tag{2}$$

where I_λ is the emission intensity at a given emission wavelength λ. Data were reported as mean ± SD from three replicates.

Dynamic light scattering was carried out in a Zetasizer μV instrument (Malvern Panalytical, Malvern, UK) using 1.5 mm path length cuvettes and 5 μM (in monomer) of protein with 25 μM FAD in 20 mM K-phosphate pH 7.4 at 25 °C. Thirty measurements with an acquisition time of 10 s were acquired for each DLS analysis, averaged and used to determine the hydrodynamic radius assuming spherical scattering particles (using the Stokes-Einstein approach). Data were reported as mean ± SD from three replicates. DLS data were processed and analysed using the ZETASIZER software (Malvern Panalytical).

Flavin-adenine dinucleotide binding affinity

Fluorescence titrations were carried out at 25 °C using 1×0.3 cm path-length cuvettes in a Varian Cary Eclipse spectrofluorometer (Agilent Technologies). Experiments were carried out in 20 mM K-phosphate, pH 7.4, essentially as described in [15]. Briefly, 20 μL of a 12.5 μM NQO1$_{apo}$ stock solution (in subunit) was mixed with 0–500 μL of FAD 10 μM, and the corresponding volume of buffer was added to yield a 1 mL of final volume. Samples were incubated at 25 °C in the dark for at least 10 min before measurements. Spectra were acquired in the 340–360 nm range upon excitation at 280 nm (slits 5 nm), and spectra were averaged over 10 scans registered at a scan rate of 200 nm·min^{-1}.

For W106R, which exhibited a very low binding affinity for FAD, the titration was carried out using near-UV CD spectroscopy. Spectra were collected in a Jasco J-710 spectropolarimeter at 25 °C in 50 mM K-HEPES pH 7.4 using 9.4 μM (in monomer) of apoprotein in the absence or presence of FAD (0–130 μM). Spectra were collected in the 300–600 nm range at 200 nm·min^{-1}, using 2 nm bandwidth, 2 s time response and 5 mm path-length cuvettes. Each spectrum was the average of 8 scans, and each sample was appropriately corrected for blanks containing the buffer and the corresponding FAD concentration.

Flavin-adenine dinucleotide binding titrations (following fluorescence intensities at 350 nm or CD ellipticity at

375 nm) were fitted using a single and identical type of binding sites as described in [15] using the GRAPHPAD PRISM 7 sofware (DotMatics, Boston, MA, USA).

Hydrogen/deuterium exchange mass spectrometry

Amide hydrogen/deuterium exchange (HDX) of NQO1 was studied for the WT and mutant variants in the NQO1$_{holo}$ and NQO1$_{dic}$ states as described previously [20] with some modifications. Briefly, to start the exchange reaction 20 μM protein solution was diluted 10 times with a D_2O-based 20 mM HEPES-NaOH, 100 mM NaCl, 0.5 mM TCEP [tris (2-carboxyethyl)phosphine], pD 7.4. The exchange was terminated after 10, 50, 250, 1250 and 6250 s by mixing (1 : 1) with 0.5 M Glycine-HCl, pH 2.3 and the samples were flash frozen in liquid nitrogen. Time points 10, 250 and 6250 s were replicated. Custom-made co-immobilized nepenthesin-2/pepsin column was used for online proteolysis which was driven by 0.4% formic acid (FA) in water pumped at 200 μL·min^{-1}. The solvent was delivered by 1260 Infinity II Quaternary pump (Agilent Technologies). Peptides were trapped and desalted on a SecurityGuard™ pre-column (ULTRA Cartridges UHPLC Fully Porous Polar C18, 2.1 mm; Phenomenex, Torrance, CA, USA) using the same solvent. The duration of digestion and desalting was 3 min. Following desalting, the peptides were separated on an analytical column (LUNA® Omega Polar C18 Column, 100 Å, 1.6 μm, 100 mm × 1.0 mm; Phenomenex) using 1290 Infinity II LC System (Agilent Technologies). Linear gradient 5–45% B in 6 min was followed by a quick step to 99% B lasting 5 min. Flow during the separation was 40 μL·min^{-1}. Solvent A was 0.1% FA/2% acetonitrile (ACN) in water, B was 0.1% FA/98% ACN in water. Digestion, desalting, and separation steps were performed at 0 °C and pH 2.3 to minimize deuterium back-exchange. LC system was directly coupled to an ESI source of 15T FT-ICR mass spectrometer (solariX XR; Bruker Daltonics, Bremen, Germany) operating in a broad-band MS mode. Correction of back-exchange was done as described in [34] using fully deuterated control prepared for each NQO1 variant [20]. The data were further processed in DATAANALYSIS 5.3, exported and assembled into a project under in-house developed program DEUTEX [35]. Peptides were identified using separate data-dependent LC–MS/MS runs carried out using an identical LC setup connected to ESI-timsTOF Pro with PASEF. MASCOT (v 2.4; Matrix Science, London, UK) was used for data searching against a custom-built database containing sequences of the proteases, NQO1 variants and common contaminants. Decoy search was enabled with a false-discovery ratio < 1% and an ion score cut-off of 20. All data were deposited to ProteomeXchange Consortium via the PRIDE database [PXD036417] [36].

To evaluate the effect of mutations, the difference in kinetics of deuterium incorporation (% D vs. time, Figs S1 and S2) of mutants and the WT protein was calculated for a given ligation state and each protein segment experimentally determined. Analysis of exchange behaviour showed that the EX2 mechanism dominates HDX in most of the peptides, variants and ligation states (Fig. S3), thus supporting that changes in HDX are associated with those in the local thermodynamic stability of the segments [20]. The average of the two most different time points (mutant-WT) was used to determine the $\Delta\%D_{av}$ values. This parameter allows to readily compare the HDX kinetics between two given NQO1 states/variants in different protein segments [12,19,20].

Enzyme kinetics for the reductive and oxidative half-reactions

For enzyme kinetic analyses of the reductive and oxidative half-reactions, we followed the procedures described for the WT protein under anaerobic conditions using a stopped-flow spectrophotometer as described [13]. Briefly, the reductive half-reaction was measured by mixing the NQO1$_{holo}$ protein with a solution of NADH, yielding final concentrations of 7.5 and 7.5–100 μM, respectively. The oxidative half-reaction was monitored after mixing NQO1$_{red}$ samples (NQO1$_{red}$ was obtained by previous mixing of NADH to the holo-NQO1, both at 7.5 μM) with an equimolar concentration of 2,6-Dichlorophenolindophenol (DCPIP). Reactions were performed in 20 mM HEPES-KOH, pH 7.4.

Multiple wavelength absorption data in the flavin absorption region were collected and processed as described [13]. Time-dependent spectral deconvolution was performed by global fitting analysis and numerical integration using previously described procedures [13] and allowed to determine observed rate constants (k_{obs}) for these steps as well as spectroscopic properties of these species (A, B and C). Despite practical limitations prevented these measurements to reach pseudo-first-order conditions [19], hyperbolic dependences of k_{obs} vs. NADH concentrations were fitted using Eqn (3):

$$k_{obs} = \frac{k_{HT} \cdot [\text{NADH}]}{K_{d(\text{NADH})} + [\text{NADH}]}, \qquad (3)$$

where k_{HT} is the limiting rate constant for HT and $K_{d(\text{NADH})}$ is the apparent equilibrium dissociation constant for NADH to a given active site. Fittings were carried out using SIGMAPLOT v.9.0 (SYSTAT Software Inc., Chicago, IL, USA).

Bioinformatic analysis

Six different readily available algorithms for the prediction of mutational effects were used. These are based on various

features such as evolutionary conservation, chemical nature of the amino acid change and structural consequences. These tools are briefly described in this section.

We used the Ensembl Variant Effect Predictor (VEP) (https://www.ensembl.org/Tools/VEP) [37] to obtain predictions for single nucleotide variants. Most of these predictions can be assessed at the database dbNSFP [38].

POLYPHEN-2 (http://genetics.bwh.harvard.edu/pph2/) predicts the impact of amino acid substitutions on the structure and function of a human protein using physical and comparative considerations [39]. Its score (0–1) yields the probability of the variation being damaging and contemplates three classes: benign, possibly damaging and probably damaging.

SIFT (https://sift.bii.a-star.edu.sg/) is based on sequence homology and the physical properties of amino acids [40]. It aligns protein sequences in numerous species and calculates normalized probabilities for all possible substitutions from the alignment. Its score also ranges 0–1. The amino acid substitution is predicted as damaging if the score is ≤ 0.05, and as tolerated if the score is > 0.05.

CADD (https://cadd.gs.washington.edu/) is a meta-predictor that takes into account many diverse annotations into a single score (C score) [41]. The higher the raw C score (between 1 and 99), the more likely is the change to be deleterious.

MUTATION ASSESOR (http://mutationassessor.org) estimates the functional impact of a missense variant based on evolutionary conservation of the affected amino acid in protein homologues [42]. A conservation score is combined with a specificity score to determine a functional impact score (0–1). Variants classed as neutral or low are predicted to have low or no impact on protein function, whereas variants classed as medium or high are predicted to result in altered function.

REVEL (integrated in VEP) [43] is a meta-predictor based on a number of individual tools: MUTPRED, FATHMM, VEST, POLYPHEN, SIFT, PROVEAN, MUTATIONASSESSOR, MUTATION-TASTER, LRT, GERP, SIPHY, PHYLOP, and PHASTCONS. The score (between 0 and 1) classifies the variations between likely benign (score < 0.5) and likely disease-causing (score ≥ 0.5).

METAL R (integrated into VEP) ([38]) uses logistic regression to integrate nine independent scores and allele frequency information to yield a score (between 0 and 1), with lower scores considered as tolerated and higher scores more likely to be damaging.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

Author contributions

JLP-G performed experiments and analysed data. EA-C performed experiments and analysed data. DSL performed experiments and analysed data. DK performed experiments and analysed data. ES analysed data and contributed reagents. PM planned experiments and analysed data. MM planned experiments and analysed data. ALP planned experiments, analysed data and wrote the paper.

Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1111/febs.16677.

Data availability statement

All HDXMS data were deposited to ProteomeXchange Consortium via the PRIDE database [PXD036417]. No new high-resolution structures were generated in this manuscript. All data in a processed form are contained in the manuscript. Any other raw data can be retrieved upon reasonable request to the corresponding author (ALP).

References

1 Arnedo-Pac C, Lopez-Bigas N, Muiños F. Predicting disease variants using biodiversity and machine learning. *Nat Biotechnol*. 2022;**40**:27–8.

2 Høie MH, Cagiada M, Beck Frederiksen AH, Stein A, Lindorff-Larsen K. Predicting and interpreting

large-scale mutagenesis data using analyses of protein stability and conservation. *Cell Rep*. 2022;**38**:110207.

3 Shendure J, Akey JM. The origins, determinants, and consequences of human mutations. *Science*. 2015;**349**:1478–83.

4 McInnes G, Sharo AG, Koleske ML, Brown JEH, Norstad M, Adhikari AN, et al. Opportunities and challenges for the computational interpretation of rare variation in clinically important genes. *Am J Hum Genet*. 2021;**108**:535–48.

5 Beaver SK, Mesa-Torres N, Pey AL, Timson DJ. NQO1: a target for the treatment of cancer and neurological diseases, and a model to understand loss of function disease mechanisms. *Biochim Biophys Acta Proteins Proteom*. 2019;**1867**:663–76.

6 Salido E, Timson DJ, Betancor-Fernández I, Palomino-Morales R, Anoz-Carbonell E, Pacheco-García JL, et al. Targeting HIF-1α function in cancer through the chaperone action of NQO1: implications of genetic diversity of NQO1. *J Pers Med*. 2022;**12**:747.

7 Ross D, Siegel D. The diverse functionality of NQO1 and its roles in redox control. *Redox Biol*. 2021;**41**:101950.

8 Faig M, Bianchet MA, Talalay P, Chen S, Winski S, Ross D, et al. Structures of recombinant human and mouse NAD(P)H:quinone oxidoreductases: species comparison and structural changes with substrate binding and release. *Proc Natl Acad Sci USA*. 2000;**97**:3177–82.

9 Li R, Bianchet MA, Talalay P, Amzel LM. The three-dimensional structure of NAD(P)H:quinone reductase, a flavoprotein involved in cancer chemoprotection and chemotherapy: mechanism of the two-electron reduction. *Proc Natl Acad Sci USA*. 1995;**92**:8846–50.

10 Lienhart W-D, Gudipati V, Uhl MK, Binter A, Pulido SA, Saf R, et al. Collapse of the native structure caused by a single amino acid exchange in human NAD(P)H: quinone oxidoreductase(1.). *FEBS J*. 2014;**281**:4691–704.

11 Medina-Carmona E, Fuchs JE, Gavira JA, Mesa-Torres N, Neira JL, Salido E, et al. Enhanced vulnerability of human proteins towards disease-associated inactivation through divergent evolution. *Hum Mol Genet*. 2017;**26**:3531–44.

12 Pacheco-Garcia JL, Anoz-Carbonell E, Vankova P, Kannan A, Palomino-Morales R, Mesa-Torres N, et al. Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue. *Redox Biol*. 2021;**46**:102112.

13 Anoz-Carbonell E, Timson DJ, Pey AL, Medina M. The catalytic cycle of the antioxidant and cancer-associated human NQO1 enzyme: hydride transfer, conformational dynamics and functional cooperativity. *Antioxidants*. 2020;**9**:1–22.

14 Asher G, Dym O, Tsvetkov P, Adler J, Shaul Y. The crystal structure of NAD(P)H quinone oxidoreductase 1 in complex with its potent inhibitor dicoumarol. *Biochemistry*. 2006;**45**:6372–8.

15 Pacheco-García JL, Cano-Muñoz M, Sánchez-Ramos I, Salido E, Pey AL. Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein. *J Pers Med*. 2020;**10**:1–31.

16 Pey AL. Biophysical and functional perturbation analyses at cancer-associated P187 and K240 sites of the multifunctional NADP(H):quinone oxidoreductase 1. *Int J Biol Macromol*. 2018;**118**:1912–23.

17 Medina-Carmona E, Palomino-Morales RJ, Fuchs JE, Padín-Gonzalez E, Mesa-Torres N, Salido E, et al. Conformational dynamics is key to understanding loss-of-function of NQO1 cancer-associated polymorphisms and its correction by pharmacological ligands. *Sci Rep*. 2016;**6**:20331.

18 Medina-Carmona E, Betancor-Fernández I, Santos J, Mesa-Torres N, Grottelli S, Batlle C, et al. Insight into the specificity and severity of pathogenic mechanisms associated with missense mutations through experimental and structural perturbation analyses. *Hum Mol Genet*. 2019;**28**:1–15.

19 Pacheco-Garcia JL, Loginov DS, Anoz-Carbonell E, Vankova P, Palomino-Morales R, Salido E, et al. Allosteric communication in the multifunctional and redox NQO1 protein studied by cavity-making mutations. *Antioxidants*. 2022;**11**:1110.

20 Vankova P, Salido E, Timson DJ, Man P, Pey AL. A dynamic Core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer. *Biomolecules*. 2019;**9**:728.

21 Medina-Carmona E, Rizzuti B, Martín-Escolano R, Pacheco-García JL, Mesa-Torres N, Neira JL, et al. Phosphorylation compromises FAD binding and intracellular stability of wild-type and cancer-associated NQO1: insights into flavo-proteome stability. *Int J Biol Macromol*. 2019;**125**:1275–88.

22 Muñoz IG, Morel B, Medina-Carmona E, Pey AL. A mechanism for cancer-associated inactivation of NQO1 due to P187S and its reactivation by the consensus mutation H80R. *FEBS Lett*. 2017;**591**:2826–35.

23 Pey AL, Megarity CF, Timson DJ. NAD(P)H quinone oxidoreductase (NQO1): an enzyme which needs just enough mobility, in just the right places. *Biosci Rep*. 2019;**39**:BSR20180459.

24 Mayorov A, Dal Peraro M, Abriata LA. Active site-induced evolutionary constraints follow fold polarity principles in soluble globular enzymes. *Mol Biol Evol*. 2019;**36**:1728–33.

25 Medina-Carmona E, Neira JL, Salido E, Fuchs JE, Palomino-Morales R, Timson DJ, et al. Site-to-site

interdomain communication may mediate different loss-of-function mechanisms in a cancer-associated NQO1 polymorphism. *Sci Rep.* 2017;**7**:44532.

26 Fraczkiewicz R, Braun W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comput Chem.* 1998;**19**:319–33.

27 Pey AL, Megarity CF, Timson DJ. FAD binding overcomes defects in activity and stability displayed by cancer-associated variants of human NQO1. *Biochim Biophys Acta.* 2014;**1842**:2163–73.

28 Frago S, Goñi G, Herguedas B, Peregrina JR, Serrano A, Perez-Dorado I, et al. Tuning of the FMN binding and oxido-reduction properties by neighboring side chains in *Anabaena* flavodoxin. *Arch Biochem Biophys.* 2007;**467**:206–17.

29 Nogués I, Campos LA, Sancho J, Gómez-Moreno C, Mayhew SG, Medina M. Role of neighboring FMN side chains in the modulation of flavin reduction potentials and in the energetics of the FMN:apoprotein interaction in *Anabaena* flavodoxin. *Biochemistry.* 2004;**43**:15111–21.

30 Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982;**157**:105–32.

31 Counterman AE, Clemmer DE. Volumes of individual amino acid residues in gas-phase peptide ions. *J Am Chem Soc.* 1999;**121**:4031–9.

32 Perkins SJ. Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur J Biochem.* 1986;**157**:169–80.

33 Echave J, Spielman SJ, Wilke CO. Causes of evolutionary rate variation among protein sites. *Nat Rev Genet.* 2016;**17**:109–21.

34 Zhang Z, Smith DL. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. Cambridge: Cambridge University Press; 1993.

35 Trcka F, Durech M, Man P, Hernychova L, Muller P, Vojtesek B. The assembly and intermolecular properties of the Hsp70-Tomm34-Hsp90 molecular chaperone complex. *J Biol Chem.* 2014;**289**:9887–901.

36 Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* 2022;**50**:D543–52.

37 McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor. *Bioinformatics.* 2010;**26**:2069–70.

38 Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: a one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Hum Mutat.* 2016;**37**:235–41.

39 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;**7**:248–9.

40 Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nat Protoc.* 2016;**11**:1–9.

41 Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.* 2014;**46**:310–5.

42 Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* 2011;**39**:e118.

43 Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: an Ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet.* 2016;**99**: 877–85.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Plot SM1. Scenarios for model 3.
Fig. S1. Rainbow heatmaps showing the deuteration profile for WT and active site mutants in their $NQO1_{holo}$ (A) and $NQO1_{dic}$ (B) states.
Fig. S2. Differential rainbow heatmaps for the deuteration profiles of WT and active site mutants in their $NQO1_{holo}$ and $NQO1_{dic}$ states.
Fig. S3. Analysis of EX1 exchange mechanism through peak width analysis.
Fig. S4. Data for HDX regarding selected segments in which kinetics of deuterium incorporation for the p.W106R and p.W106C mutants differ from that of WT NQO1 ($|\Delta\%D_{av}| > 10\%$).
Fig. S5. Effect of active-site mutations on the stability of the MMI in the $NQO1_{holo}$ state determined by HDX-MS.
Fig. S6. Effect of active-site mutations on the stability of the MMI in the $NQO1_{dic}$ state determined by HDX-MS.
Fig. S7. Effect of active-site mutations on the stability of the FBS in the $NQO1_{holo}$ state determined by HDX-MS.
Fig. S8. Effect of active-site mutations on the stability of the FBS in the $NQO1_{dic}$ state determined by HDX-MS.

Fig. S9. Effect of active-site mutations on the stability of the DBS in the $NQO1_{holo}$ state determined by HDX-MS.

Fig. S10. Effect of active-site mutations on the stability of the DBS in the $NQO1_{dic}$ state determined by HDX-MS.

Fig. S11. Time-dependent spectra of the NQO1 flavin reduction by NADH.

Fig. S12. Deconvolution of spectral species (A→B→C) observed during flavin reduction with NADH.

Fig. S13. Kinetics of NQO1 flavin reduction by NADH.

Fig. S14. Time-dependent spectra of the NQO1 flavin oxidation by DCPIP.

Fig. S15. Spectral deconvolution of intermediate species (A→B→C) observed during NQO1 flavin oxidation by DCPIP.

Fig. S16. Kinetics of NQO1 oxidation by DCPIP.

Fig. S17. Data for HDX regarding selected segments in which kinetics of deuterium incorporation for p.F107C and p.M155I mutants differ from that of WT NQO1 ($|\Delta\%D_{av}| > 10\%$).

Fig. S18. Data for HDX regarding selected segments in which kinetics of deuterium incorporation for p.W106R and p.W106C mutants differ from that of WT NQO1 ($|\Delta\%D_{av}| > 10\%$).

MATERIAL SUPLEMENTARIO DE LA PUBLICACIÓN 2

Counterintuitive structural and functional effects due to naturally occurring mutations targeting the active site of the disease-associated NQO1 enzyme.

Counterintuitive structural and functional effects due to naturally-occurring mutations targeting the active site of the disease-associated NQO1 enzyme

Juan Luis Pacheco-García, Ernesto Anoz-Carbonell, Dmitry S. Loginov, Daniel Kavan, Eduardo Salido, Petr Man, Milagros Medina and Angel L. Pey.

Supplementary methods

Analysis of deuterium exchange kinetic mechanism

Deuterium incorporation into amide protons of protein main-chain is interpreted using the Linderstrøm-Lang model (Model 1):

$$\text{NE-NH} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} \text{E-NH} \overset{k_{exch}}{\longrightarrow} \text{ND} \qquad \textbf{Model 1}$$
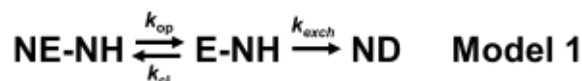
$$K_{op} = \frac{k_{op}}{k_{cl}}$$

where NE-NH and E-NH stand for non-exchanging and exchanging states and ND for the state that underwent deuterium exchange. The interconversion between states is described by the first-order rate constants k_{op}, k_{cl} and k_{exch}, and the opening reaction at equilibrium by the constant K_{op}. In this model, there are two limiting scenarios, named EX1 and EX2 exchange mechanisms. In the EX1 mechanism, $k_{exch} >> k_{cl}$, and thus, the observed exchange rate constant $k_{obs} \approx k_{op}$ (i.e. the exchange rate is determined by the *kinetics* or *dynamics* of the opening equilibrium). In the EX2 mechanism, $k_{exch} << k_{cl}$ and thus, $k_{obs} \approx K_{op} \cdot k_{exch}$ (i.e. the exchange rate is partly determined by the *thermodynamic stability* of the opening equilibrium).

To detect the presence EX1 exchange kinetics we utilized the peak width approach [1]. Briefly, peak width at 10% of the intensity was extracted for each data point in DeutEx software. Each value was then divided by an average that was calculated for every peptide using all exchange times and experimental conditions. Then, the final value indicates whether a particular number deviates from the average. For ideal peptides, exchanging using via the EX2 mechanism under all studied conditions and in identical way, the normalized value should be close to 1. Systematic shift to higher values are then indicative of peak broadening and thus of the EX1 opening kinetics being present. The

data were then plotted for each protein form (WT and mutants) and state (holo/dic). Peptides are represented as individual horizontal lines and the time dimension is captured through the red gradient. Using this approach, we demonstrate that significant EX1 contribution can only be seen in a few segments of the p.M155I and p.H162N mutants in the NQO1$_{holo}$ state and that Dic binding mostly abolishes the EX1.

Supplementary models

A plausible and simple model to explain enzyme kinetics of active-site mutants

Rapid enzyme kinetics for the FAD reduction by NADH has been modeled for WT and several NQO1 mutants under pre-steady conditions using model 2:

$$A \xrightarrow{k_{obs1}} B \xrightarrow{k_{obs2}} C \quad \textbf{Model 2}$$
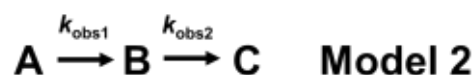
Where A, B and C are *spectral species* (not truly *states*) and k_{obs1} and k_{obs2} are the observed rate constants for *fast* (A→B) and *slow* (B→C) FAD reduction pathways at a fixed NQO1$_{holo}$ concentration. Both rate constants show a hyperbolic dependence on [NADH] and allow to determine apparent K_d values ($K_{d(NADH)}$) and the limiting values of hydride transfer (HT) rate constants (k_{HT}) using equation 1:

$$k_{obs} = \frac{k_{HT} \cdot [NADH]}{K_{d(NADH)} + [NADH]} \quad \textbf{(Equation 1)}$$

We must note that this modeling is phenomenological and applies a Ockham´s razor approach (this three-state kinetic model provided the best results for spectral deconvolution and time evolution of FAD spectra)[2–4].

As for all kinetic models, we aim to provide the simplest model that describes experimental data and hopefully get insight into reaction mechanisms. For instance, in the WT system pseudo-first order conditions are only satisfied when NADH concentration largely exceeds NQO1 concentration (e.g. 100 µM NADH). Under these conditions, the *fast* reduction pathway is close to detection limit of the instrument (an apparent half-life of 2 ms at 6ºC). In addition, we are not truly dealing with states but with spectral species. Therefore, our results provide pseudo-mechanistic information suitable for comparison and discussion of the results of WT and mutants enzyme kinetics.

Consider a simple and general model that may explain the different behavior between WT (pseudo-first order kinetics) and some mutants investigated in this work (pseudo zero-order kinetics) (Model 3) for each step (fast, A→B; slow, B→C):

$$NQO1_{holo} + NADH \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} NQO1_{holo}:NADH \xrightarrow{k_{HT}} NQO1_{red} + NAD^+ \qquad \text{Model 3}$$

$$K_{d(NADH)} = \frac{k_{off}}{k_{on}}$$
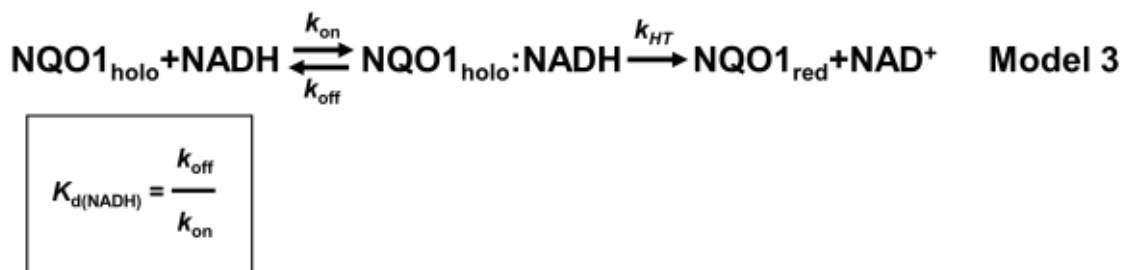
The first step represents the binding equilibrium between $NQO1_{holo}$ (defined by a dissociation constant $K_{d(NADH)}$) and the NADH, whereas the second step is the reductive step for FAD (described by the first-order constant, k_{HT}). The hyperbolic response of k_{obs} for $NQO1_{holo}$ WT on [NADH] simply reflects a change in the rate-limiting step as the [NADH] raises (Plot SM1). At low [NADH] ($<< K_{d(NADH)}$) [2–4], this resembles an EX2 scenario (the red line in the plot). Here the k_{obs} linearly grows on [NADH], and thus we observe first order kinetics regarding [NADH] (note that in this scenario, k_{obs} depends on k_{on}, k_{off} and k_{HT}). At [NADH] $>> K_{d(NADH)}$, we observe an EX1 scenario (the horizontal green line in the plot). In this scenario, as we are reaching the k_{HT} the kinetics becomes of zero-order vs. [NADH] (i.e. a horizontal straight line).

How a mutation can change the hyperbolic dependence (observed for WT NQO1) to [NADH]-independent behavior or to a linear dependence on [NADH]?. Reaction Model 3 provides simple and reasonable answers: the mutation affects the values of one or several of the rate constants (k_{on}, k_{off} and/or k_{HT}) to shift the observed behavior to an apparently pure EX1 or EX2 mechanism under the experimental range of NADH concentrations (losing the EX2 to EX1 transition with [NADH] eliminates the hyperbolic behavior). Obviously, this may be an oversimplification for NQO1 enzyme kinetics but nicely illustrates how mutations could alter the [NADH]-dependent behavior.

Plot SM1. Scenarios for model 3. The black line shows the hyperbolic saturation function on [NADH] concentration. Green and red lines show the two limiting scenarios: at low [NADH], k_{obs} linearly grows on [NADH] (EX2 mechanism) whereas at high [NADH], k_{obs} becomes independent on [NADH] (EX1 mechanism). The curvature observed in the hyperbolic behavior (evident at [NADH] $> 0.2 \cdot K_{d(NADH)}$) indicates the onset of the EX1 mechanism.

EX2 Mechanism; [NADH] $\ll K_{d(NADH)}$; $k_{obs} = k_{HT} \cdot$ [NADH]$/K_{d(NADH)}$

EX1 Mechanism; [NADH] $\gg K_{d(NADH)}$; $k_{obs} = k_{HT}$

Hyperbolic behavior
[NADH]-dependent EX2→EX1 shift

$$k_{obs} = \frac{k_{HT} \cdot [NADH]}{K_{d(NADH)} + [NADH]}$$

$K_{d(NADH)}$ = [NADH]
at which $k_{obs} = k_{HT}/2$

k_{obs}/k_{HT}

[NADH] (µM)
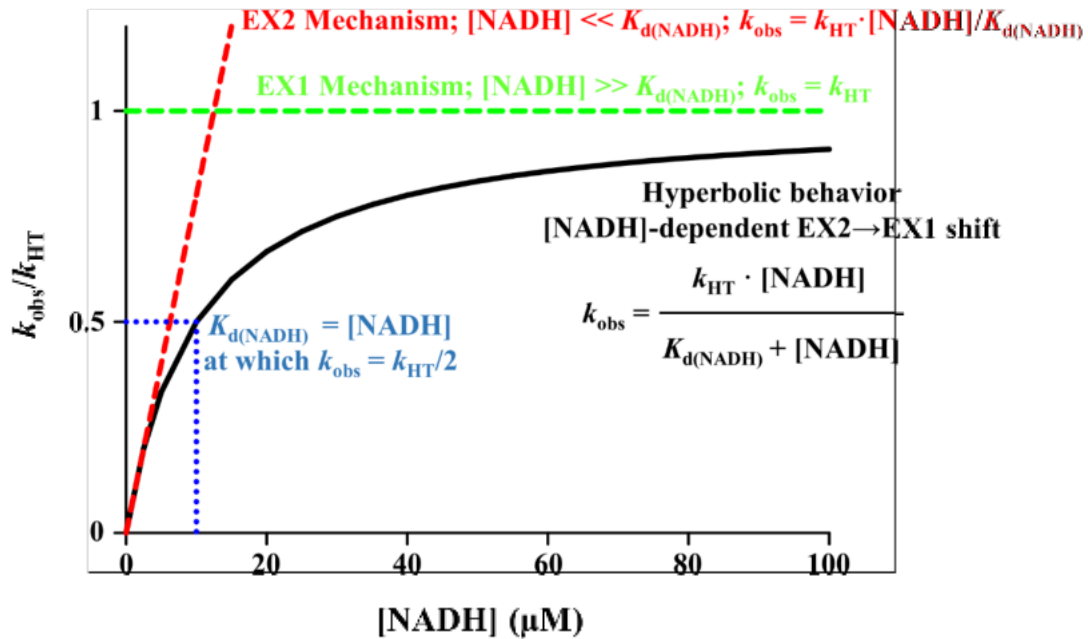
Supplementary data

Figure S1. Rainbow heatmaps showing the deuteration profile for WT and active site mutants in their NQO1$_{holo}$ (A) and NQO1$_{dic}$ (B) states. Each thick section represents a NQO1 variant (indicated on the left) and is subdivided into the individual time points. The x-axis represents the primary sequence. The secondary structure elements are shown above the heatmap according to [5] (PDB: 1D4A) as well as the sites of mutation. The residues forming the MMI, the DBS and the FBS are also depicted. Color scale in %D is displayed below the heatmap. Heatmaps were plotted with MSTools - http://peterslab.org/MSTools/index.php [6].

A
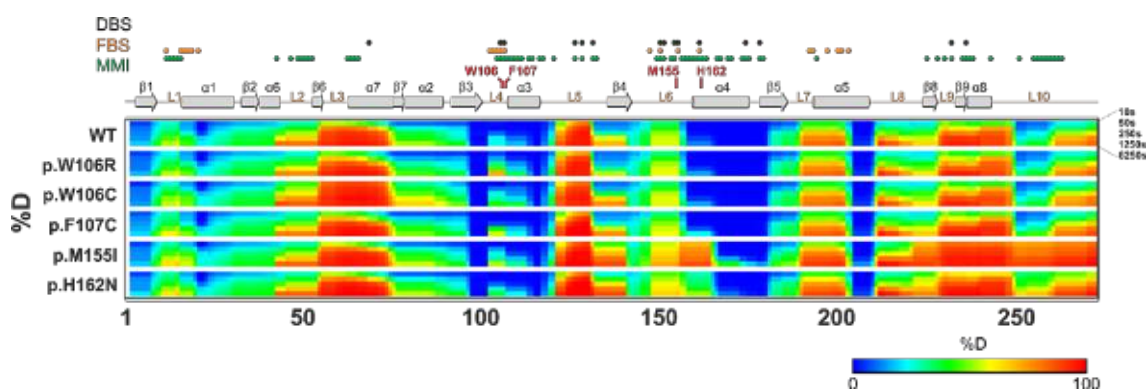
B
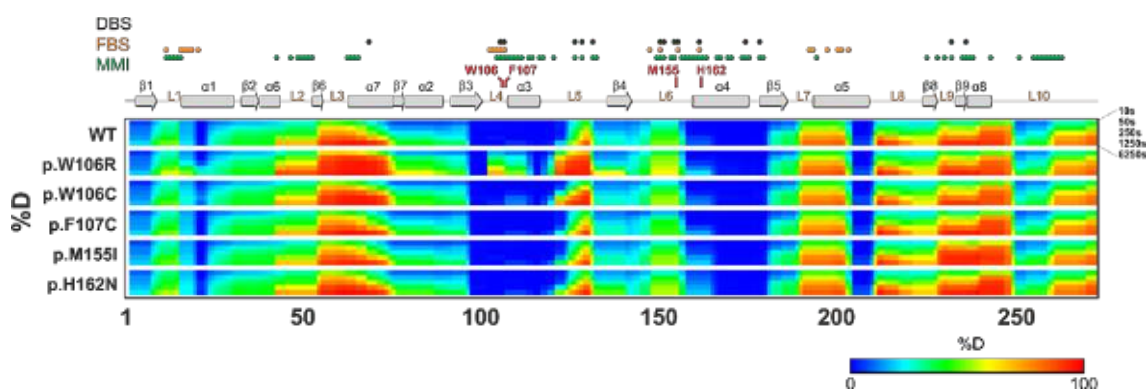
Figure S2. Differential rainbow heatmaps for the deuteration profiles of WT and active site mutants in their NQO1holo and NQO1dic states. A) Differential heatmaps between NQO1dic and NQO1holo states for individual NQO1 variants representing the effect of Dic binding. B and C) Differential heatmaps between a given mutant and the WT protein in the NQO1holo (B) and NQO1dic (C) states representing the effect of a given mutation on the stability of the corresponding state *vs.* the WT protein. Each thick section represents a NQO1 variant (indicated on the left of panel A) or the difference between a variant and the WT protein (indicated on the left of panels B-C) and is subdivided into the individual time points. The x-axis represents the primary sequence. The secondary structure elements are shown above the heatmap according to [5] (PDB: 1D4A) as well as the sites of mutation. The residues forming the MMI, the DBS and the FBS are also depicted. Color scale in %D is displayed below the heatmap. Heatmaps were plotted with MSTools - http://peterslab.org/MSTools/index.php [6].
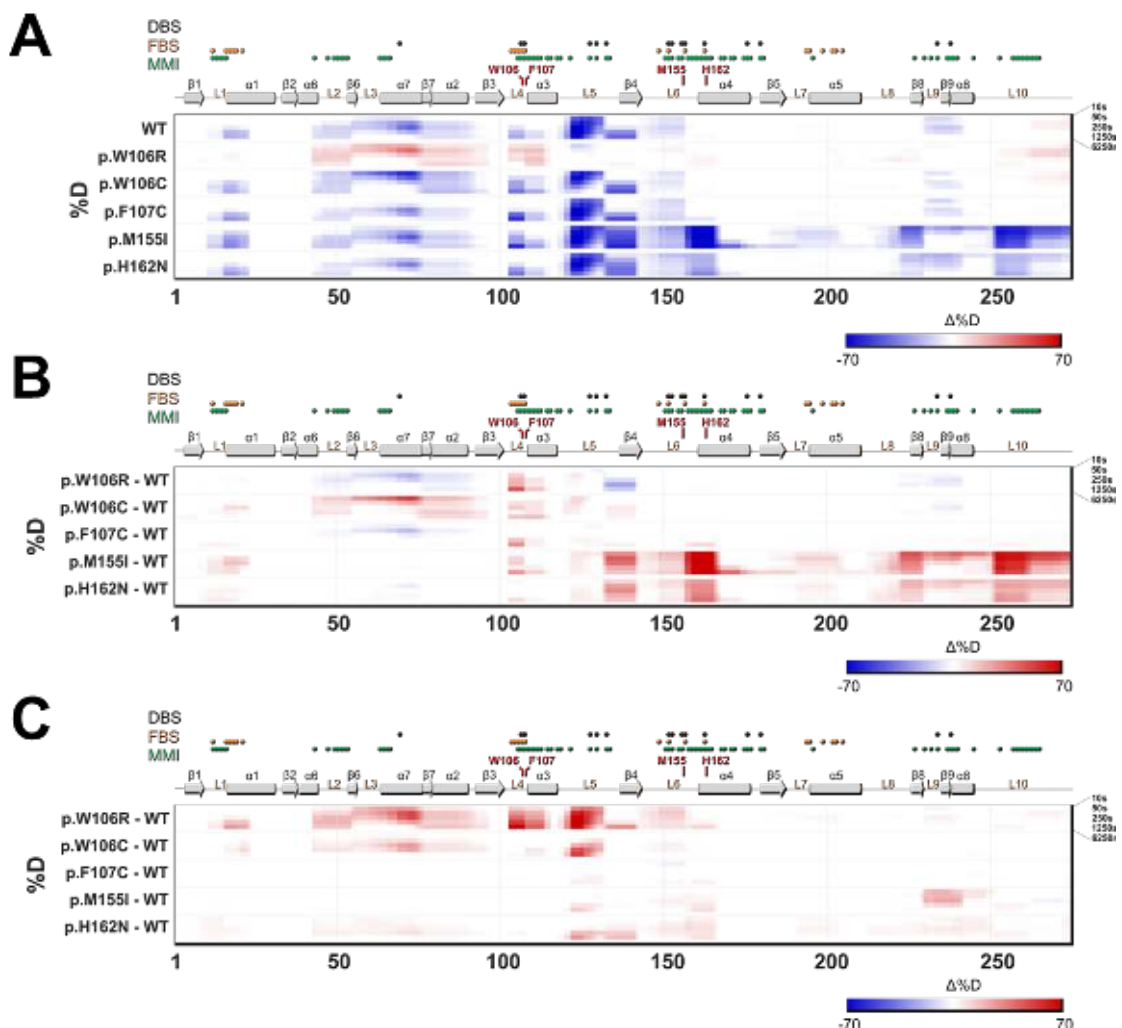
Figure S3. Analysis of EX1 exchange mechanism through peak width analysis.
Details on these analyses can be found in supplementary methods. Left and right panels show data for NQO1$_{holo}$ (h) and NQO1$_{dic}$ (d) states, respectively. Individual lines are the peptides and the time-resolved kinetics is captured through the red gradient (shown above each graph, in seconds). Systematic shift toward values > 1 indicates presence of EX1 exchange kinetics. Analysis was performed at 10% of the peak height. Only a few regions show more prominent EX1 in the NQO1$_{holo}$ state of M155I and H162N. Heatmaps were plotted with MSTools - http://peterslab.org/MSTools/index.php [6].

hF107C

dF107C

hM155I

dM155I

hH162N

dH162N
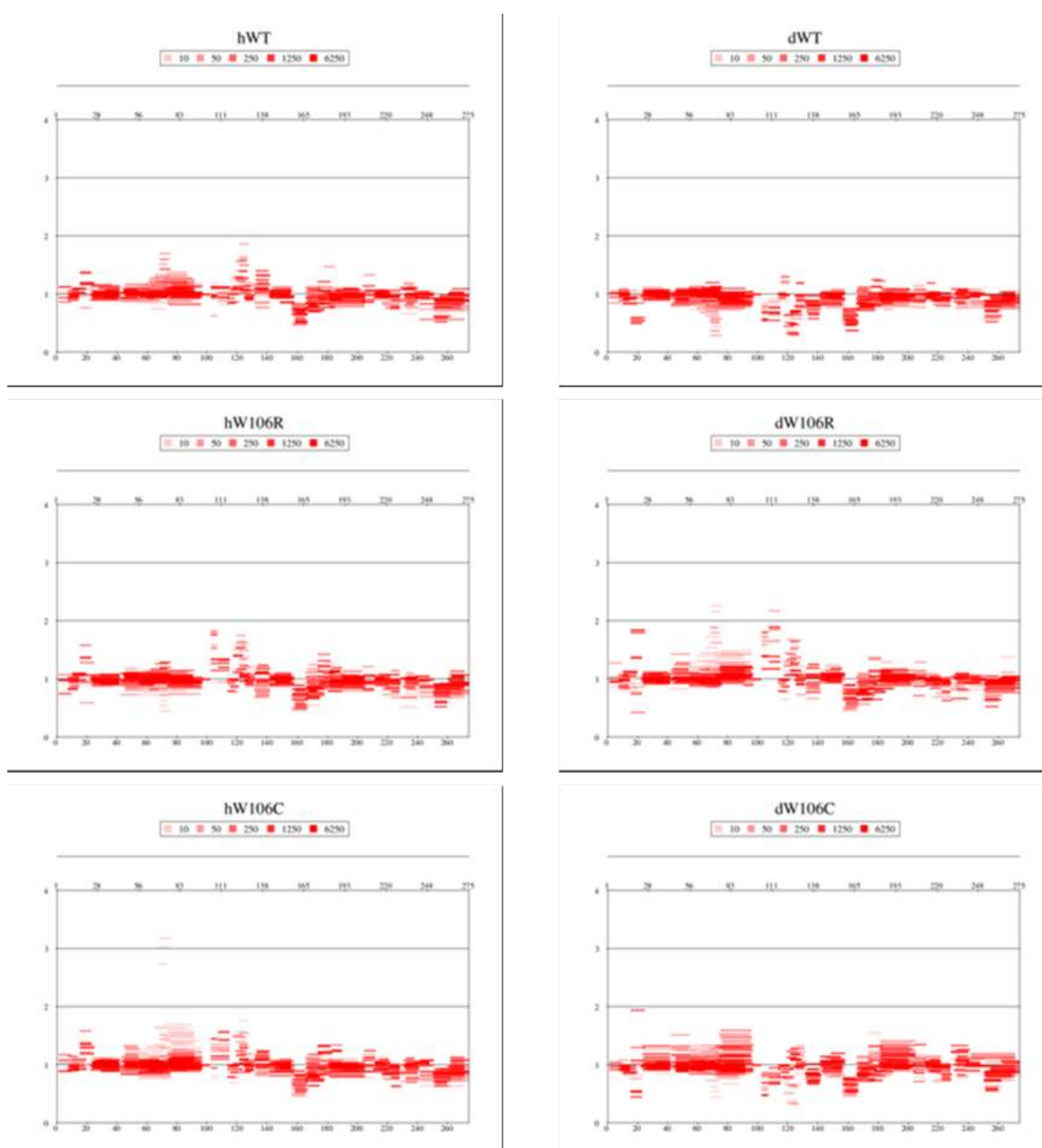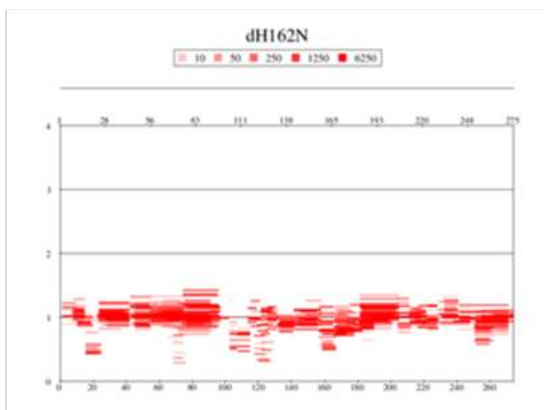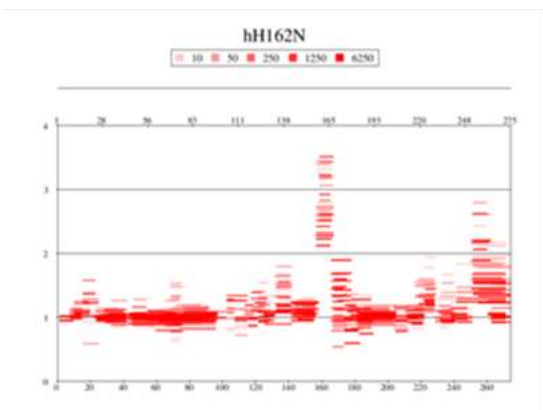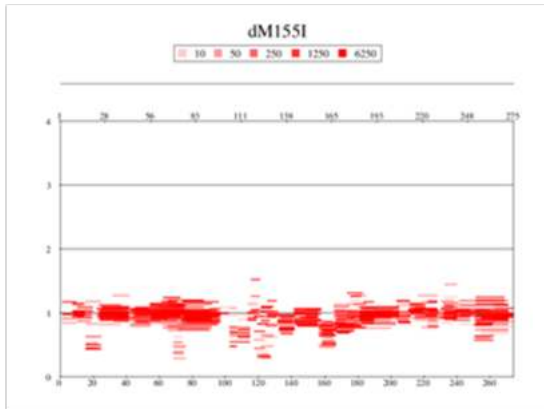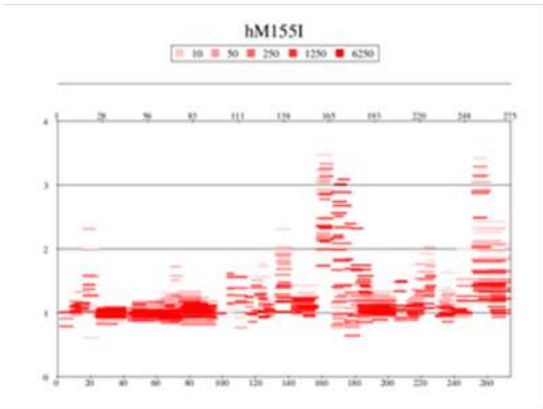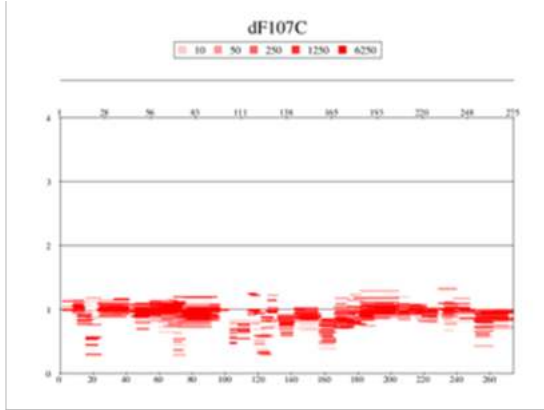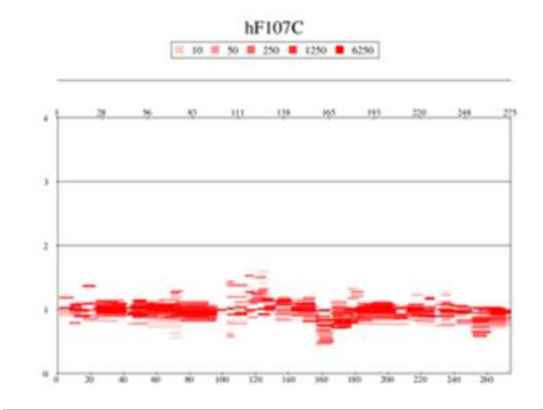
Figure S4. Data for HDX regarding selected segments in which kinetics of deuterium incorporation for the p.W106R and p.W106C mutants differ from that of WT NQO1 (|Δ%D$_{av}$| > 10%). Lines show best-fits to either single or double exponential functions with an initial burst-phase [7].
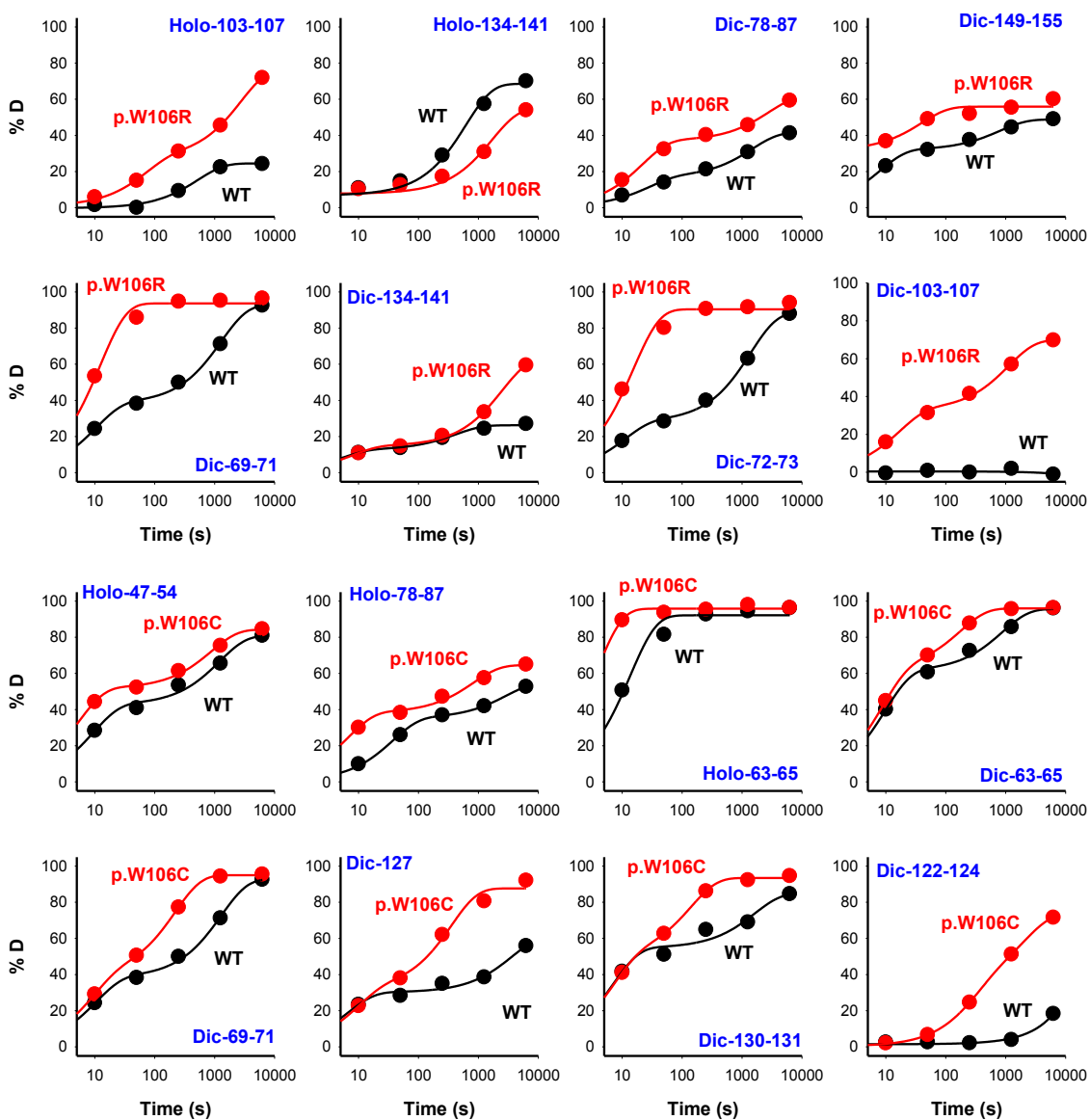
Figure S5. Effect of active-site mutations on the stability of the MMI in the NQO1holo state determined by HDX-MS. Panel A shows the location of the FAD (in black spheres) and the MMI (in blue spheres) in the WT structure. Panels B-F show the mutational effects on the MMI stability. Coloring of B-F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [8].
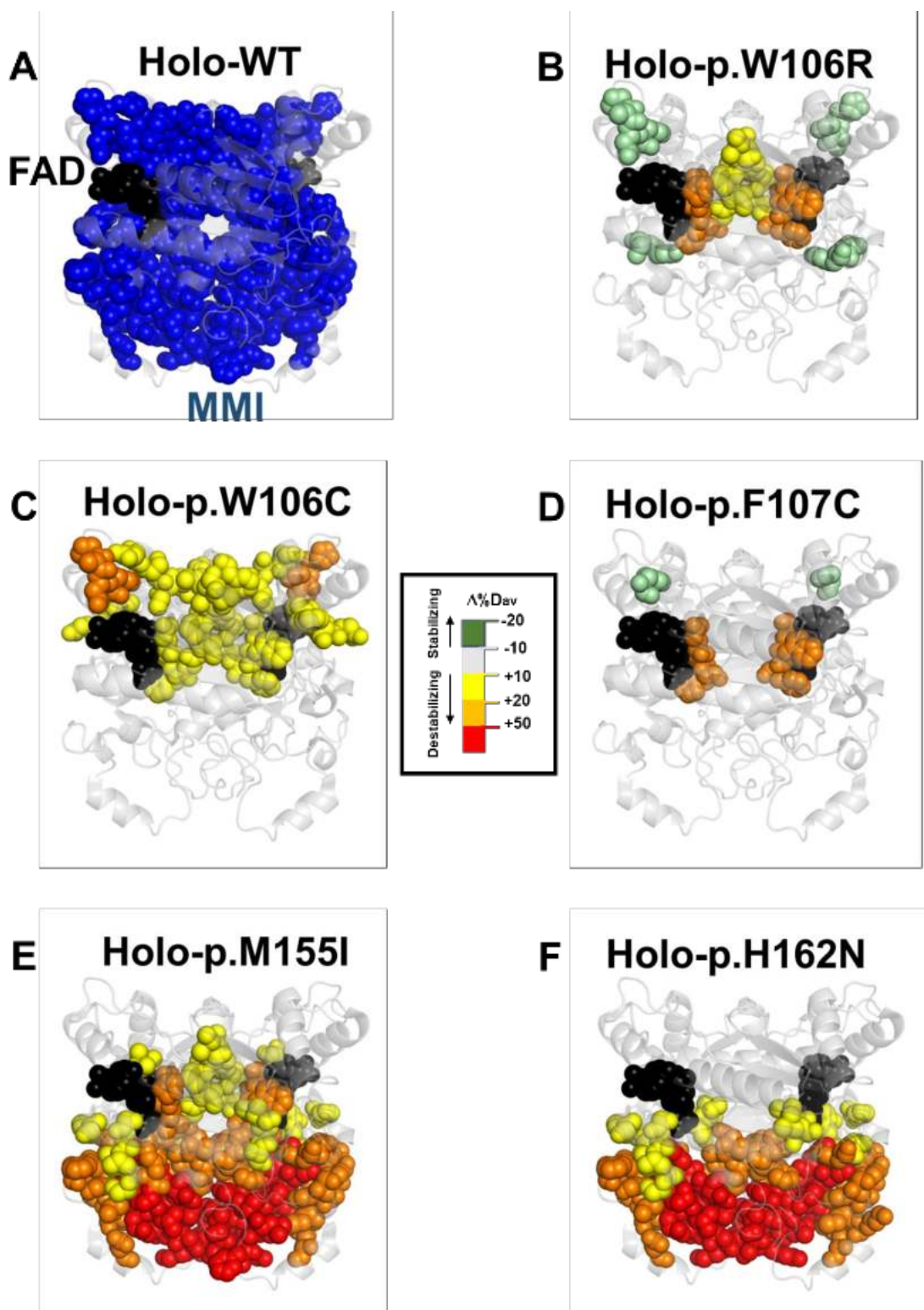
Figure S6. Effect of active-site mutations on the stability of the MMI in the NQO1$_{dic}$ state determined by HDX-MS. Panel A shows the location of the FAD (in grey spheres), Dic (in black spheres) and the MMI (in blue spheres) in the WT structure. Panels B-F show the mutational effects on the MMI stability. Coloring of B-F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [8].
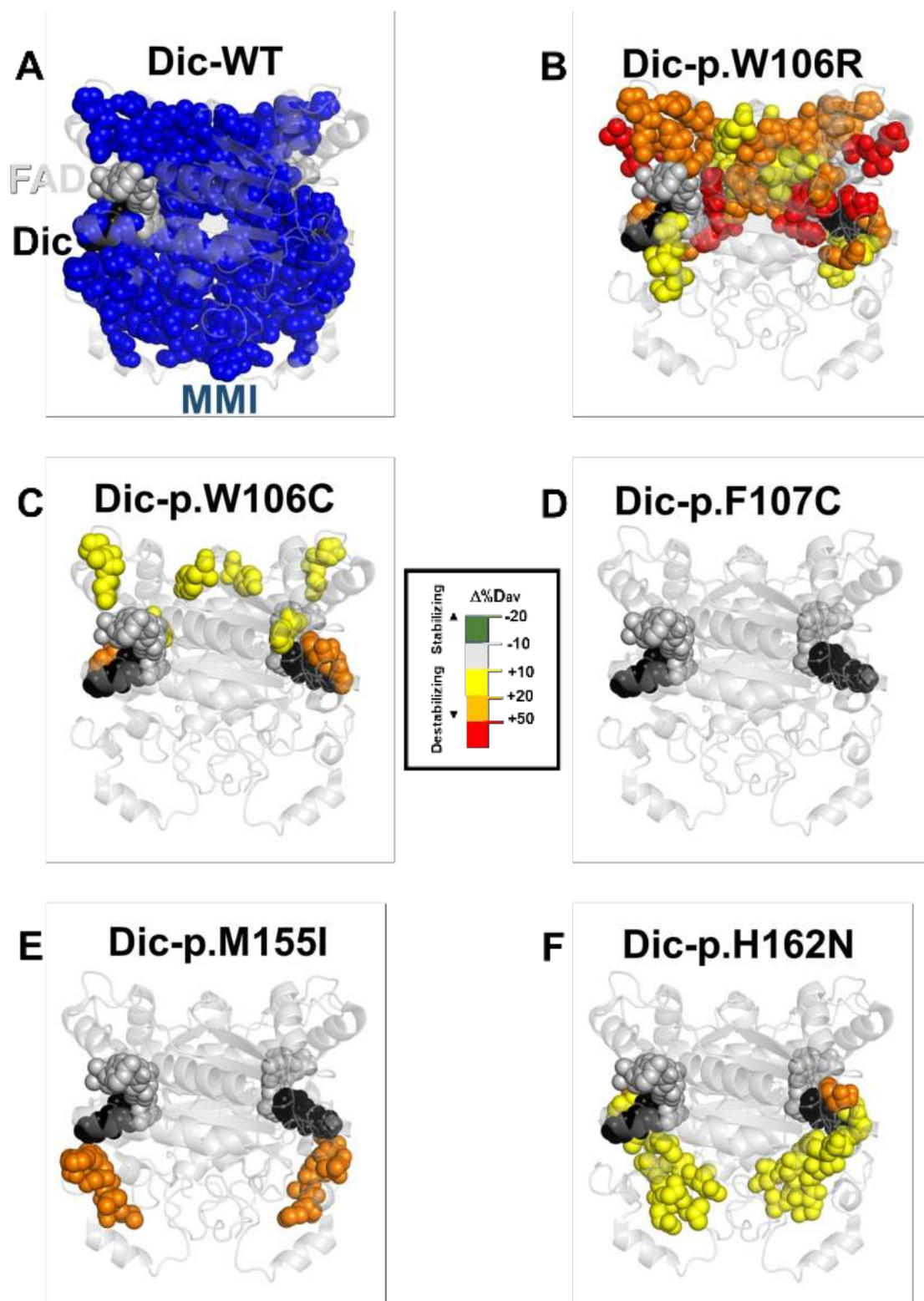
Figure S7. Effect of active-site mutations on the stability of the FBS in the NQO1$_{holo}$ state determined by HDX-MS. Panel A shows the location of the FAD (in black spheres) and the FBS (in blue spheres) in the WT structure. Panels B-F show the mutational effects on the FBS stability. Coloring of B-F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [8].
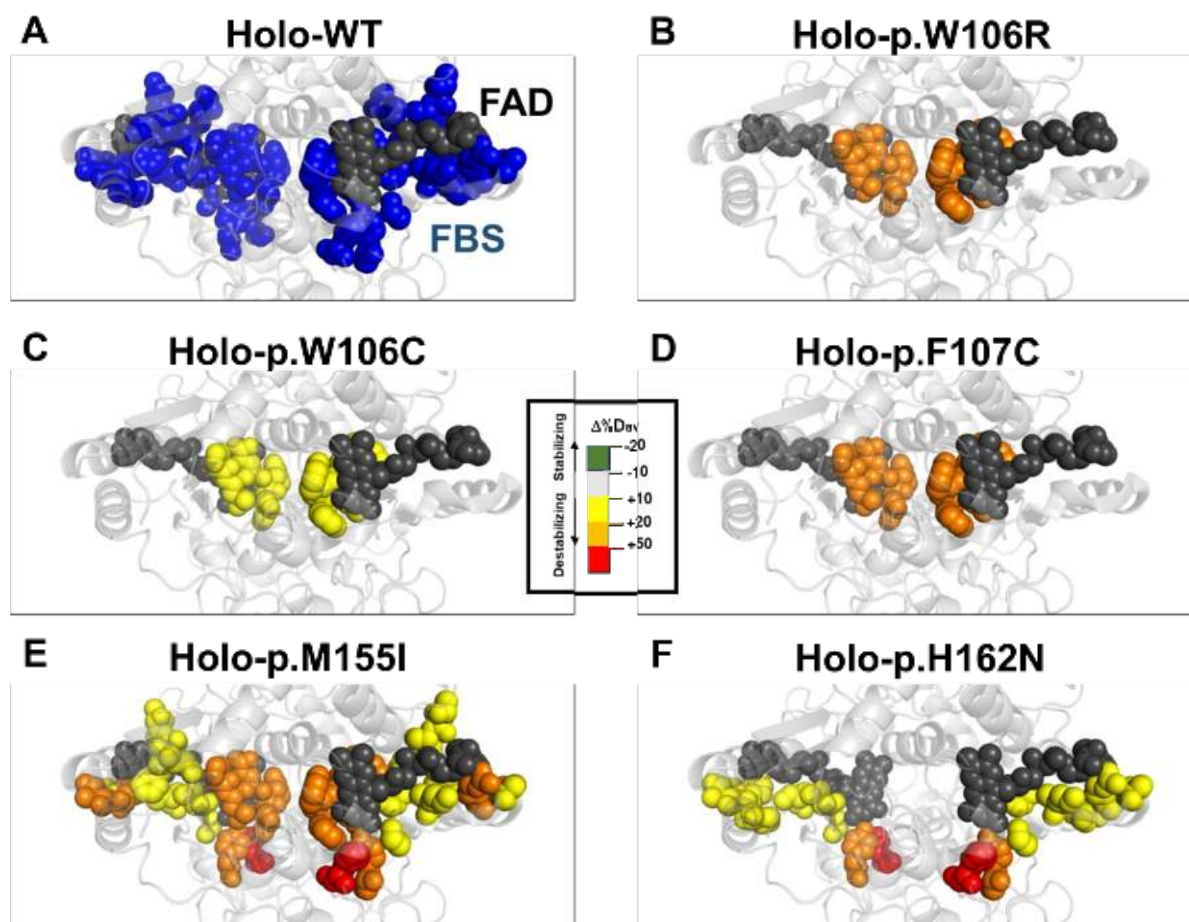
Figure S8. Effect of active-site mutations on the stability of the FBS in the NQO1$_{dic}$ state determined by HDX-MS. Panel A shows the location of the FAD (in grey spheres), Dic (in black spheres) and the FBS (in blue spheres) in the WT structure. Panels B-F show the mutational effects on the FBS stability. Coloring of B-F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [8].
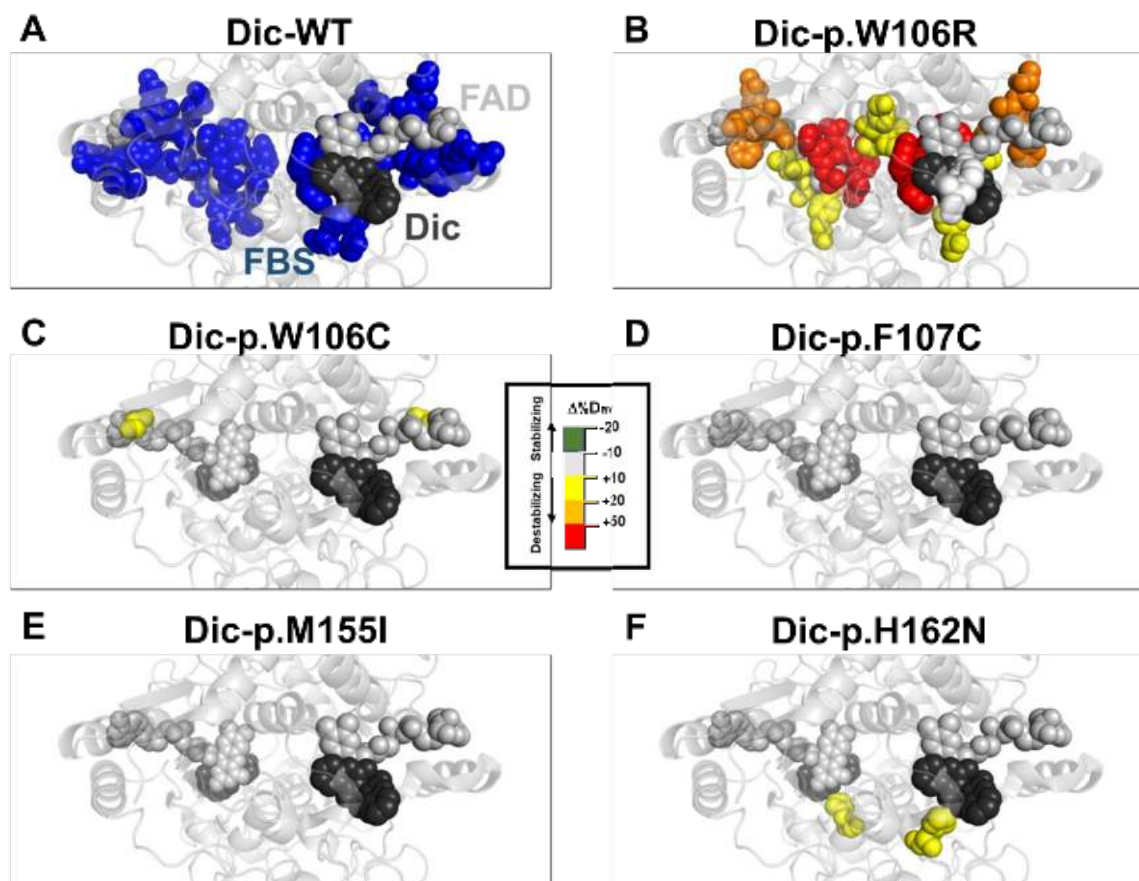
Figure S9. Effect of active-site mutations on the stability of the DBS in the NQO1$_{holo}$ state determined by HDX-MS. Panel A shows the location of FAD (in black spheres) and the DBS (in blue spheres) in the WT structure. Panels B-F show the mutational effects on the DBS stability. Coloring of B-F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [8].
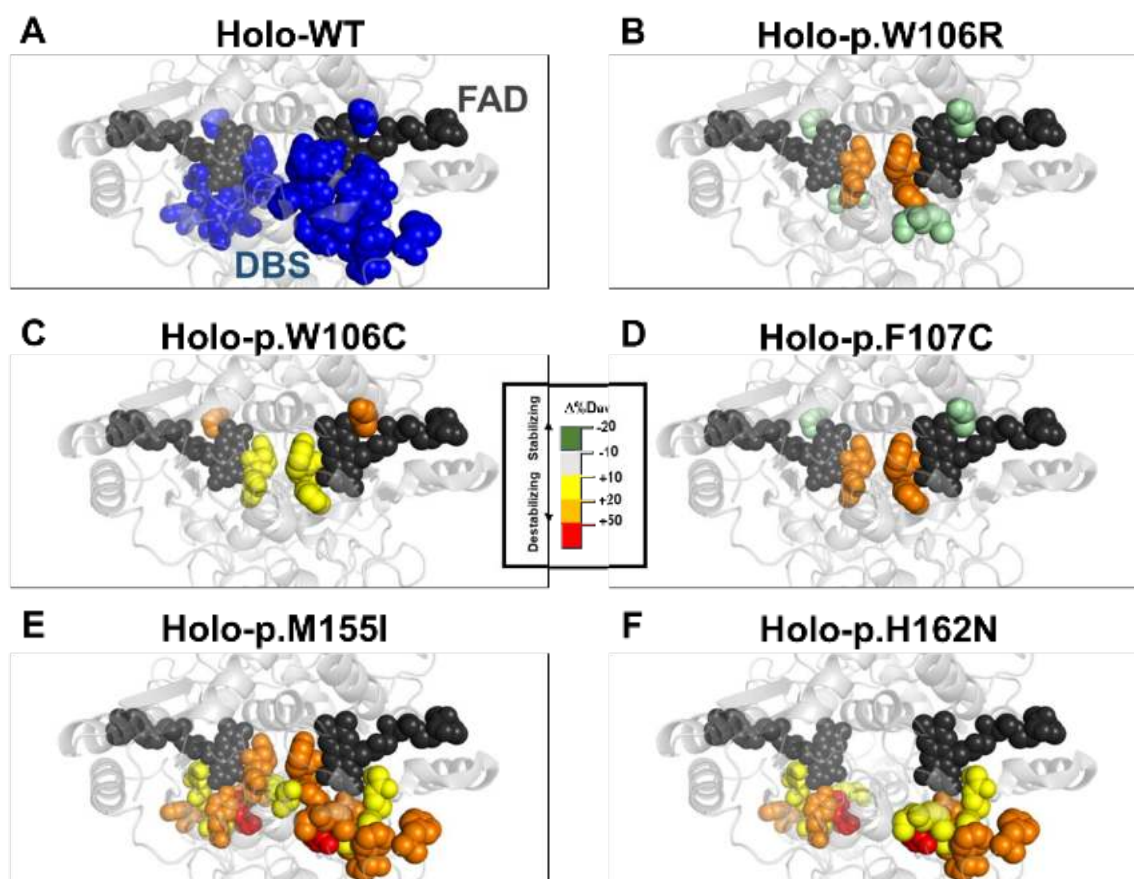
Figure S10. Effect of active-site mutations on the stability of the DBS in the NQO1$_{dic}$ state determined by HDX-MS. Panel A shows the location of FAD (in grey spheres), Dic (black spheres) and the DBS (in blue spheres) in the WT structure. Panels B-F show the mutational effects on the DBS stability. Coloring of B-F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [8].
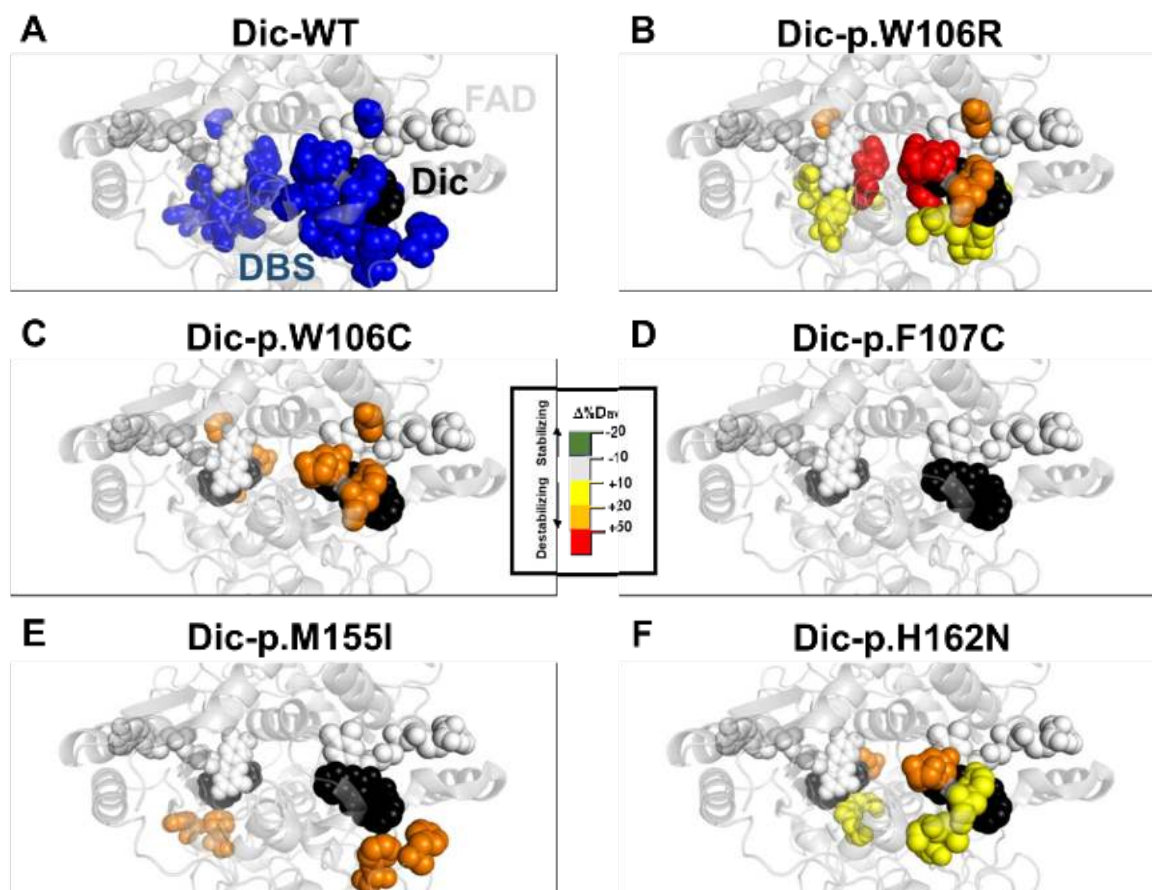
Figure S11. Time-dependent spectra of the NQO1 flavin reduction by NADH. Spectral evolution after mixing NQO1$_{holo}$ (7.5 μM) with NADH (7.5 μM) in 20 mM HEPES-KOH, pH 7.4, at 6 °C. Panels A-F correspond to the variants indicated. Different colored lines correspond to the spectra at different reaction times. The blackline represents the spectrum of NQO1$_{holo}$ before mixing. Data are from a single measurement and representative from n > 3.
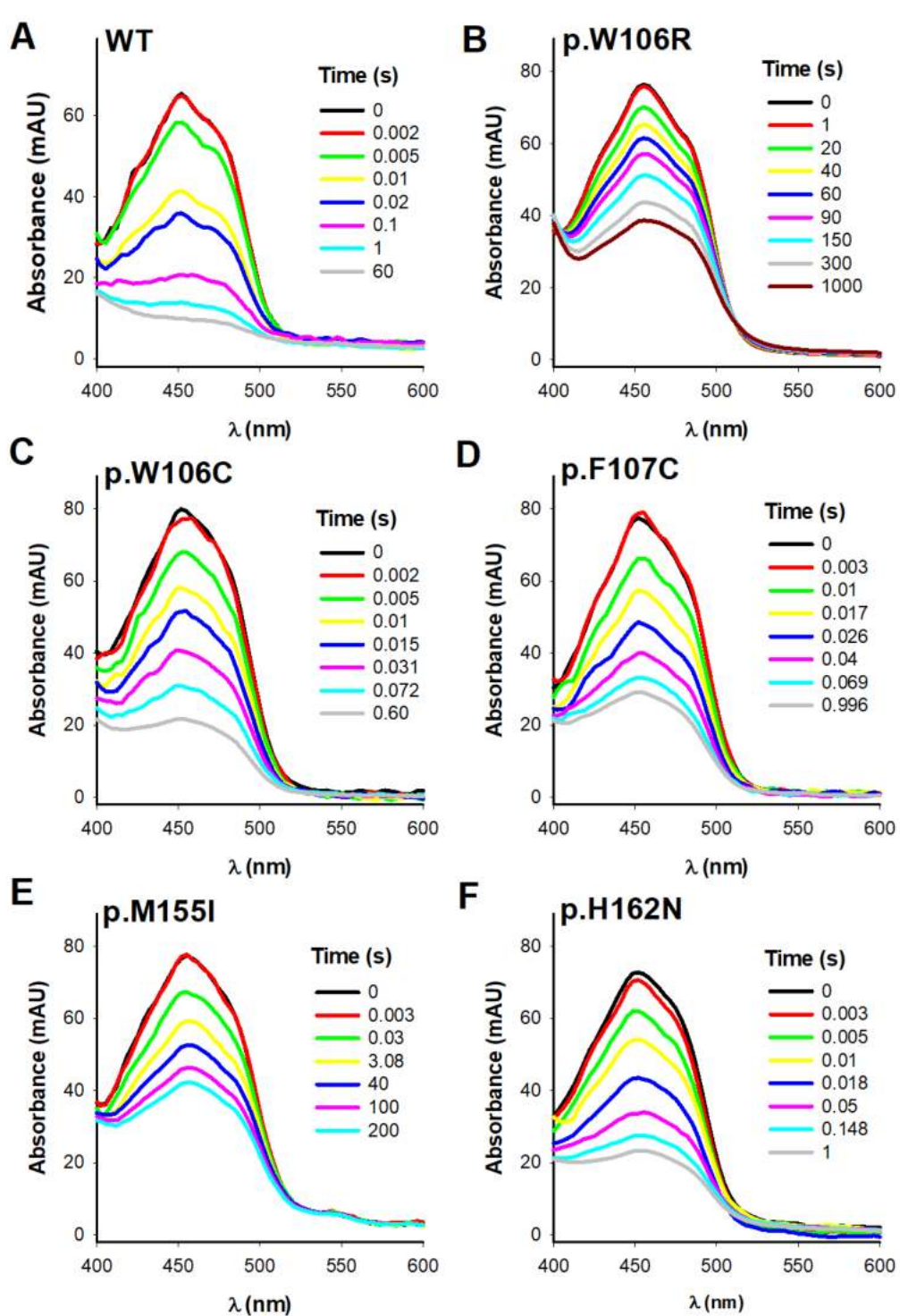
Figure S12. Deconvolution of spectral species (A→B→C) observed during flavin reduction with NADH. A-F show deconvolution for the NQO1 variants indicated from a single measurement and representative from n > 3. The % of the A_{450} for the reaction (A→C) corresponding to the fast (A→B) and slow (B→C) steps are indicated. The % of the total A_{450} (corresponding to that of A) evolved during the entire reaction (A→B→C) is also shown. Data are from a single measurement and representative from n > 3. Temperature was 6°C.
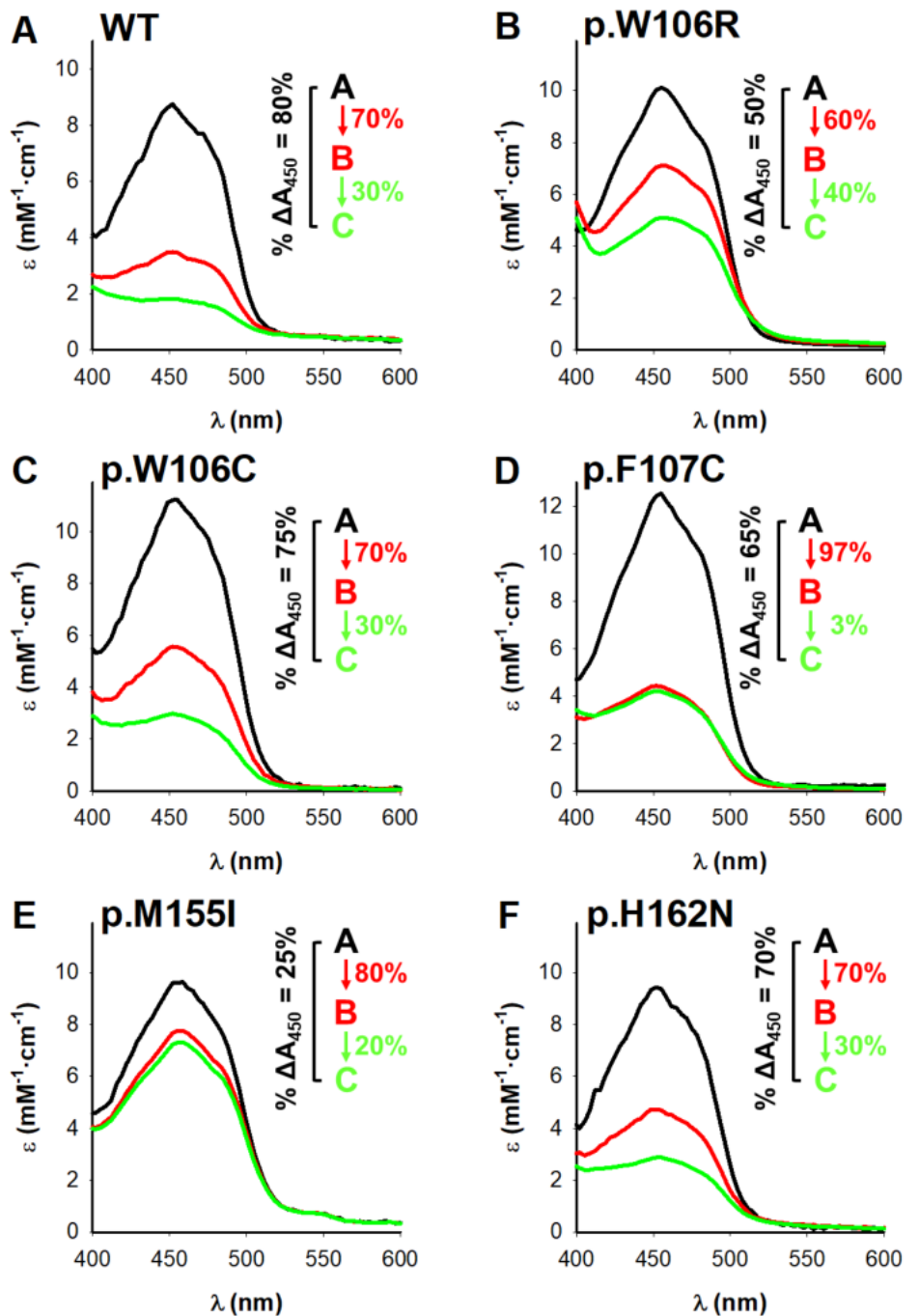
Figure S13. Kinetics of NQO1 flavin reduction by NADH. Reduction was followed by changes in A_{450} (black circles). Lines are best fits to a two-step model (A→B→C) (red lines). Insets show the corresponding residuals (red circles, in mAU). Transients were recorded after mixing $NQO1_{holo}$ (7.5 µM) with NADH (7.5 µM) in 20 mM HEPES-KOH, pH 7.4, at 6 °C. Panels A-F correspond to the variants indicated.
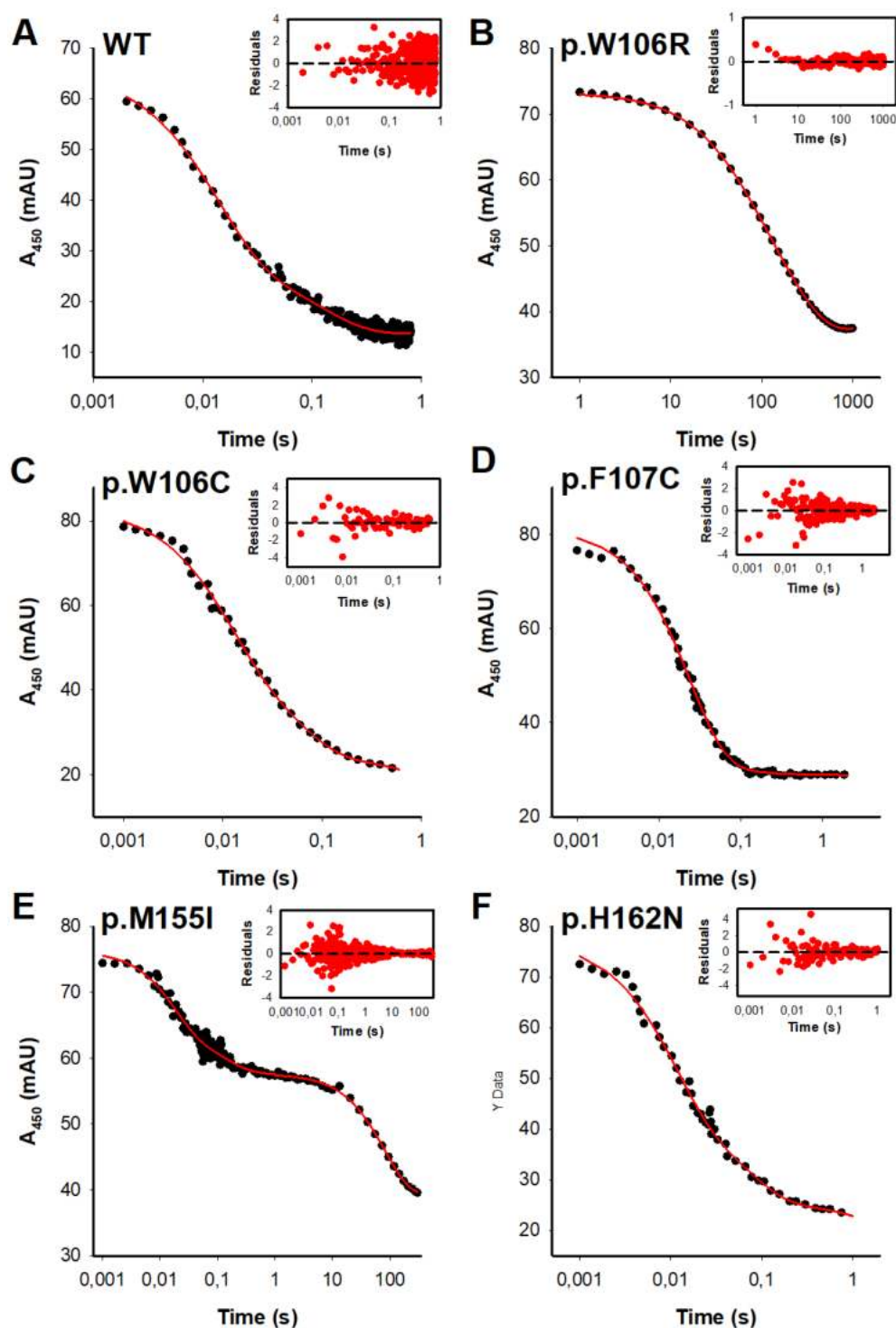
Figure S14. Time-dependent spectra of the NQO1 flavin oxidation by DCPIP. Spectral evolution after mixing NQO1$_{red}$ (produced by previous mixing 7.5 μM protein with 7.5 μM NADH) with DCPIP (7.5 μM) in 20 mM HEPES-KOH, pH 7.4, at 6 °C. Panels A-D correspond to the variants indicated. The mutations p.W106R and p.M155I were not investigated due to the inability to produce of their fully FADH$_2$ states upon NADH-reduction. Different colored lines correspond to the spectra at different reaction times. Data are from a single measurement and representative from n > 3.
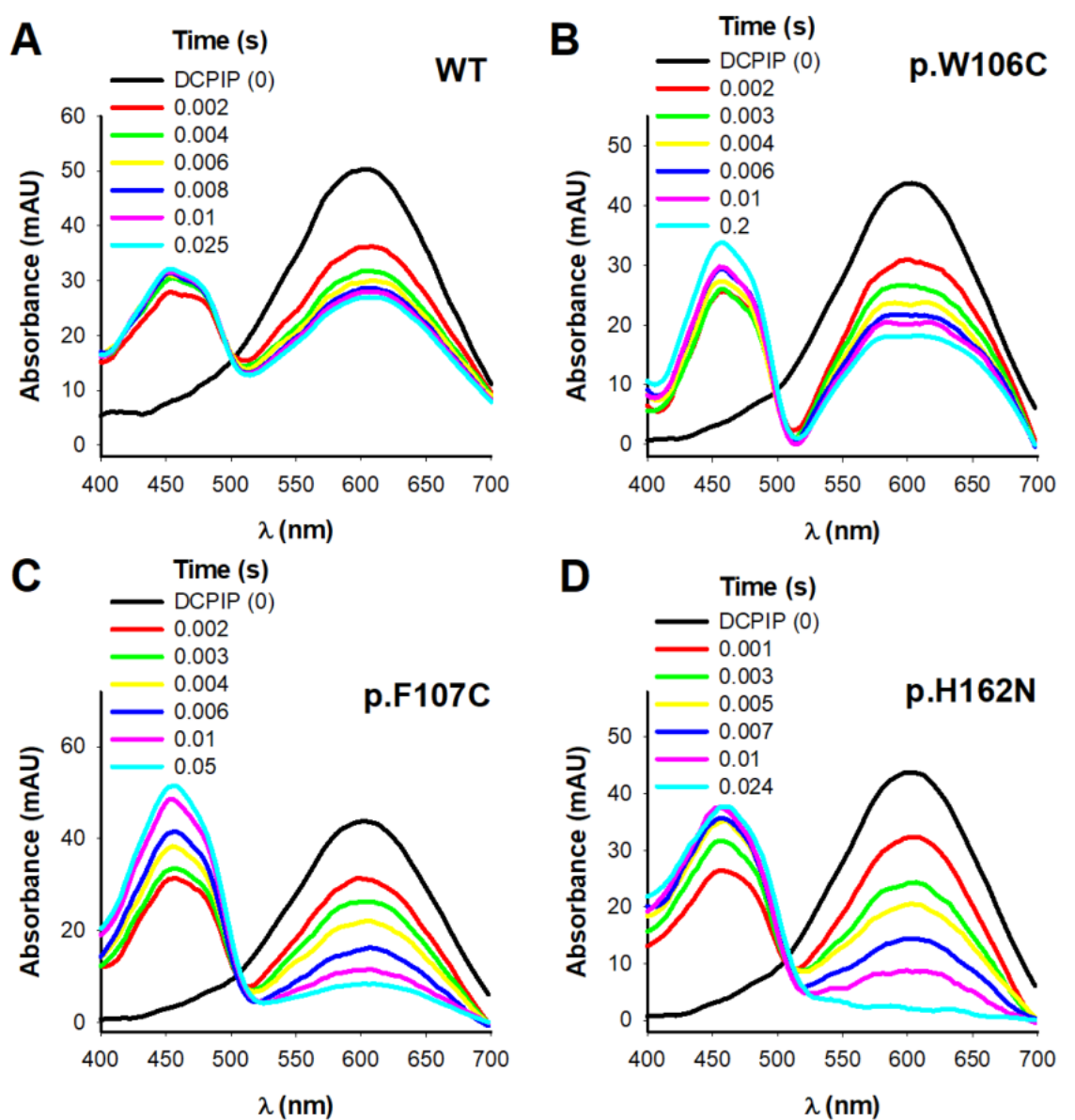
Figure S15. Spectral deconvolution of intermediate species (A→B→C) observed during NQO1 flavin oxidation by DCPIP. Panels A-D show data for the variants indicated from a single measurement and representative from n > 3. Temperature was 6°C.
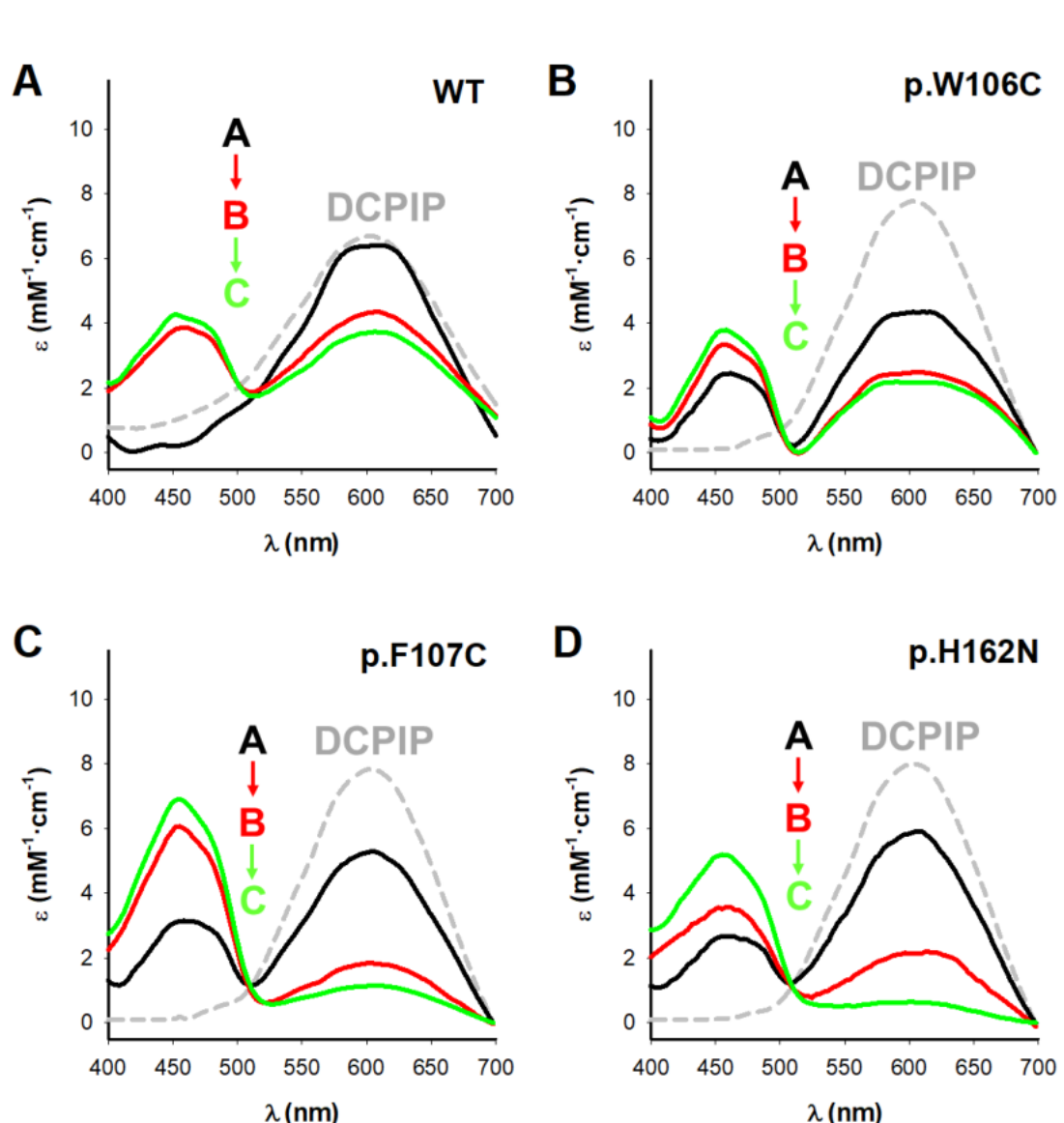
Figure S16. Kinetics of NQO1 oxidation by DCPIP. Oxidation kinetics was followed by changes in A_{450} (black circles) and A_{600} (red circles). Lines are best fits to a two-step model (A→B→C). Insets show the corresponding residuals (red circles, in mAU for fittings using A_{600}). Transients were recorded after mixing $NQO1_{red}$ (produced by previously mixing 7.5 µM protein with 7.5 µM NADH) with DCPIP (7.5 µM) in 20 mM HEPES-KOH, pH 7.4, at 6 °C. Panels A-D correspond to the variants indicated.
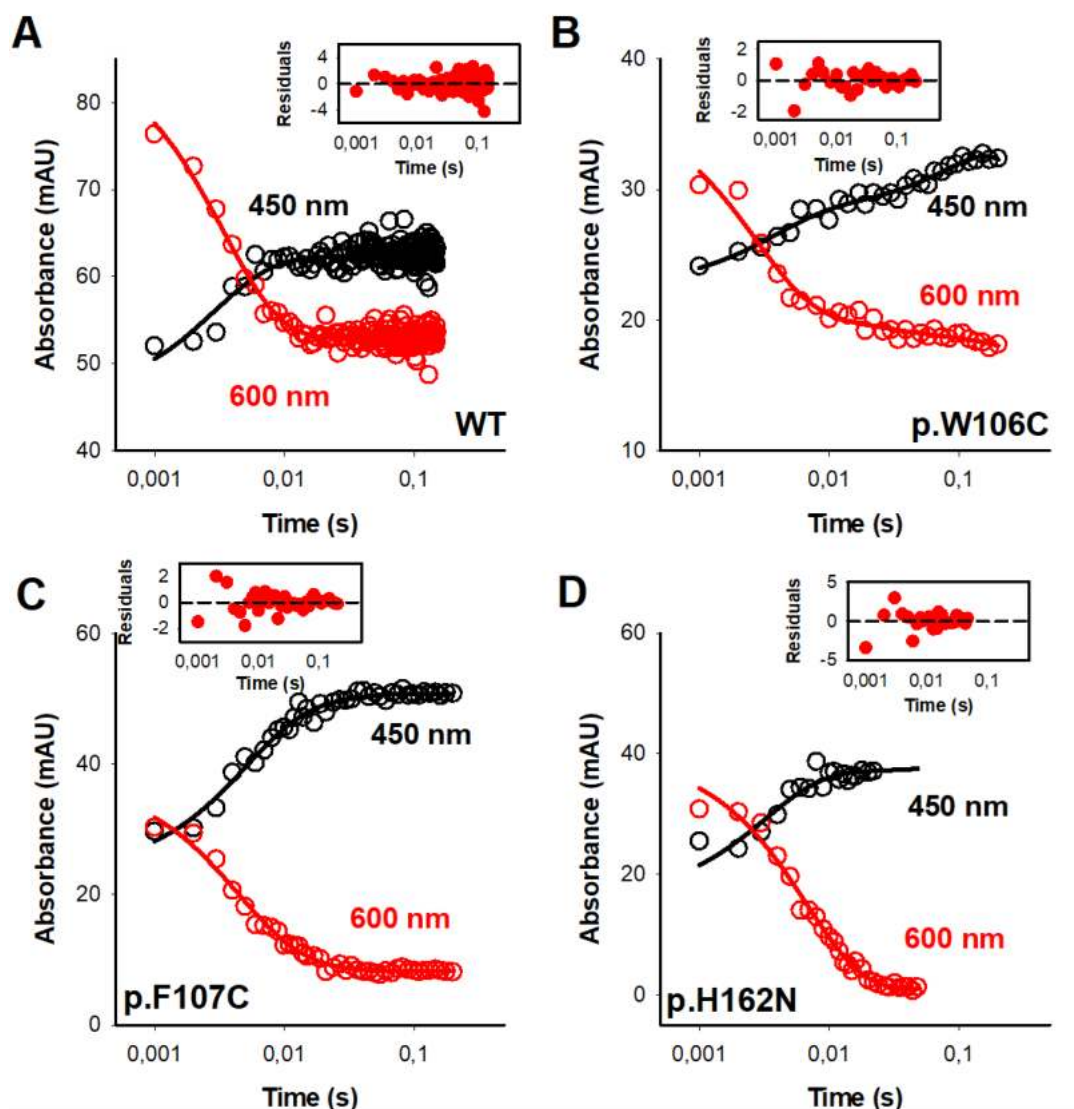
Figure S17. Data for HDX regarding selected segments in which kinetics of deuterium incorporation for p.F107C and p.M155I mutants differ from that of WT NQO1 (|Δ%D_av| > 10%). Lines show best-fits to either single or double exponential functions with an initial burst-phase [7].
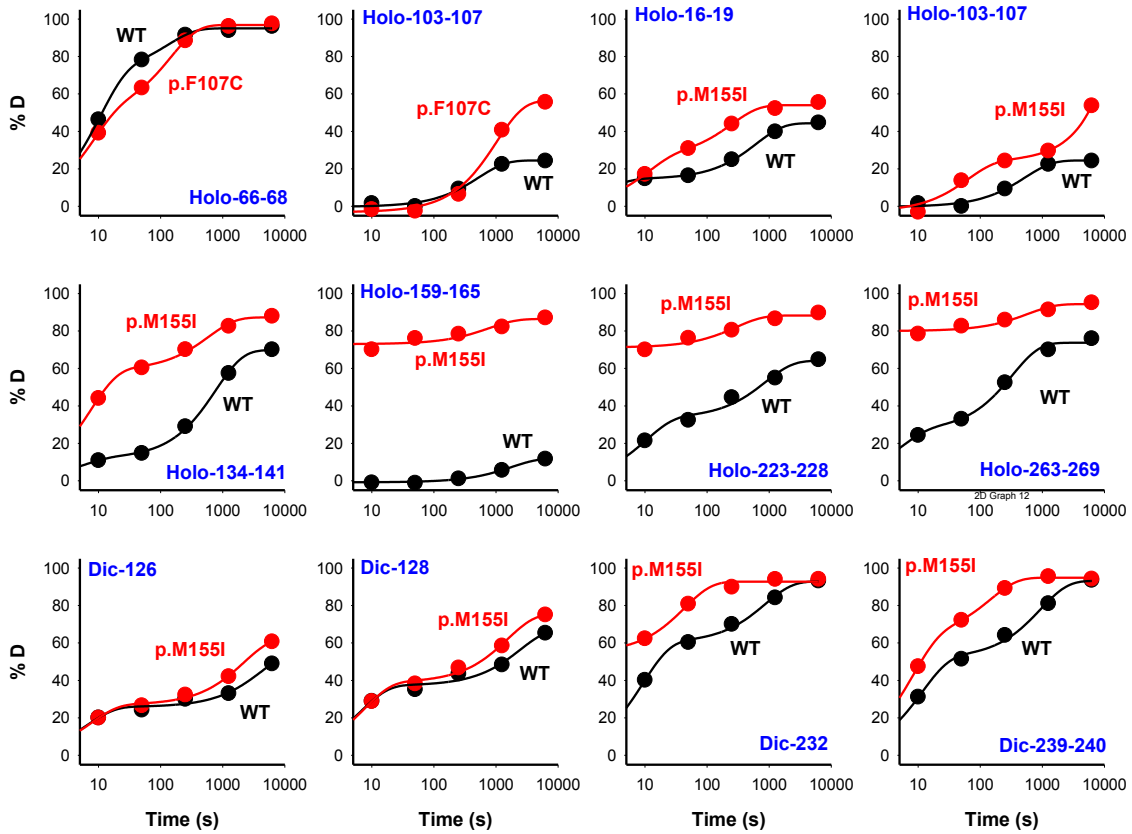
Figure S18. Data for HDX regarding selected segments in which kinetics of deuterium incorporation for p.W106R and p.W106C mutants differ from that of WT NQO1 ($|\Delta\%D_{av}| > 10\%$). Lines show best-fits to single or double exponential functions with an initial burst-phase [7].
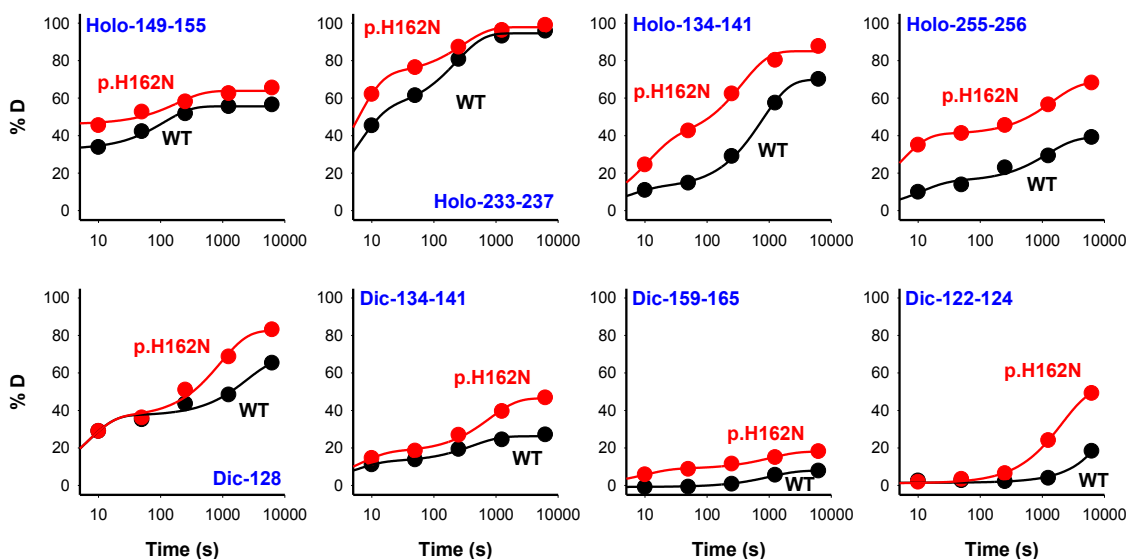
Supplementary references

1 Weis DD, Wales TE, Engen JR, Hotchko M & ten Eyck LF (2006) Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. *J Am Soc Mass Spectrom* **17**, 1498–1509.

2 Anoz-Carbonell E, Timson DJ, Pey AL & Medina M (2020) The catalytic cycle of the antioxidant and cancer-associated human NQO1 enzyme: Hydride transfer, conformational dynamics and functional cooperativity. *Antioxidants* **9**, 1–22.

3 Pacheco-Garcia JL, Loginov DS, Anoz-Carbonell E, Vankova P, Palomino-Morales R, Salido E, Man P, Medina M, Naganathan AN & Pey AL (2022) Allosteric Communication in the Multifunctional and Redox NQO1 Protein Studied by Cavity-Making Mutations. *Antioxidants* **11**, 1110.

4 Pacheco-Garcia JL, Anoz-Carbonell E, Vankova P, Kannan A, Palomino-Morales R, Mesa-Torres N, Salido E, Man P, Medina M, Naganathan AN & Pey AL (2021) Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue. *Redox Biol* **46**, 102112.

5 Faig M, Bianchet MA, Talalay P, Chen S, Winski S, Ross D & Amzel LM (2000) Structures of recombinant human and mouse NAD(P)H:quinone oxidoreductases:

species comparison and structural changes with substrate binding and release. *Proc Natl Acad Sci U S A* **97**, 3177–82.

6 Kavan D & Man P (2011) MSTools—Web based application for visualization and presentation of HXMS data. *Int J Mass Spectrom* **302**, 53–58.

7 Vankova P, Salido E, Timson DJ, Man P & Pey AL (2019) A Dynamic Core in Human NQO1 Controls the Functional and Stability Effects of Ligand Binding and Their Communication across the Enzyme Dimer. *Biomolecules* **9**, 728.

8 Asher G, Dym O, Tsvetkov P, Adler J & Shaul Y (2006) The crystal structure of NAD(P)H quinone oxidoreductase 1 in complex with its potent inhibitor dicoumarol. *Biochemistry* **45**, 6372–8.

PUBLICACIÓN 3

Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue.

Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue

Juan Luis Pacheco-Garcia [a,1], Ernesto Anoz-Carbonell [b,1], Pavla Vankova [c,d,1], Adithi Kannan [e], Rogelio Palomino-Morales [f], Noel Mesa-Torres [a], Eduardo Salido [g], Petr Man [c], Milagros Medina [b], Athi N. Naganathan [e], Angel L. Pey [h,*]

[a] Departamento de Química Física, Universidad de Granada, Av. Fuentenueva s/n, 18071, Granada, Spain
[b] Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias, Instituto de Biocomputación y Física de Sistemas Complejos (GBsC-CSIC and Joint Unit), Universidad de Zaragoza, 50009, Zaragoza, Spain
[c] Institute of Microbiology, Academy of Sciences of the Czech Republic, Videnska 1083, Prague 4, 142 20, Czech Republic
[d] Department of Biochemistry, Faculty of Science, Charles University, Hlavova 2030/8, Prague 2, 128 43, Czech Republic
[e] Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras (IITM), Chennai, 600036, India
[f] Departamento de Bioquímica y Biología Molecular I, Facultad de Ciencias y Centro de Investigaciones Biomédicas (CIBM), Universidad de Granada, Granada, Spain
[g] Center for Rare Diseases (CIBERER), Hospital Universitario de Canarias, Universidad de la Laguna, 38320, Tenerife, Spain
[h] Departamento de Química Física, Unidad de Excelencia en Química Aplicada a Biomedicina y Medioambiente e Instituto de Biotecnología, Universidad de Granada, Av. Fuentenueva s/n, 18071, Granada, Spain

ARTICLE INFO

ABSTRACT

The multifunctional nature of human flavoproteins is critically linked to their ability to populate multiple conformational states. Ligand binding, post-translational modifications and disease-associated mutations can reshape this functional landscape, although the structure-function relationships of these effects are not well understood. Herein, we characterized the structural and functional consequences of two mutations (the cancer-associated P187S and the phosphomimetic S82D) on different ligation states which are relevant to flavin binding, intracellular stability and catalysis of the disease-associated NQO1 flavoprotein. We found that these mutations affected the stability locally and their effects propagated differently through the protein structure depending both on the nature of the mutation and the ligand bound, showing directional preference from the mutated site and leading to specific phenotypic manifestations in different functional traits (FAD binding, catalysis and inhibition, intracellular stability and pharmacological response to ligands). Our study thus supports that pleitropic effects of disease-causing mutations and phosphorylation events on human flavoproteins may be caused by long-range structural propagation of stability effects to different functional sites that depend on the ligation-state and site-specific perturbations. Our approach can be of general application to investigate these pleiotropic effects at the flavoproteome scale in the absence of high-resolution structural models.

1. Introduction

The human flavoproteome consist of about a hundred of different proteins [1,2] that can be regarded as *multifunctional* proteins. In a wide sense, we refer to *function* as the ability to carry out enzymatic reactions, to mediate biomacromolecular interactions and to be transported for operating in different subcellular locations, among others (Table S1). As a principle, this multifunctionality must be intrinsically linked to the ability of human flavoproteins to populate different conformational states (i.e. to their conformational lansdscape). Regarding biochemical reactions, the capacity of human flavoproteins to catalyze (mostly redox) reactions requires the presence of a bound flavin cofactor (primarily FAD and FMN) as well as the binding of other substrates and cofactors. In this sense, the diversity of biochemical reactions catalyzed by flavoproteins is remarkable (Table S1). Biomolecular association

---

* Corresponding author.
  E-mail address: angelpey@ugr.es (A.L. Pey).
[1] These authors contributed equally to this work.

Abbreviations

A	Arrhenius frequency factor
ASA	accessible surface area
CTD	C-terminal domain
ΔC_p	apparent binding heat capacity
% D	percentage of deuterium incorporated
Dic	dicoumarol
DBS	dicoumarol binding site
DCPIP	2,6-dichlorophenolindophenol
DAD	donor-acceptor-distance
DT	deuterium-transfer
E_a	activation energy
FAD	flavin-adenine dinucleotide
FBS	FAD binding site
FMN	flavin mononucleotide
$\Delta\Delta G_{prot}$	change in thermodynamic stability towards proteolysis
ΔH	apparent binding enthalpy
HDX	hydrogen/deuterium exchange
HT	hydride-transfer
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
K_d	dissociation constant
k_{HT}	limiting hydride-transfer rate constant
KIE	kinetic isotope effect
k_{obs}	observed rate constant
k_{prot}	second-order proteolysis rate constant
LDA model	ligand-dependent anisotropic model
LDI model	ligand-dependent isotropic model
MCI model	mutation-centered isotropic model
MMI	monomer:monomer interface
NADH	Nicotinamide adenine dinucleotide
NADD	deuterated nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide 2′-phosphate
N_{conf}	residues folded upon ligand binding
NQO1	NADP(H):quinone oxidoreductase 1
$NQO1_{apo}$	NQO1 with no bound ligand
$NQO1_{holo}$	NQO1 with bound FAD
$NQO1_{dic}$	NQO1 with bound FAD and dicoumarol
NTD	N-terminal domain
PDB	protein data bank
SDS-PAGE	polyacrylamide gel electrophoresis in the presence of SDS
Rib	riboflavin
WT	wild-type

with other proteins and nucleic acids has potential regulatory effects on different cellular processes (Table S1). In addition, a large fraction of human flavoproteins operate in the mitochondria [1] and therefore, most of them must be synthesized in the cytosol and imported (i.e. through *unfolding*) to this organelle [3]. Nonetheless, human flavoproteins display a substantial diversity of subcellular locations (about two-thirds show more than one cellular location; Table S1), and upon cytosolic synthesis must fulfill different requirements for translocation through dedicated macromolecular machineries to different subcellular locations such as the nucleus, peroxisomes or endoplasmic reticulum among others [4–6]. Flavin binding is also required to modulate their enzyme activity, intracellular stability and protein:protein interactions [7,8]. In addition, the redox status of the flavin is not only required for their enzymatic function, but can also modulate others features such as subcellular location [9]. Interestingly, although many high-resolution structures have been described for human flavoproteins with different bound ligands (i.e. ligation states), there is virtually no such information on their flavin-free apo-state (Table S1). This is an important issue, since the structure and energetics of the apo-state can contribute to different functions, such as flavin binding affinity, intracellular stability, bio-macromolecular interactions and subcellular transport.

Due to their multiple functions, a large fraction of human flavoproteins has been associated with inherited diseases through loss-of-function mutations [2]. In addition, a vast number of potential post-translational modifications (mostly phosphorylation events) have been reported by high-throughput techniques, although the site-specific functional consequences of these events are largely unknown [10]. Considering the evidence on the multifunctional nature of human flavoproteins, it is plausible that missense mutations and post-translational modifications will affect to different extents certain of these multiple functional features [11–14]. Indeed, structure-based stability analyses support that disease-associated mutations and phosphorylation may share a similar destabilizing effect on the structure of human flavoproteins, particularly when such events affect residues buried within the protein interior (Figure S1 and Table S2).

To study the *allosteric* interplay between ligand binding, post-translational modifications and disease-associated mutational effects, we have investigated here the multifunctional human protein NAD(P)H: quinone oxidoreductase 1 (NQO1) [15]. Alterations in NQO1

functionality are associated with a variety of diseases, including cancer, Parkinson's and Alzheimer diseases, diabetes, multiple sclerosis, schizophrenia, metabolic syndrome and benzene toxicity [15,16]. NQO1 is a flavoprotein that catalyzes the two-electron reduction of a wide variety of natural and synthetic quinones (in the presence of NAD(P)H as coenzyme) and interacts with and stabilizes transcription factors critical for cancer development (such as p53, p73α and HIF-1α) [15–18]. NQO1 is primarily cytosolic, although it can also be found in mitochondria, nucleus and perinuclear regions (Table S1) [9]. NQO1 forms functional homodimers with a molecular size of 62 kDa and containing two functional domains within each monomer (Fig. 1A) [19–21]. The N-terminal domain (NTD, residues 1–225) is capable of forming stable dimers that bind FAD, although the presence of the C-terminal domain (CTD) is required to complete the active site [21,22]. NQO1 is a remarkable example of human flavoprotein that exists in different conformational (i.e. ligation-dependent) states with very different functionalities [7–9,12,15,16,18,23]. In the absence of FAD (the $NQO1_{apo}$ state) the WT protein shows a remarkably low conformational stability, with a minimal stable core that holds the protein dimer and the FAD and NAD(P)H binding sites populating non-competent binding conformations [23]. The $NQO1_{apo}$ state is rapidly degraded inside the cell by ubiquitin-dependent proteasomal activity and display altered interactomic patterns [7,8]. Consequently, FAD binding (forming the $NQO1_{holo}$ state) triggers a population-shift in the conformational ensemble (i.e. a conformational change) leading to a precatalytic competent state in which the FAD and substrate/NADP(H) binding sites are stabilized, providing a boost in intracellular stability through stabilization of the CTD [7,12,15]. Binding of NAD(P)H is the rate-limiting step in the reductive half-reaction of NQO1 [22,24], and this state with the flavin reduced is associated with binding to microtubules and α-tubulin acetylation [9], and more generally, it promotes the interaction of NQO1 with protein partners [16,18]. Binding of dicoumarol (Dic, forming the state $NQO1_{dic}$), a competitive inhibitor of NAD(P)H that may resemble a transition state analogue in both reductive and oxidative half-reactions, also lead to significant changes in the local stability in the protein ensemble without causing gross conformational changes [23–25]. Binding of this inhibitor seems to prevent the interaction of NQO1 with other proteins [18].

Several single amino acid substitutions in NQO1 have been shown to
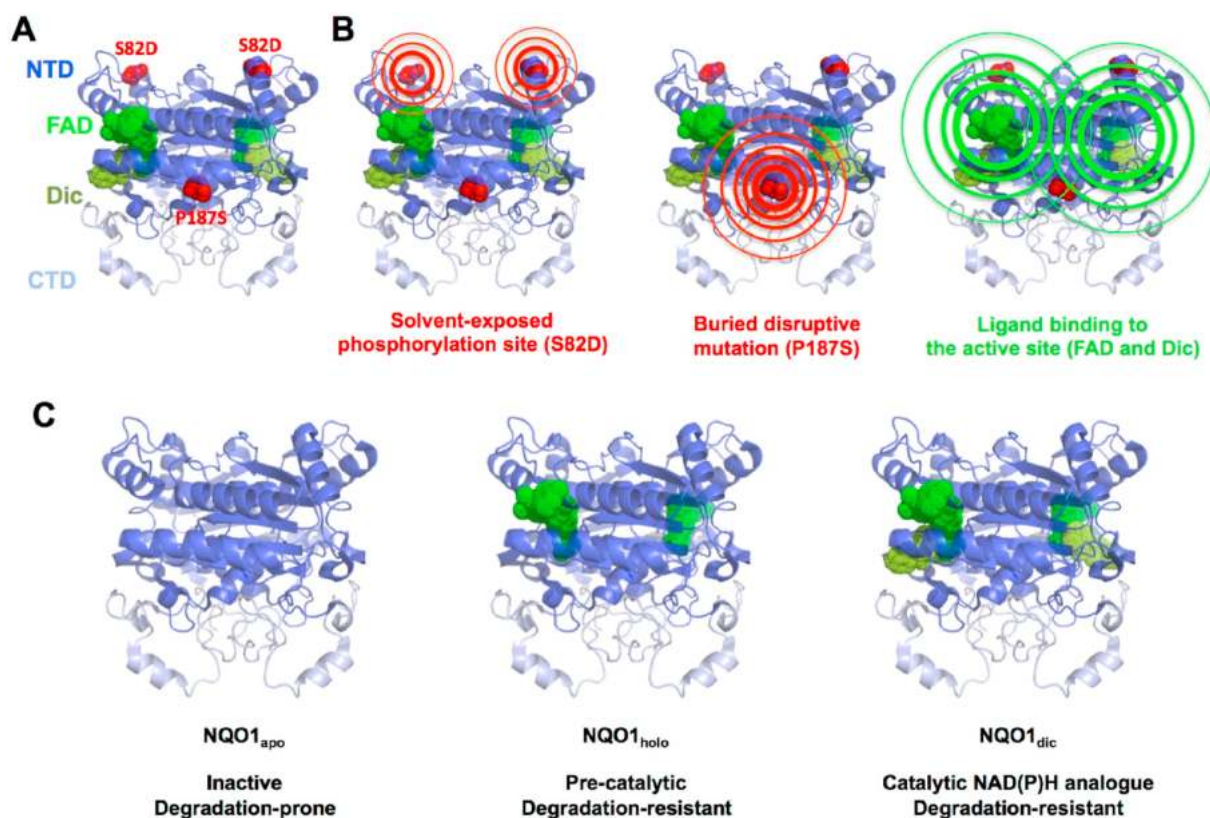
Fig. 1. Structural features of the NQO1 protein dimer and the sites perturbed in this work. A) Crystallographic model of the NQO1 dimer (PDB code 2F1O [25]) showing its two-domain structure (NTD, N-terminal domain and CTD, C-terminal domain) and the location of the S82D and P187S sites as well as FAD and Dic binding sites. Panel B shows the potential propagation of the *allosteric perturbation events* studied here (S82D, P187S, FAD and Dic binding). C) Relevant ligation states investigated in this work show different catalytic and stability features in WT NQO1.

affect different functional sites (and features) of NQO1, sometimes located far from the mutated site [8,10–14,26,27]. The common cancer-associated polymorphism Pro187Ser/P187S [16,28] is known to severely affect FAD binding and to reduce the intracellular stability of the protein, affecting also the interactomics of the protein [7,8,12,15,21, 22,26]. This polymorphism affects a residue that is fully buried in the structure (% ASA strictly 0; using PDB code 2F1O and the software GetArea; http://curie.utmb.edu/getarea.html; average of eight monomers) and close to the dimer interface (Fig. 1B). Therefore, we may expect that its local structural and energetic effects will efficiently propagate across the entire structure of the NQO1 dimer (likely amplified through some *constructive interference* between the two mutated sites in the homodimer) (Fig. 1). The phosphomimetic mutation Ser82-Asp/S82D, affects a partially exposed site (% ASA of $16 \pm 8\%$; using the same procedure) (Fig. 1A) and seems to cause more local effects, causing a 30-fold decrease in binding affinity for FAD [10](Fig. 1). However, this decrease in affinity and/or alterations in structural stability may lead to a concomitant decrease in intracellular stability [10]. Here, we have used these two mutations to unravel how different type of mutations may lead to different phenotypic consequence in a multifunctional flavoprotein, and how these effects are dependent on bound ligands (Fig. 1). To this end, we carried out extensive structure-function analysis using a wide array of computational, biochemical, biophysical, cell and structural biology approaches.

2. Materials and methods

2.1. Protein expression and purification

Expression and purification of NQO1 WT, S82D and P187S were carried out as described [23,24]. *E. coli* BL21 (DE3) cells were transformed with the pET46 Ek/LIC vector containing the cDNA of NQO1 WT, S82D or P187S [10,11]. A preculture of single clones was grown in 400 mL LBA medium (LB containing 0.1 mg·ml^{-1} ampicillin) for 16–20 h at 37 °C. Cells were harvested and resuspended in 4.8 L of LBA and grown at 37 °C for 3 h and NQO1 expression was induced by the addition of 0.5 mM IPTG (isopropyl β-ᴅ-1-thiogalactopyranoside) and lasted for 6 h at 25 °C. Induced cells were harvested by centrifugation and frozen overnight at −80 °C. NQO1 proteins were purified using immobilized nickel affinity chromatography columns (GE Healthcare) and size-exclusion chromatography as described [24]. Isolated dimeric fractions of NQO1 variants were buffer exchanged to K-HEPES buffer 50 mM pH 7.4 using PD-10 columns (GE Healthcare). The UV–visible spectra of purified NQO1 proteins were registered in a HP8453 UV–Visible spectrophotometer (Agilent Technologies, Waldbronn, Germany) and used to quantify the content of FAD as described [24]. Prior to pre-steady state kinetic analyses, proteins were incubated with 1 mM FAD and the excess of FAD removed using PD-10 columns, obtaining a saturation fraction (FAD:NQO1 monomer) higher than 95% based on UV–visible spectra. Apo-proteins were obtained by treatment with 2 M urea and 2 M KBr as described [23], obtaining samples with less than 2% saturation fraction of FAD based on UV–visible spectra. Samples were stored at −80 °C upon flash freezing in liquid N$_2$. Protein purity and integrity was checked by polyacrylamide gel electrophoresis in the presence of sodium dodecylsulphate (SDS-PAGE) and intact mass analysis using direct infusion on 15T ESI-FT-ICR MS as described previously [23].

2.2. Hydrogen-deuterium exchange (HDX) monitored by mass spectrometry

Backbone amide hydrogen/deuterium exchange (HDX) was

monitored for NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$ forms of NQO1 using mass spectrometry. Data were collected for S82D and P187S variants whereas data for WT are those previously published [23]. It should be noted that all the data were acquired within one set of experiments which ensures their full comparability and diminishes possible variations eventually arising due to different experimental settings. Therefore all procedures including sample preparation, sample measurement and data processing were identical and performed as described previously [23]. To generate NQO1$_{holo}$ and NQO1$_{dic}$ states, 20 μM purified NQO1 protein was preincubated with 10-molar excess of FAD (NQO1$_{holo}$) for 5 min. Generation of NQO1$_{dic}$ required an additional 5 min incubation with 10-molar excess of dicoumarol. Isotope exchange reaction was started by 10-fold dilution of the protein solution into D$_2$O-based 50 mM K-HEPES, pD 7.4, 1 mM TCEP (tris(2-carboxyethyl)phosphine). The exchange was followed for 10 s, 30 s, 2 min, 5 min, 20 min, 1 h and 3 h. The reaction was quenched by the addition of 0.5 M Glycine-HCl, pH 2.3 in a 1:1 ratio and flash frozen in liquid N$_2$. The 10 s, 5 min and 3 h aliquots were replicated three times. LC-MS analysis was started by rapid thawing of the sample and injection onto the custom-made nepenthesin-2 (Nep2) and pepsin columns coupled in tandem (each of them having a bed volume 66 μL). Generated peptides were trapped and desalted on a VanGuard Pre-column (ACQUITY UPLC BEH C18, 130 Å, 1.7 μm, 2.1 mm × 5 mm, Waters, Milford, MA, USA). The digestion and desalting (total time 3 min) were driven by 0.4% formic acid (FA) in water pumped by 1260 Infinity II Quaternary pump (Agilent Technologies, Waldbronn, Germany) at a flow rate of 200 μL min^{-1}. Desalted peptides were then separated on an analytical column (ACQUITY UPLC BEH C18, 130 Å, 1.7 μm, 1 mm × 100 mm, Waters, Milford, MA, USA) using linear gradient (5%–45% B in 7 min) followed by a quick step to 99% B lasting 5 min, where the solvent A was 0.1% FA/2% acetonitrile (ACN) in water, B was 0.1% FA/98% ACN in water. The gradient was delivered by the 1290 Infinity II LC System (Agilent Technologies, Waldbronn, Germany) at a flow rate 40 μL ·min^{-1}. The analytical column was coupled to an ESI source of 15T FT-ICR mass spectrometer (solarix XR, Bruker Daltonics, Bremen, Germany) operating in broad-band MS mode. Acquired data were exported using DataAnalysis v. 5.0 (Bruker Daltonics, Bremen, Germany), processed by in-house designed software program Deutex (unpublished) and handled as described previously [23]. Peptide identification was performed using separate LC-MS/MS analyses followed by MASCOT database searching (version 2.4, Matrix Science, London, United Kingdom) against a custom-built database containing sequences of the proteases and the NQO1 variants under investigation. Fully deuterated controls were prepared as described previously [23] and correction for back-exchange was applied according to Ref. [29]. HDX-MS related data, graphs and images in the Supplementary material include sequence coverage map (plotted with MSTools - http://peterslab.org/MSTools/), all deuterium uptake plots for apo, holo and dicoumarol states of wild-type, S82D and P187S NQO1 variants, selected examples of isotopic profiles (all time points plus non- and fully-deuterated samples) for NQO1$_{apo}$ and tables (Tables S3-S4) summarizing digestion metrics and a full HDX dataset (averaged values - %D - after back-exchange correction). HDX results are presented in two different ways: i) %D at a given time, that just reflects the % of total H exchanged for a given protein segment; ii) Δ%D$_{av}$, a single metric that compares the kinetic behavior of a given segment between two variants in a given ligation state (i.e. S82D$_{apo}$ vs. WT$_{apo}$) as the average of the three time points that show maximal difference in % D along a time course. This metric provides a simple way to identify those segments with different stabilities between two protein variants in a given ligation state as well as a semiquantitative ranking of these differences [23].

2.3. Structure-based models for the propagation of mutational and ligand binding effects from HDX

The native conformational landscape of NQO1$_{apo}$ was generated from the Ising-like and native-centric WSME (Wako-Saitô-Muñoz-Eaton) model [30,31]. This employs a binary treatment of folding wherein folded- and unfolded-like residues are assigned the binary variables *1* and *0*, respectively. We used a version that allows for only a single stretch of folded residues (single sequence approximation, SSA), two stretches folded residues (double sequence approximation, DSA) and DSA with interaction across the folded islands despite allowing for the intervening residues to be unfolded (DSAw/L), to generate the conformational landscape of NQO1. Moreover, instead of considering every residue as a folding unit that results in >400 million microstates, we employ an approximation where consecutive residues are treated as blocks (in this case, a block length of 5 is considered) [32]. This reduces the total number of microstates to 1,443,586 making it computationally less intensive. The model includes contributions from van der Waals interactions (identified with a 6 Å cutoff excluding nearest neighbors), electrostatics (via a Debye-Hückel formalism), simplified solvation and sequence-structure dependent conformational entropy [33]. We employ a mean-field van der Waals interaction energy of −49 J mol^{-1}, entropic penalty for fixing a residue in the native conformation of −14 J mol^{-1} K^{-1} per residue, heat capacity change on forming a native contact of −0.36 J mol^{-1} K^{-1} per native contact and an excess conformational entropy (for disordered residues and glycine) of −6.1 J mol^{-1} K^{-1} per residue. Free energy surfaces were generated by accumulating partial partition functions with fixed number of folded blocks or their combination.

First, we applied a *mutation-centered-isotropic model* (MCI model). Briefly, the effects of mutations on the local stability were derived from analysis of HDX data by calculating Δ%D$_{av}$ values for different segments between a mutant and the WT protein for each ligation state (NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$). Then, the distance between the mutated site (residue 82 or 187) and any given protein segment derived from HDX data, was calculated as the minimal distance (d) between the C$_\alpha$ atoms in the center of the segments corresponding to the mutated site and the rest of (104) segments analyzed experimentally. Therefore, the reference system for these calculations is the same for all ligation states (i.e. the mutated site). For these calculations, we used the structure of dimeric NQO1 in the NQO1$_{dic}$ state (PDB: 2F1O). The results (Δ%D$_{av}$ vs. distance, d) were eventually clustered in four groups according to their distance to the mutated site: < 10 Å, 10–20 Å, 20–30 Å and >30 Å. Data pairs (Δ%D$_{av}$ vs. distance) corresponding to each group were averaged (±s.d.) and fitted to the following equation (Equation (1)):

$$A(d) = A_0 \cdot e^{\left(-\frac{d}{d_c}\right)} \tag{Eq 1}$$

where A(d) is the Δ%D$_{av}$ at a given d (for each of the four groups), A$_0$ is the Δ%D$_{av}$ at d = 0 (i.e. at the mutated site) and d$_c$ is a characteristic distance value for the propagation of stability effects.

In a second approach, we used reference systems that changed according to the ligation state, while keeping the dependent variable (i.e. Δ%D$_{av}$) as it was described in the MCI model. Briefly, the reference systems were: i) NQO1$_{apo}$.- the reference system was the Cα of the mutated site; ii) NQO1$_{holo}$ and NQO1$_{dic}$.- the reference system was the center of mass of the ligand (i.e. FAD and Dic, respectively). In all cases, the centroid of the Cα atoms of the experimentally determined segments were used to calculate distances from the corresponding reference point. This approach was implemented in two different ways. In the first one, the pairs of Δ%D$_{av}$ and d values were calculated and clustered in four groups as described in the MCI model. This way was termed as *ligand-dependent-isotropic model* (LDI model). In a second one, we applied a zone method in which the 3D space was divided into 6 zones (angle of 60° each) using 3 planes at an angle of 120° to each. Zone 1 contains the Cα – Cβ bond of the mutated residue in case of the NQO1$_{apo}$ and the imaginary line connecting the center of masses of the ligand and its interacting residues in case of NQO1$_{holo}$ and NQO1$_{dic}$. Zones 2 and 6 are adjacent to Zone 1 and together account for residues in the direction of

perturbation. Zones 4, 5 and 3 fall exactly opposite to zones 1, 2 and 6 respectively and are structurally opposite to the direction of perturbation. Residues inside each zone were clustered into 4 groups as abovementioned and averaged (±s.d.). We refer to this method as *ligand-dependent-anisotropic model* (LDA model).

2.4. Pre-steady state enzyme kinetic analysis

Fast hydride-(HT) or deuteride-(DT) transfer reactions were carried out under anaerobic conditions using a stopped-flow spectrophotometer (SX.18 MV, Applied Physics Ltd.) interfaced with a photodiode array detector, essentially as described [24]. Briefly, the reductive half-reaction was measured by mixing NQO1$_{holo}$ variants (7.5 μM) with either NADH ranging from 7.5 to 100 μM or with NADPH or 4R-^{2}H-NADH (NADD) at stoichiometric concentrations (7.5 μM). To study the oxidative half-reaction, 7.5 μM NQO1$_{hq}$ (formed by reaction of NQO1$_{holo}$ with stoichiometric amounts of NADH) was mixed with 7.5 μM of 2,6-Dichlorophenolindophenol (DCPIP). Reactions were performed in 20 mM HEPES-KOH, pH 7.4. Multiple wavelength absorption data in the flavin absorption region were collected and processed as described [24]. Time-dependent spectral deconvolution was performed by global analysis and numerical integration methods using previously described procedures [24]. Basically, this deconvolution procedure was carried out considering sequential and irreversible steps in the context of two (A→B→C) or three (A→B→C→D) step mechanisms, where A-D are spectral species (not necessarily a given *state*), and allowed to determine observed rate constants (k_{obs}) for these steps as well as spectroscopic properties of A and D are initial and final, but not intermediate states A to D species. According to a recent study, catalytically relevant processes involved steps A→B→C [24]. Hyperbolic dependences of k_{obs} vs. NADH concentrations were fitted using equation (2):

$$k_{obs} = \frac{k_{HT} \cdot [NADH]}{K_d^{NADH} + [NADH]}$$ (Eq 2)

where k_{HT} is the limiting rate constant for HT and K_d^{NADH} is the equilibrium dissociation constant to a given active site.

For estimation of primary kinetic isotopic effects (KIEs) in the HT process [34], HT or DT k_{obs} from NADH/D to NQO1$_{holo}$ were evaluated at different temperatures in samples containing equimolecular mixtures of protein and coenzyme (7.5 μM of each component) using NADH and [4R-^{2}H]-NADD. KIEs were thus determined as the ratio of the k_{obs} values using NADH and NADD at a given temperature. Activation parameters (frequency factor, A, and the activation energy, E$_a$) were determined using the Arrhenius equation as described [24].

2.5. Proteolysis by trypsin

Stock solutions of trypsin from bovine pancreas (T1426, from Sigma-Aldrich) were prepared at 80–100 μM in 50 mM HEPES-KOH, pH 7.4 and stored at −80 °C in small aliquots upon freezing in liquid nitrogen. Trypsin concentration was determined using $\varepsilon_{280} = 35100$ M^{-1}·cm^{-1} (according to the manufacturer's instructions). NQO1 samples were diluted to a final concentration of 15 μM in protein subunit in 50 mM HEPES-KOH, pH 7.4. For NQO1$_{holo}$ samples, a final concentration of FAD 100 μM was added, while in NQO1$_{dic}$, final concentrations of FAD and dicoumarol were 100 μM. The reaction mixtures were incubated for 10 min at 25 °C prior to addition of a concentrated trypsin solution (to a final concentration ranging from 5 nM to 2 μM in protease). Aliquots were withdrawn at different time points, mixed with 1/10th volume of 10 mM Phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich) in ethanol and then with 1 volume of 4 x Laemmli's buffer and immediately denatured for 5 min at 95 °C. Samples were resolved in 12% acrylamide SDS-PAGE gels, stained with Coomassie® Brilliant Blue R250 (Sigma-Aldrich). Gels were analyzed using ImageJ (https://imagej.nih.gov/ij/). The intensity of the uncleaved (full-length) protein (I) vs. time was used

to determine the observed (apparent) rate constant k_{obs} from fittings using the following exponential function:

$$I = I_0 \cdot exp^{-k_{obs} \cdot t}.$$ (Eq 3)

From the linear dependence of k_{obs} on trypsin concentration, the second-order rate constants k_{prot} were obtained. Thus, the change in thermodynamic stability of the CTD due to a mutation or ligand binding (state i), $\Delta\Delta G_{prot}$, can be calculated using a given reference state (for instance, WT NQO1$_{apo}$) using equation 4:

$$\Delta\Delta G_{prot} = -R \cdot T \cdot ln\left(\frac{k_{prot(i)}}{k_{prot(ref)}}\right).$$ (Eq 4)

Therefore, a negative/positive value of $\Delta\Delta G_{prot}$ reflects that the i state has lower/higher CTD thermodynamic stability than the reference state.

2.6. Thermodynamics of dic binding by isothermal titration calorimetry (ITC)

ITC experiments were carried out using an ITC$_{200}$ microcalorimeter (Malvern). NQO1 samples were loaded in the cell at a concentration of 15–17 μM (in monomer) as holo-proteins (+100 μM of FAD). NQO1$_{holo}$ proteins were titrated using Dic (150–180 μM, plus 100 μM FAD) by performing an initial injection (0.5–1 μL) followed by 18–23 injections of 1.5–2 μL each, spaced by 150–210 s. Experiments were carried at 25 °C (at least three independent titrations), 20 °C and 15 °C (a single titration for these temperatures). Data analyses were performed as previously described [13,21]. Briefly, binding isotherms were analyzed considering a single type of independent binding sites, and dilution heats were considered as a fitting parameter. This yielded all apparent binding thermodynamic parameters (K$_a$, ΔG, ΔH, ΔS) whereas the apparent change in heat capacity (ΔC_p) was determined from the linear dependence of the apparent ΔH on temperature.

Apparent ΔH and ΔC_p were used to estimate the magnitude of the conformational change triggered by Dic binding as recently described [13]. Briefly, it was considered that these two variables contain a contribution from intrinsic binding (i.e. lock-and-key mechanism) ΔH_{int} and $\Delta C_{p,int}$, plus a contribution arising from the conformational change, ΔH_{conf} and $\Delta C_{p,conf}$ [35,36]. Intrinsic binding parameters were determined from changes in apolar and polar surface upon Dic binding (−969 Å^2 and -621 Å^2, respectively) estimated from high-resolution X-ray structural models of NQO1 with Dic and using well-known structure-energetics correlations [13,36,37]. These correlations provided values of $\Delta H_{int} = -5.8$ kcal mol^{-1} and $\Delta C_{p,int} = -0.14$ kcal mol^{-1}. Thus, the difference between experimental (apparent) ΔH and ΔC_p and those from intrinsic binding provided the contribution from the conformational change [13,35,36,38]. Previous studies have supported that for Dic binding to NQO1, ΔH_{conf} and $\Delta C_{p,conf}$ mainly arise from the conformational transition (i.e. unfolded/partially folded to folded) induced by Dic binding on the CTD and thus, these can be used using well-known structure-energetics for protein folding thermodynamics [13,21,39] to estimate in two different manners the apparent number of residues folded upon binding (N$_{conf}$) that provides an estimate of the magnitude of the conformational change associated with Dic binding (Equations (5) and (6)):

$$\Delta H_{conf} = 0.215 \cdot N_{conf}$$ (Eq 5)

$$\Delta C_{p,conf} = 0.0138 \cdot N_{conf}$$ (Eq 6)

Where ΔH_{conf} and $\Delta C_{p,conf}$ are expressed in kcal·mol^{-1} and kcal·mol^{-1}·K^{-1}, respectively. These values of N$_{conf}$ are expressed per NQO1 monomer.

2.7. Expression analysis in eukaryotic cells

The mutations P187S and S82D were introduced by site-directed mutagenesis on the wild-type (WT) NQO1 cDNA cloned into the pCI-NEO plasmid by GenScript (Leiden, The Netherlands). Mutagenesis was confirmed by sequencing the entire cDNA. The resulting construct produces the full-length WT protein containing three extra amino acids (MLA) prior to Met1 but no additional tags.

HAP1 NQO1 knockout cells (HAP-1 NQO1-KO; Horizon, Waterbeach, UK; ref. K9HZGHC006138C004) were cultured in Iscove's modified Dulbecco's medium (IMDM, Lonza, Barcelona, Spain) supplemented with 10% fetal bovine serum (HyClone, GE Healthcare, Barcelona, Spain), 100 $U \cdot mL^{-1}$ penicillin and 100 $\mu g \cdot mL^{-1}$ streptomycin (Sigma Aldrich, Madrid, Spain) and cultured at 37 °C in a humidified incubator with 5% CO_2. Cells were transfected using Lipofectamine LTX with Plus Reagent (Thermo Fisher Scientific, Madrid, Spain) and selected using 1 $mg \cdot mL^{-1}$ of G418 (Sigma Aldrich). Cells were treated with riboflavin (100 μM) or riboflavin + Dic (both at 100 μM and from Sigma Aldrich) for 24 h, using untreated cells under the same conditions as controls. For proteasomal inhibition studies, MG-132 (Calbiochem, Merck, Madrid, Spain) was added to the medium (at 10 $\mu g \cdot mL^{-1}$) for 4 h at 37 °C. After treatment, cells were scrapped and lysed in RIPA buffer (50 mM Tris-HCl, 150 mM NaCl, 0.1% Triton X-100, 0.1% sodium dodecyl sulphate, 1 mM sodium orthovanadate, 1 mM NaF pH 8) with protease inhibitors (COMPLETE, from Roche, Spain). Soluble extracts were denatured using Laemmli's buffer and resolved by SDS-PAGE and transferred to polyvinylidene difluoride membranes (GE Healthcare). Immunoblotting was carried out using primary monoclonal antibodies anti-NQO1 (sc-393736, Santa Cruz Biotechnology) and anti-β-actin (sc-47778, Santa Cruz Biotechnology) from mouse at 1:200 and 1:5000 dilutions, respectively. As a secondary antibody, we used chicken anti-mouse IgG-HRP (sc-516102, Santa Cruz Biotechnology) at 1:2000 dilution. Samples were visualized using luminol-based enhanced chemiluminescence (from BioRad Laboratories). All experiments included untreated cells transfected with WT NQO1 as reference.

3. Results and discussion

3.1. The mutants S82D and P187S increase the population of partially folded conformations in the native state ensemble of NQO1$_{apo}$

Unlike NQO1$_{holo}$ and NQO1$_{dic}$, no high resolution structural information is available for NQO1$_{apo}$ (likely due to its high intrinsic flexibility) [12]. Importantly, local stability analyses by hydrogen/deuterium exchange (HDX) kinetics of NQO1$_{apo}$ showed remarkable heterogeneity across the structure. Different regions of the WT protein showed widely different extent and kinetics of deuterium incorporation (%D) in a 3 h experiment [23]. This behavior suggests that NQO1$_{apo}$ populates different conformational substates under native conditions with very different local stabilities. Presumably, the population of these partially folded states is connected with certain functional features such as FAD binding, proteasomal degradation of the protein, or biomacromolecular interactions [7,8,12]. To further characterize this conformational heterogeneity, we carried out a statistical mechanical analysis with the Ising-like and native-centric WSME (Wako-Saitô-Muñoz-Eaton) model, in which the conformational ensemble of NQO1$_{apo}$ WT is generated from ~1.4 million microstates. The resulting conformational landscape consistently reproduces the heterogeneity of states expected from HDX experiments (Fig. 2A and S7) [23]. In addition to the Native (N) and Unfolded (U) ensembles, at least two partially folded ensembles (I$_1$ and I$_2$) are significantly populated (Fig. 2A). The I$_1$ intermediate contains residues 1–220 folded, whereas the I$_2$ intermediate has residues 90–220 folded.

What are the structural-functional features of these intermediates? The intermediate I$_1$ is predicted to exhibit a conformation in which the stable core of NQO1$_{apo}$ is folded including those residues critical for FAD

binding (i.e. it would have *high affinity* for FAD), whereas the intermediate I$_2$ contains 74% of this core in a folded conformation but only a few of the residues key for FAD binding (Fig. 2A–B). The population of conformational states similar to the I$_1$ intermediate in NQO1$_{apo}$ is further supported by previous mutagenesis studies in which the CTD (residues 225–274) was excised (i.e. Δ50-NQO1). Δ50-NQO1 stably forms dimers and retains significant binding affinity for FAD (with 8-fold lower than the full-length apo-protein) but reduces NQO1 activity by two-orders of magnitude [21,22,40]. Additionally, the population of these intermediates can be critical in determining efficient ubiquitination and degradation of NQO1 inside cells, which is mainly dictated by the presence of an unfolded CTD [7,12,22,41].

The NQO1$_{apo}$ state shows a marginal stability and the dimeric state is held by a stable core that undergoes minimal HDX after 3 h (Fig. 2B and [23]). Importantly for our study, the mutations S82D and P187S affected this core, and their effects were quantitatively different (Fig. 2B). S82D mildly destabilized segments 21–23 (helix α1) and 166–181 (helix α4) whereas P187S largely destabilized segments 166–178 (helix α4) and moderately destabilized segments 21–23 (helix α1) and 179–185 (sheet β5). Hence, the mutation P187S causes a stronger structural destabilization that also propagates more extensively through the stable core of NQO1$_{apo}$ (Fig. 2B). In addition, HDX kinetics and statistical mechanical analyses support that both P187S and S82D may increase the population of the I$_1$ intermediate, and to a lesser extent of I$_2$ intermediate, and these population-shift should be more pronounced in the case of the P187S mutant (Fig. 2 and S7-8).

3.2. Structural perturbations in the NQO1$_{holo}$ and NQO1$_{apo}$ states largely differ between the S82D and P187S mutants and explain their reduced affinity for FAD

FAD binding to NQO1 WT causes remarkable changes in protein structural stability to form a pre-catalytic and intracellularly stable NQO1$_{holo}$ state [7,12,23,42–44]. Perturbations in the structural stability of the NQO1$_{holo}$ and NQO1$_{apo}$ states by P187S and S82D are likely associated with their altered function and stability [7,10,12,21,22,42, 44]. Particularly, alterations of the flavin binding site (FBS) stability are associated with the 10–40 fold lower affinity for FAD observed for the S82D and P187S mutants [10–12,22]. In addition, perturbations of the structural stability of the CTD by these mutations in the NQO1$_{holo}$ and NQO1$_{apo}$ states could also be associated with increased intracellular degradation and/or catalytic alterations. Thus, we might expect that specific structural perturbations in NQO1$_{holo}$ and NQO1$_{apo}$ due to S82D and P187S would lead to different molecular consequences and to diverse molecular mechanisms associated with their loss-of-function phenotypes. However, to date, no high-resolution experimental study has explained these effects.

To provide quantitative analysis on the perturbations caused by the mutations S82D and P187S on NQO1$_{holo}$ and NQO1$_{apo}$, we evaluated the overall effects of these sequence variations on the structural stability of these two states by comparing their HDX kinetics using the Δ%D$_{av}$ parameter [23](Fig. 3, Figures S8-S9). This parameter reflects changes in amplitudes and/or kinetics of HDX in a single metric, allowing us to compare two given protein species; in this case, we compared a given mutant vs. WT NQO1 in a given ligation state (NQO1$_{holo}$ and NQO1$_{apo}$, respectively) (Fig. 3 and S9-S10).

In NQO1$_{apo}$, both S82D and P187S led to significant local destabilization that propagated to almost the entire structure of the protein, although these effects were much more intense for P187S (Fig. 3A and B). Interestingly, FAD binding led to markedly different effects in these two variants vs. the WT protein. In S82D, structural destabilization in the N-terminal region (residues 1–110) was also observed upon FAD binding, and in the region 57–76 this was even larger than in NQO1$_{apo}$, whereas the mild destabilization found in the C-terminal region of NQO1$_{apo}$ was largely reduced upon FAD binding (Fig. 3A). Thus, perturbations of NQO1$_{holo}$ due to S82D should be considered *more local to*
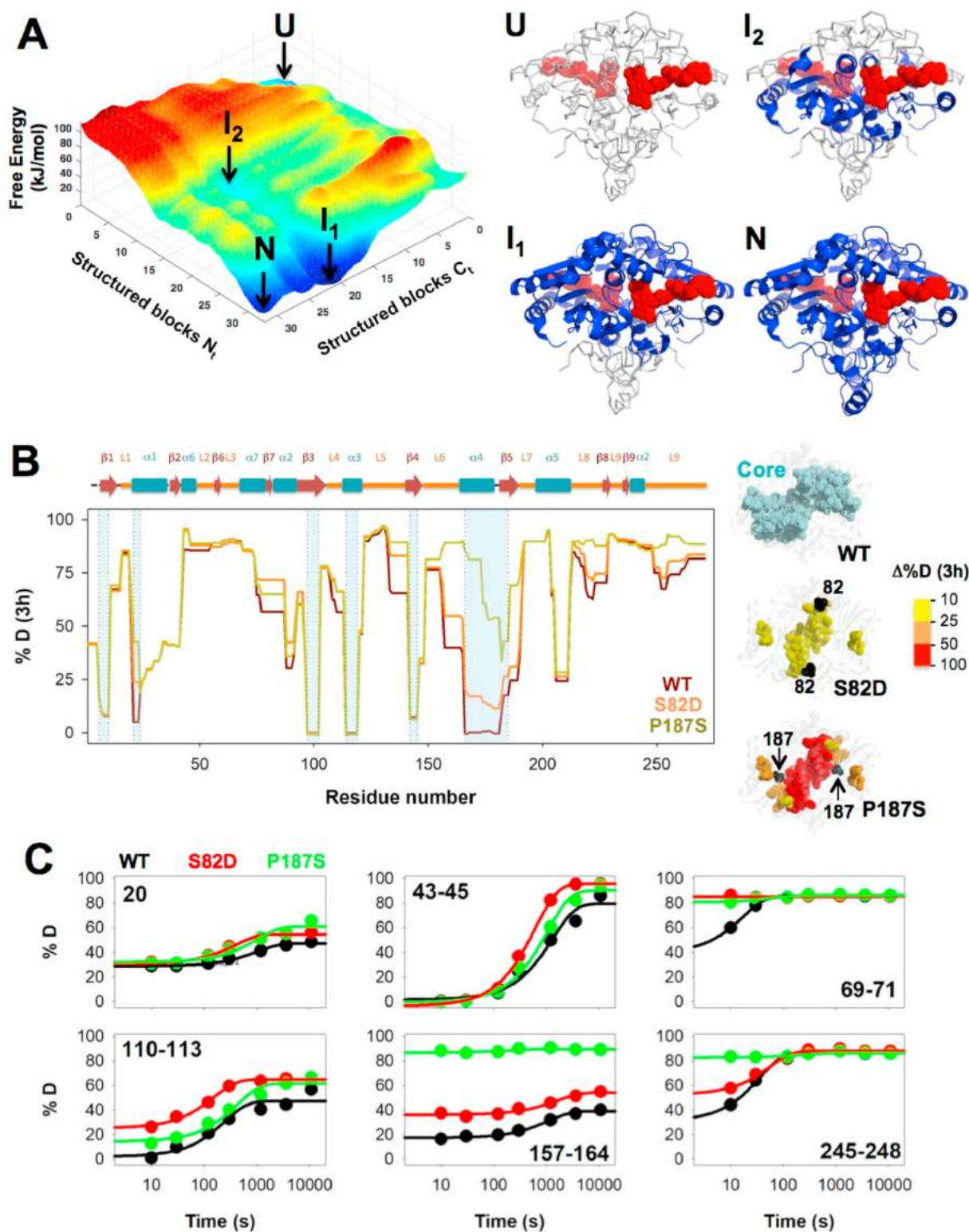
Fig. 2. Effects of S82D and P187S on the conformational ensemble of NQO1_apo. A) The left panel shows the free energy landscape of WT NQO1_apo generated as a function of the number of structured blocks in the NTD and CTD; In the right panel, the structured regions for four relevant conformational substates (N, I_1, I_2 and U) are mapped onto the crystal structure of NQO1 as blue cartoon (PDB code 2F1O [25]). For reference, the bound FAD is shown as dot representation in red. B) HDX (% D) after 3 h at 25 °C for WT_apo, S82D_apo and P187S_apo. Secondary structure elements are shown according to Ref. [19]. Shadowed regions indicate those belonging to the stable core (<20% D after 3 h) [23]. Colour codes in structural representations indicate values of Δ%D after 3 h according to the scale using WT_apo as reference. Structural representations (right panels) show the location of the stable core in WT_apo as well as the destabilization by S82D and P187S (as a Δ%D after 3 h); C) HDX analysis for WT, S82D and P187S in the NQO1_apo showing kinetic heterogeneity along 3 h exchange time and the effect of these two amino acid substitutions. HDX data for WT are from Ref. [23]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 3. The effects of S82D and P187S on the structural stability of NQO1$_{holo}$ and NQO1$_{apo}$ using the mutated site as reference (MCI model). Effects of S82D (A) and P187S (B) on $\Delta\%D_{av}$ for NQO1$_{holo}$ and NQO1$_{apo}$. Plots show the values of $\Delta\%D_{av}$ using the WT protein as a reference in each ligation state as a function of the distance between the mutated site and the different segments evaluated by HDX. Structural representations show the location in the structure of perturbed regions. C) Distance-dependent perturbations shown in panels A and B are grouped according to their distance to the mutated site in four groups ($<$10 Å, 10–20 Å, 20–30 Å and $>$30 Å). For segments in each group, distances and $\Delta\%D_{av}$ values are represented as mean $\pm$ s.d. Lines are exponential fits. The value of d_c (in Å) is only reported in those cases in which a clear exponential behavior is observed. D) Structural representation of the effects of S82D and P187S on the FBS. Values of $\Delta\%D_{av}$ larger than 10% are depicted as dot representation. Note that a positive value of $\Delta\%D_{av}$ indicates destabilization vs. the WT protein. Colour codes in structural representations indicate values of $\Delta\%D_{av}$ according to the scale shown. In all cases, the structural model PDB code 2F1O [25] was used for display and calculations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

the mutated site. In P187S, part of the destabilizing effects in the N-terminal region (residues 1–110) were reduced upon FAD binding, while the C-terminal region remains largely perturbed even in the NQO1$_{holo}$ state (Fig. 3B). P187S also destabilized to a larger extent the monomer: monomer interface (MMI) than S82D, and this effect was similarly observed in the NQO1$_{holo}$ and NQO1$_{apo}$ states of both mutants (Figure S11). An interesting finding is that both mutations S82D and P187S altered the stability of regions located far from the mutated site in both NQO1$_{holo}$ and NQO1$_{apo}$, although the magnitude of this effect was evidently stronger for P187S and depended on the ligation state (Fig. 3A–B).

To globally evaluate the different impact of S82D and P187S on the structural stability of the NQO1 structure, we first investigated the propagation of stability effects using a model (MCI model) in which the reference system was the mutated site and propagation was considered isotropic (i.e. it just depended on the distance to the mutated site). These analyses further illustrated the different effects of S82D and P187S on protein structural stability (Fig. 3A–C). Destabilization caused by S82D was rather mild in the NQO1$_{apo}$ state, and showed a very weak (and linear) dependence on the distance to the mutated site (i.e. propagated *similarly* to regions close and far from the mutated site). Conversely, in the NQO1$_{holo}$ state, these effects were much stronger and exponentially depended on the distance to the mutated site, with a characteristic distance of ~13 Å (Fig. 3C). In the case of P187S, the perturbations were on average moderate for both NQO1$_{holo}$ and NQO1$_{apo}$ states, and also showed a weak (and linear) distance-dependence from the mutated site affecting regions located in virtually the entire 3D structure (Fig. 3C).

To understand the large decrease in FAD binding affinity caused by S82D and P187S, we analyzed their effects in the structural stability of FBS in NQO1$_{apo}$ and NQO1$_{holo}$ (Fig. 3D). Alterations in the stability of the FBS on either of these states may contribute to reduce FAD binding affinity. Decreased stability of the FBS in NQO1$_{apo}$ would reflect a reduced population of FAD binding competent states and thus, an increased penalization to binding due to the conformational transition (i.e. from binding non-competent to competent states). Reduced stability of the FBS in the bound state would facilitate FAD release (i.e. increase the dissociation rate constant) reflecting destabilization of the complex. The S82D mutation affected mildly the stability of the set of residues in the FBS, both in the NQO1$_{apo}$ and NQO1$_{holo}$ states (Fig. 3D). The effects due to P187S were somewhat different. Although P187S also affected a set of residues in the FBS in both NQO1$_{apo}$ and NQO1$_{holo}$, in some residues these effects were very large (particularly relevant for Tyr156 and His162), and the set of residues affected in this variant as NQO1$_{holo}$ was more extensive (Fig. 3D). Thus, even though both S82D and P187S affect the FAD binding affinity to a similar extent [10,11,22], the molecular details of the alterations in the structural stability of the FBS in NQO1$_{holo}$ and NQO1$_{apo}$ caused by these mutations are clearly different.

3.3. Structural perturbations caused by S82D and P187S are differently modulated by dicoumarol binding

Binding of dicoumarol (Dic) to P187S is 6 to 10-fold weaker than to the WT protein [13,21]. The origin of this lower affinity seems to be the thermodynamic destabilization of the CTD of P187S$_{holo}$, that thus required a *folding-coupled-to-binding* process of the CTD to reach binding competent states for Dic binding [12,13,21]. This alteration in the structural stability of the CTD has additional phenotypic consequences: P187S$_{holo}$ is efficiently degraded inside cells through recognition of the unstable CTD by the proteasomal degradation pathways [12,21]. Consequently, the abundance of this variant (in contrast to the WT protein) strongly responds to supplementation with Dic that may reduce ubiquitination of the CTD and consequently its degradation [12,21]. The effect of Dic supplementation on the intracellular stability of the S82D variant has not been investigated so far.

To investigate the effects of Dic binding on the structural stability of S82D and P187S, we carried out HDX kinetic experiments in the presence of Dic (the NQO1$_{dic}$ state) (Figure S12). Binding of Dic to S82D leads to a quite similar structural stabilization to that found for the WT protein (Fig. 4A and Figure S12). Most of the regions that were destabilized in the N-terminal part of the protein (residues 1–110) in S82D$_{holo}$ remained destabilized upon Dic binding, with the exception of two regions that showed enhanced stabilization upon inhibitor binding to S82D vs. WT (regions 11–20 and 107–113). The effects of Dic binding to P187S$_{holo}$ were much more dramatic (Fig. 4B and S12). Dic binding overcome most of the destabilizing effects observed in P187S$_{holo}$ across the entire protein, with the main exception being the CTD in which these effects were reduced but still noticeable (Fig. 4B). Similar analyses focused on these effects on the MMI (Figure S13) yielded essentially the same results (note the remarkable reduction in the MMI of P187S$_{holo}$ upon binding the inhibitor). Application of the MCI model additionally supported the different global response of S82D and P187S upon Dic binding in terms of structural stability (Fig. 4A–C). The exponential dependence of the S82D was similar in the NQO1$_{holo}$ and NQO1$_{dic}$, whereas Dic binding to P187S in the NQO1$_{holo}$ state strongly diminished the destabilizing effect vs. the WT protein (Fig. 4C).

NQO1$_{holo}$ and NQO1$_{dic}$ resemble two functionally relevant states of the enzyme during its catalytic cycle. The former is a pre-catalytic state that is ready for NAD(P)H binding and fast FAD reduction, whereas the latter may represent either the ternary complex with the NAD(P)H or the transition state for the NAD(P)H bound towards the reduced flavin [23]. According to this view, alterations in the structural stability of the FBS and DBS due to S82D and P187S may underlie alterations in the catalytic cycle, in both the reductive and oxidative half-reactions [24]. In WT$_{holo}$, the FBS is particularly stabilized upon FAD binding and also some stabilization is found for the DBS, while binding of Dic leads to strong stabilization of both sites [23]. We have previously interpreted these effects as the formation of a highly *rigid* ternary complex that may allow proper orientation between NAD(P)H and FAD and highly efficient hydride transfer, thus explaining the high rate constant of this process in the WT enzyme [24]. It must be noted that the reductive half-reaction is the rate-limiting step of this catalytic cycle and that communication between the two active sites seems to exist (consistent with negative cooperativity towards NAD(P)H in both steady- and pre-steady state kinetics [24,45]). An important role of the MMI in this allosteric communication has been proposed due to the strong effects of FAD and Dic binding on the stability of this region [23].

We then focused our analyses on the effects of S82D and P187S on the structural stability of the FBS and DBS in NQO1$_{holo}$ and NQO1$_{dic}$ (Fig. 4D). The mutant S82D only showed significant structural destabilization of the FBS as NQO1$_{holo}$, thus suggesting that this mutant might only moderately affect the rate of the reductive half-reaction. The effects of P187S were again much stronger. First, in NQO1$_{holo}$, the structural stability of both the FBS and DBS were largely decreased in this variant, supporting that to form the ternary complex with NAD(P)H competent for flavin reduction, we might expect a larger structural reorganization (and possibly a large entropic penalization; actually this penalization has been observed by titration calorimetry [13,21]). In NQO1$_{dic}$, the FBS of P187S still remains destabilized to some extent which might imply an increased flexibility in the ternary complex (or the transition state for hydride-transfer reaction) and thus a lower efficiency in the flavin reduction step. All this might contribute to increase the kinetic barrier for the rate-limiting step of the reductive half-reaction due to P187S.

3.4. Anisotropic propagation of mutational and ligand binding effects

Visual inspection of the propagation of mutational effects due to S82D and P187S in different ligation states as well as global analysis by the isotropic MCI model (Figs. 3 and 4), strongly supported the existence of anisotropy in the propagation of stability effects which is distinct for different mutations and ligation states. To get a deeper insight into this notion, we have applied two additional models to interpret the results from HDX regarding the propagation of stability effects across the NQO1
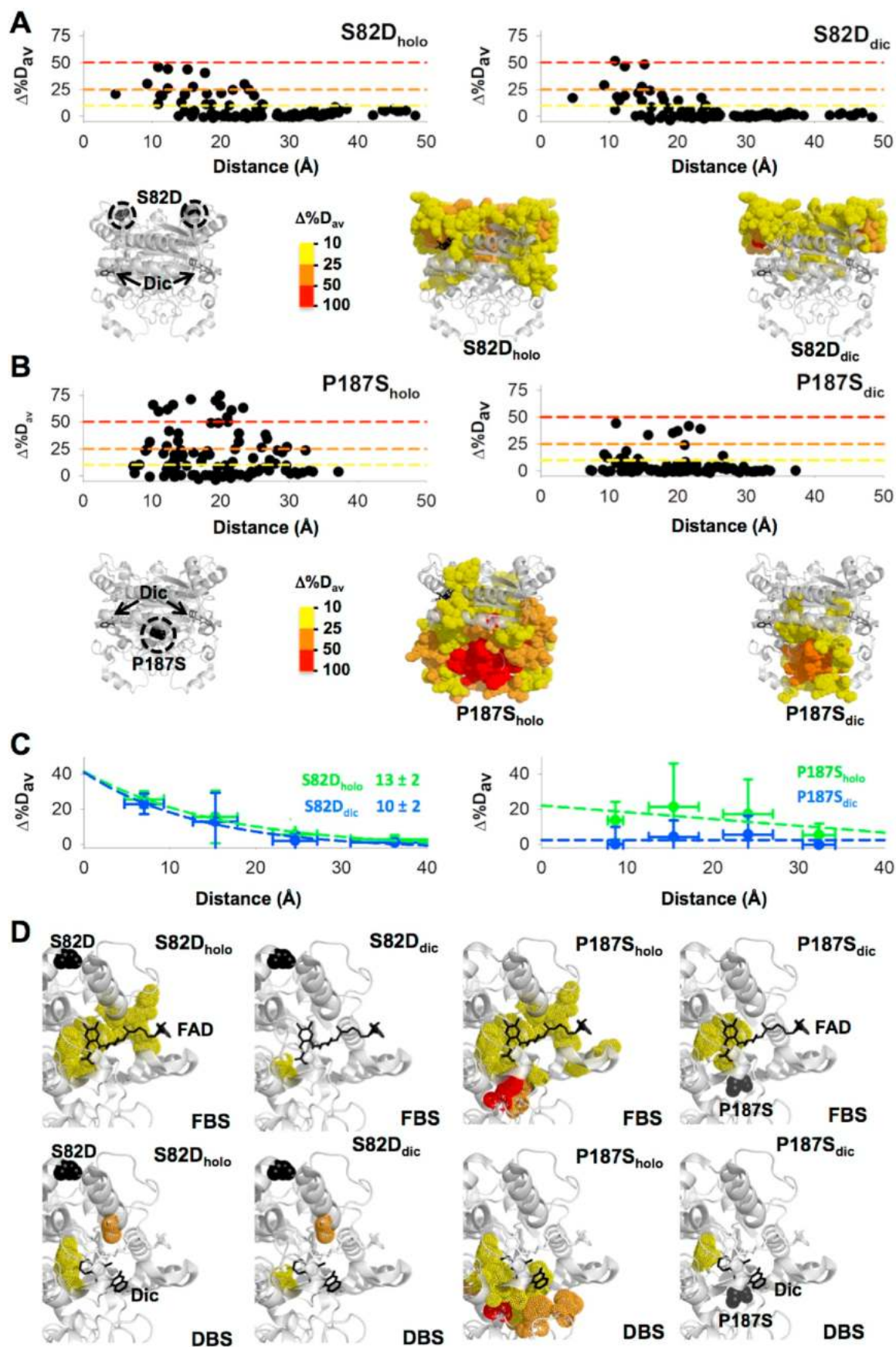
(caption on next page)

Fig. 4. The effects of S82D and P187S on the structural stability of NQO1$_{holo}$ and NQO1$_{dic}$ (MCI model). Effects of S82D (A) and P187S (B) on Δ%D$_{av}$ for NQO1$_{holo}$ and NQO1$_{dic}$. Plots show the values of Δ%D$_{av}$ using the WT protein as a reference in each ligation state as a function of the distance between the mutated site and the different segments evaluated by HDX. Structural representations show the location in the structure of perturbed regions. C) Distance-dependent perturbations shown in panels A and B are grouped according to their distance to the mutated site in four groups (<10 Å, 10–20 Å, 20–30 Å and >30 Å). For segments in each group, distances and Δ%D$_{av}$ values are represented as mean ± s.d. Lines are exponential fits. The value of d$_c$ (in Å) is only reported in those cases in which a clear exponential behavior is observed. D) Structural representation of the effects of S82D and P187S on the FBS and DBS. Values of Δ%D$_{av}$ larger than 10% are depicted in dot representation. Colour codes in structural representations indicate values of Δ%D$_{av}$ according to the scale shown. In all cases, the structural model PDB code 2F1O [25] was used for display and calculations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

structure. In these two models, the reference system depends on the ligation state: for the NQO1$_{apo}$ the reference is 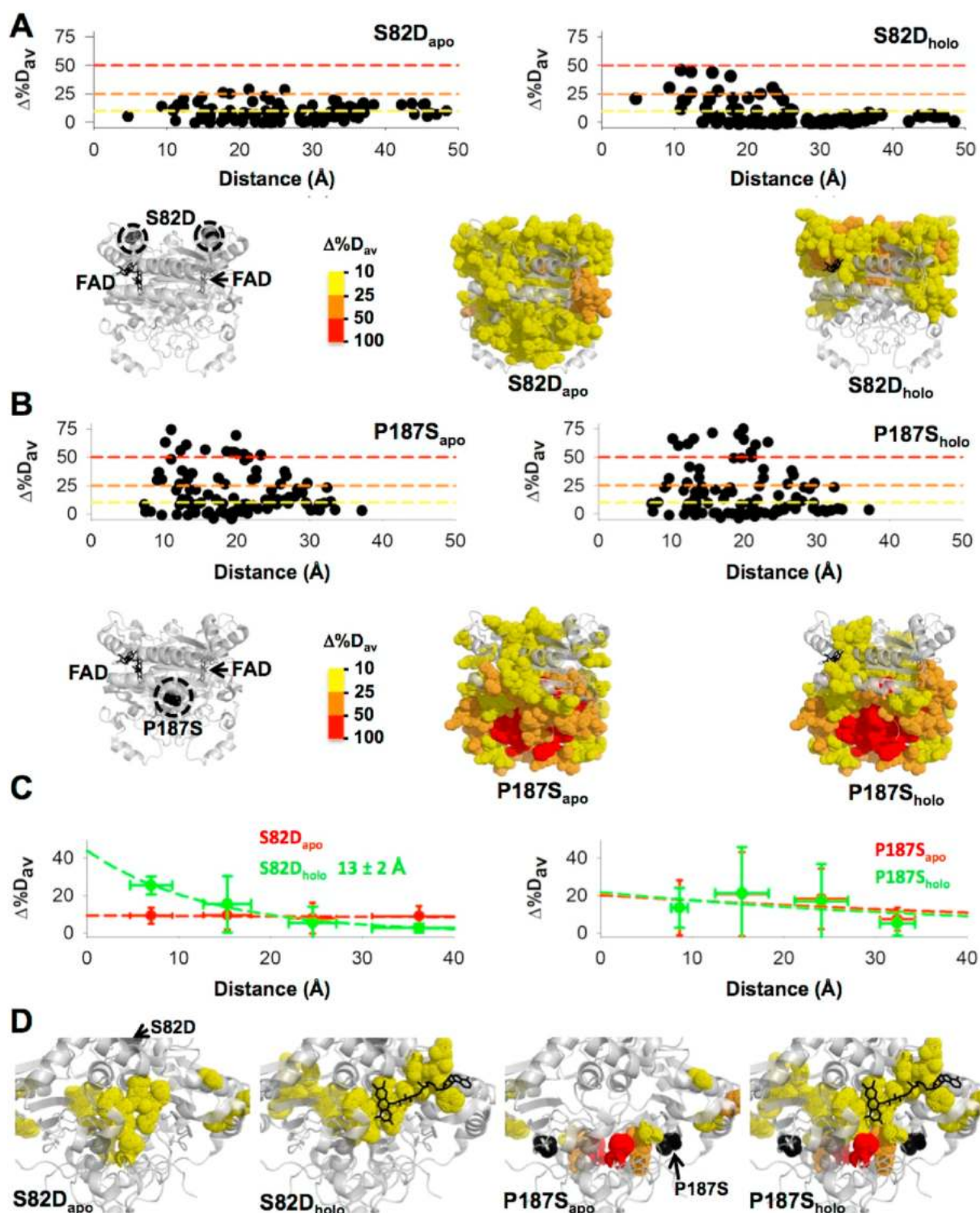the mutated site while in NQO1$_{holo}$ and NQO1$_{dic}$ the reference is the center of mass of the corresponding ligand (for FAD and Dic, respectively). These two models differ regarding the directional propagation of mutational and ligand-binding effects: the LDI model considers isotropic propagation (all directions are considered as equal, and thus, distance is simply related to the perturbed site) whereas the LDA method explicitly considers anisotropic propagation by dividing the structure in six different directions or zones, but maintaining the same concept of distance (Fig. 5A).

The LDI method provides similar results to those of the MCI method (an expected result since both are isotropic approaches) for S82D$_{apo}$ and P187S$_{apo}$ (Fig. 5B). Interestingly, the LDI method shows some differences between S82D and P187S regarding ligand binding (Fig. 5B). The distance dependence of S82D is similar in the NQO1$_{apo}$ and NQO1$_{holo}$ states, whereas upon Dic binding the decay is somewhat more strongly distance-dependent than in the WT protein. P187S showed similar behaviour in the NQO1$_{holo}$ and NQO1$_{apo}$ states, but Dic binding essentially abolished the destabilizing effects of P187S (vs. WT NQO1) (Fig. 5B).

The LDA method provides some interesting insight into anisotropic propagation of stability effects. The entire protein structure is divided into six equal sections called *zones* intersecting at the reference point of each ligation state. The zone 1 contains either the side chain of the mutated site in case of NQO1$_{apo}$ or the center of mass of the ligand interacting residues in case of the bound forms. Subsequent zones were marked in a clockwise manner (Fig. 5A). In simple words, zone 1 along with the adjacent zones 2 and 6 encompass residues in the direction of the perturbation whereas the other zones contain residues opposite to the perturbation direction. Using this method, mutational and ligand-binding effects on stability showed significant directionality (Fig. 5B). Regarding S82D$_{apo}$, zones 1, 5 and 6 show the strongest effects (although milder than those observed for P187S), with zones 1 and 5 showing certain distance-dependence whereas zone 6 does not (Fig. 5B). In S82D$_{holo}$, the strongest effects are observed for zones 1 and 6, although they seem to differ in the distance-dependence propagation (Fig. 5B). In S82D$_{dic}$, the strongest effects are found for zone 6, and to a lesser extent for zone 1, with a noticeable distance-dependence (Fig. 5B). In the case of P187S$_{apo}$, zone 1 and 6 show by far the strongest effects on stability, although their distance-dependences are different (Fig. 5B). The behavior of P187S$_{holo}$ is a little odd, since multiple zones show significant stability perturbations, and for zones 1 and 2, these perturbations are essentially distance-independent (Fig. 5B). The situation for P187S$_{dic}$ is apparently simpler, because destabilizing effects are basically abolished with the exception of some long-range perturbations (about 25 Å) of moderate intensity observed in zones 2 and 3. Overall, these anisotropic perturbation analyses provide at least two interesting conclusions: i) propagation of mutational and ligand-binding effects are quite anisotropic: these occurring in the direction of the perturbation (i. e. via zones 1, 2 and 6) are much more frequent and stronger; ii) the directional preference (in terms of magnitude of the *original* perturbation and the distance-dependence dissipation) for the propagation of mutational and ligand-binding effects strongly depend on both the mutation and the ligand bound.

3.5. Effects of S82D and P187S on enzyme catalysis

The reductive and oxidative half-reactions of WT NQO1 are not simple processes [24]. The two flavin molecules located in the two active sites of NQO1 WT are reduced with widely different kinetics: one of the FAD molecules is reduced 10-fold faster than the other one in the dimer (*fast* vs. *slow* steps) [24]. In the oxidative half-reaction, one molecule of reduced flavin, FADH$_2$, is oxidized extremely fast (*very fast* step) whereas the second FADH$_2$ molecule is oxidized more slowly (*fast* step). These features of the catalytic cycle of WT NQO1 are summarized in Fig. 6A. The non-equivalence of the FAD/FADH$_2$ molecules bound to the two active sites has been proposed to represent a case of functional negative cooperativity [24].

Our analyses using HDX (NQO1$_{holo}$ vs. NQO1$_{dic}$, regarding the FBS and DBS; Fig. 4D) actually suggested that P187S, and to lesser extent S82D, may affect the catalytic cycle of NQO1. To test this, we first carried out experiments on the reductive half-reaction with the S82D and P187S mutants using equimolar concentrations of NQO1$_{holo}$ and NADH (Figure S14). These experiments revealed some interesting effects (Fig. 6B). These two mutations did not alter the existence of two different pathways in the reduction of the flavin and hardly affected the observed rate constant (k_{obs}) for the fast step (step A→B, k_{obs1}) (Fig. 6B). However, at these conditions, the reduction of the second flavin molecule (step B→C, k_{obs2}) occurred 4.5- and 70-fold more slowly for S82D and P187S, respectively (Fig. 6B). Further experiments to evaluate the NADH concentration dependence of the reductive half-reaction supported mild effects of these mutations on the limiting value of k_{obs1} (i.e. k_{HT1}, where HT refers to hydride-transfer) and revealed a modest 4-fold decrease in K_{d1}^{NADH} for P187S (Fig. 6C–D and Table S5). Thus, the catalytic efficiency (k_{HT}/K_d^{NADH}) of S82D in the fast reduction step was similar to that of WT NQO1, while P187S decreased it by 2.4-fold (Table S5). Importantly, the effects of S82D and P187S were much more pronounced for the slow FAD reduction process (Fig. 6C–D and Table S5). The mutation S82D had a marginal effect on k_{HT2} but increased by 5-fold the K_{d2}^{NADH}. The effects of P187S were much more pronounced, with a decrease in k_{HT2} of 8-fold and an increase of 9-fold in the K_{d2}^{NADH}. Therefore, both mutations significantly affected the slow FAD reduction step, with a decrease in catalytic efficiency of 8-fold and 80-fold for S82D and P187S, respectively (Table S5). These observations were also supported by experiments carried out with NADPH in the reductive half-reaction (Figure S15), or using DCPIP as substrate in the oxidative half-reaction (in this case, NQO1$_{holo}$ had been previously reduced with NADH) (Fig. 6E, Figure S16 and Table S6), although these reactions were too fast to carry out a more detailed characterization [24]. Since the reductive half-reaction is rate-limiting in the catalytic cycle of NQO1 [24], these results suggest that these two mutations (particularly P187S) exacerbate the negative functional cooperativity found in NQO1 WT under pre-steady state conditions. In particular for P187S, one of the active sites became essentially *useless*, thus likely contributing to reduce the specific activity inside the cell and constituting some sort of half-of-sites reactivity. We may speculate that these remarkable effects of P187S, particularly on the slow step, as well as this half-of-sites reactivity, could result from: i) the structural destabilization of NADH binding site (inferred from the effects observed in the FBS and DBS of P187S$_{holo}$, Fig. 4D) that could affect directly catalysis; ii) destabilization of the MMI (Figure S13) that may contribute to alter the
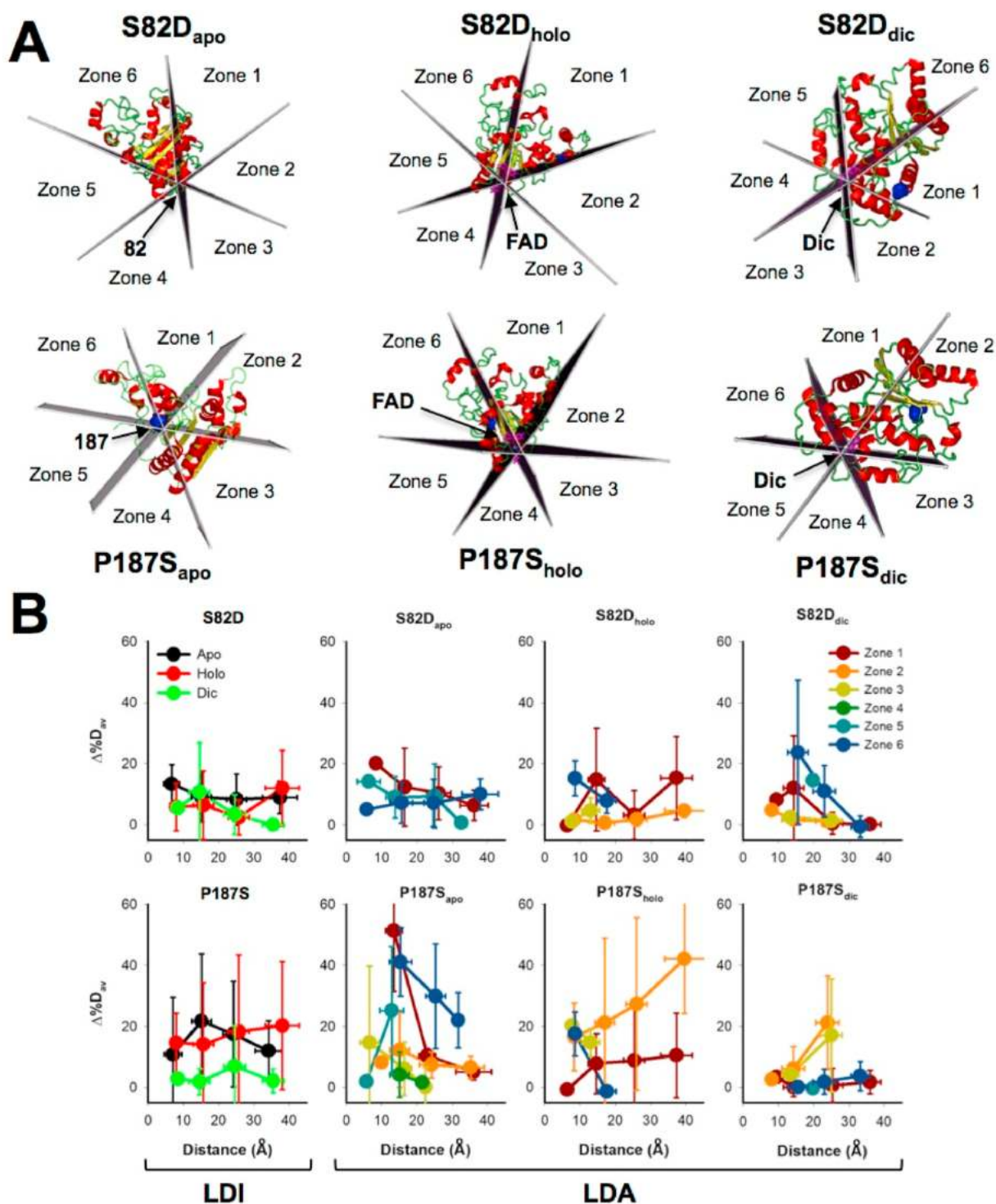
Fig. 5. Anisotropic propagation of mutational effects in different ligation states. A) Representation of the procedure used to generate different zones for the anisotropic analysis (LDA model). The mutated residue is shown as a blue sphere. B) Isotropic (LDI model) vs. anisotropic (LDA model) analysis of the distance-dependent propagation of mutation- and ligand-binding effects. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

communication of stability effects between active sites during the catalytic cycle of NQO1.

To further explore these effects on the catalytic cycle of NQO1, we then carried out experiments on the reductive half-reaction to compare the behaviour towards NADH and NADD. These experiments were carried out using equimolecular concentration of enzyme and coenzyme to overcome technical limitations (particularly for the temperature-dependent studies) [24]. These experiments yielded the kinetic isotope

effects (KIE) for both the fast and slow FAD reduction steps (Fig. 7A–B and Table S7). The first point to note is that when using NADD the fast reduction step is slowed down, resulting for all variants in similar KIEs (1.8–2.0). These relative low KIEs have been associated with the transition state being asymmetrical or nonlinear [24,46]. However, some differences were observed again for the slow reduction step. Basically, this slow step shows the same KIE as the fast step for WT NQO1, but its value is lower in the mutants, becoming essentially 1 for P187S
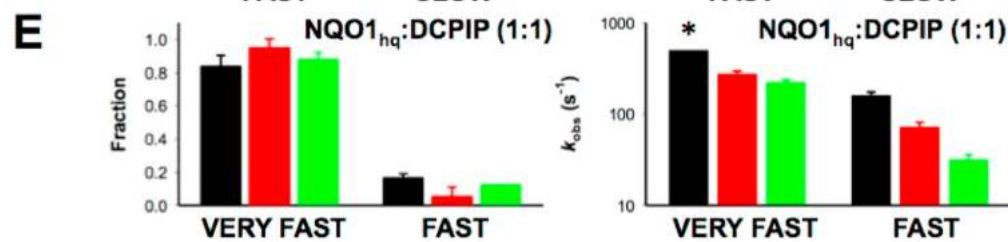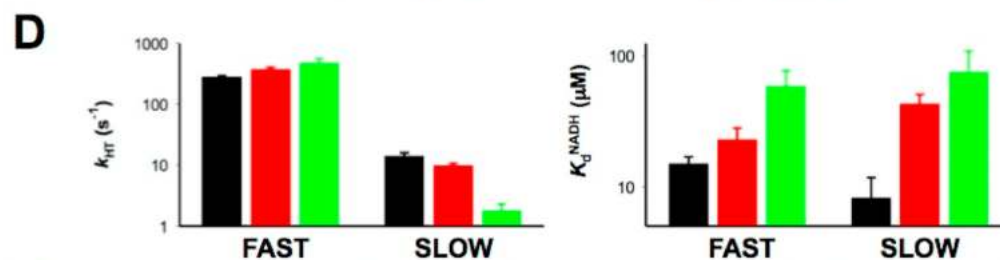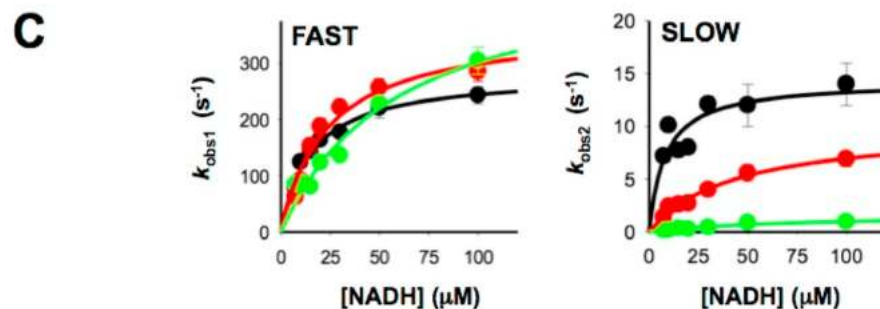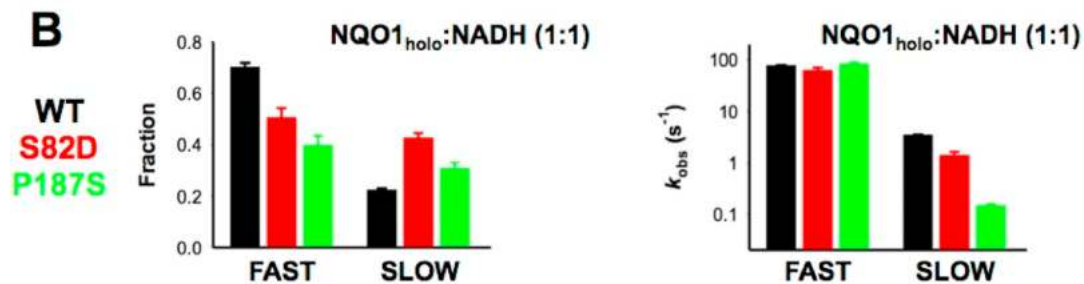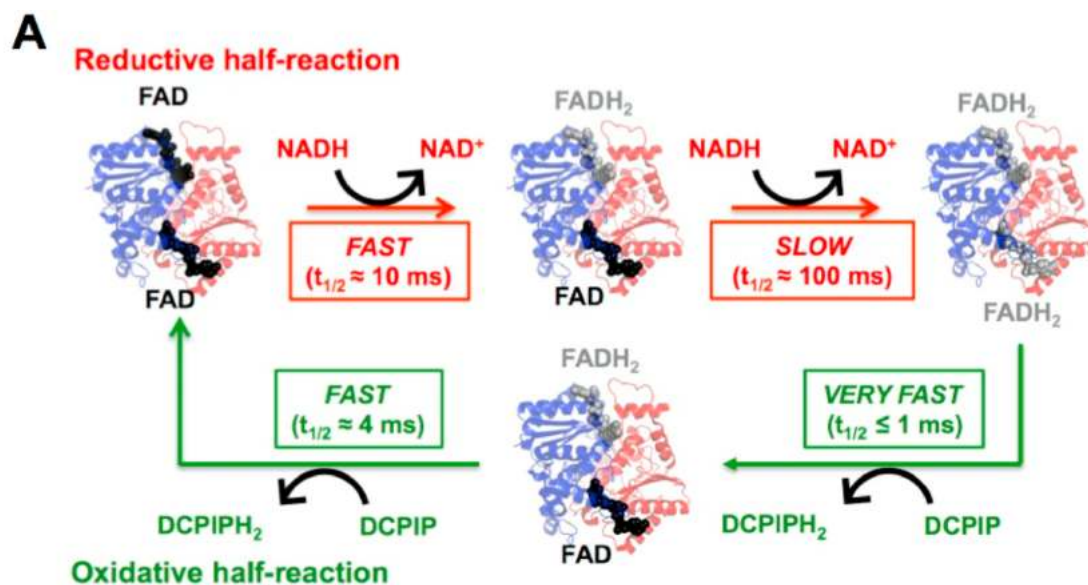
(caption on next page)

Fig. 6. Effects of S82D and P187S on pre-steady state enzyme kinetics. A) Schematic representation of the NQO1 WT catalytic cycle (from Ref. [24]). The reductive half-reaction occurs through two main consecutive processes (referred to as *fast* and *slow* steps), likely reflecting the sequential reduction of the two FAD molecules in the NQO1 dimer by NADH with widely different kinetics. Similarly, the oxidative half-reaction occurs through two pathways with very high rate constants, (referred to as *very fast* and *fast* steps) using DCPIP as substrate; B-D) Kinetics for the reductive half-reaction of NQO1 variants by NADH; B) Kinetics using 7.5 μM NQO1:NADH. The left panel shows the fraction corresponding to the fast and slow processes determined as the fractional change in absorbance (average of 445–455 nm) corresponding to the spectral deconvolution (Figure S14). The right panel shows the observed rate constants (k_{obs}) for each process from kinetic analysis (Figure S14). C) Dependence of the k_{obs} values on NADH concentration for the fast (k_{obs1}) and slow (k_{obs2}) processes in the reductive half-reaction. Lines are fits to equation (2). D) Limiting values of k_{obs} (k_{HT}) and NADH equilibrium binding constants (K_d^{NADH}) for the fast and slow processes in the reductive half-reaction with NADH. E) Kinetics for the oxidative half-reaction of reduced variants (NQO1$_{hq}$) using DCPIP as substrate (7.5 μM NQO1:NADH). The left panel shows the fraction corresponding to the very fast and fast processes determined as the fractional change in absorbance (average of 445–455 nm) corresponding to the spectral deconvolution (see Figure S16 and Table S6). The right panel shows the k_{obs} values for the very fast and fast processes. The asterisk indicates that $k_{obs} > 500\ s^{-1}$. Data for WT NQO1 are reproduced from Ref. [24]. All experiments were carried out at 6 °C. In panels B–E, data correspond to WT (black), S82D (red) and P187S (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(Fig. 7A–B and Table S7). This result reinforces the notion of P187S (and to a lower extent, S82D) affecting much more strongly the kinetics and activation energetics of the slow reduction pathway of FAD. In addition, these KIEs were essentially temperature-independent for both fast and slow FAD reduction steps in all variants (Fig. 7A–B), with the exception of the slow step for S82D. Further analyses of these results in the context of the Arrhenius equation revealed some additional effects of these mutations on the activation parameters of the reaction (Fig. 7C–D and Table S7). Using either NADH or NADD as coenzyme, the *fast* step was marginally affected by S82D and P187S mutations regarding the values of frequency factors (A, less than one order of magnitude) and activation energies (E$_a$). However, more interesting differences were found (again) in the *slow* FAD reduction step: the S82D mutation increased the value of the frequency factor and activation energy to some extent, but the effects of P187S were much more dramatic, with a decrease in the pre-exponential factor of 6–7 orders of magnitude and a remarkable decrease in the activation energy of over 5 kcal mol^{-1} (vs. WT) (Fig. 7C). A simple calculation (based on the transition state theory) yield an increase in the activation free energy for the slow step of 2.0–2.5 kcal mol^{-1} for P187S (using NADH or NADD), thus supporting a noticeable entropic penalization for this step in P187S vs. the WT protein. A potential explanation for this is that P187S requires a larger structural reorganization to bind NADH/NADD to engage in the catalytic process, which actually agrees with our HDX analyses that showed a larger change in structural stability from NQO1$_{holo}$ to NQO1$_{dic}$ in the DBS of P187S (Fig. 4D).

KIEs and activation parameters serve to prove the nature of the chemical step and the effects of mutations on the structural organization and dynamics at the active site during catalysis [47,48]. For the fast FAD reduction step in all three variants, KIEs are temperature independent and above unity. Together with values of A$_H$/A$_D$ $\neq$ 1, these observations indicate that the mutations do not prevent the tunnel contribution to the fast HT process as reported for the WT protein [24]. In addition, the ΔE_{aD-H} values close to zero indicate in general a *tunnelling ready state* (TRS) where the donor-aceptor-distance (DAD) has been reduced during complex pre-organization motions. In this situation, *passive dynamics* (i. e. movements of active site heavy atoms that do not actively contribute to the tunnel probability but increase that to achieve tunnel ready conformations) is sufficient to bring the donor and acceptor to an adequately short distance for efficient tunnelling and HT. This suggests that S82D and P187S hardly impact the architecture and dynamics of the active site for the fast FAD reduction, and thus, the decreased HT efficiency in P187S simply arises from the lower affinity for NADH (Fig. 6C–D and Table S5). However, different contributions from dynamics to the slower FAD reduction process were observed for the S82D and P187S mutants. In WT NQO1, data were consistent with a native-like pre-organization and increased contribution of passive dynamics (ambiental) reorganization (i.e. increased A$_H$/A$_D$ ratio) to achieve HT relative to the faster process. On the contrary, the temperature-dependent KIE, low A$_H$/A$_D$ ratio and increased ΔE_{aD-H} for S82D are indicative of only H-tunnelling and of a DAD sampling coordinate (*gating* or active dynamics; i.e. environmental vibrational

enhancement that alters the DAD and as consequence energy barrier for tunnel) starting to significantly contribute when packing defects both alter/enlarge the initial DAD and decrease the force constant of the local DAD sampling mode. This is a typical situation when thermal energy is required to maintain short DADs and H-tunnelling. Finally, lack of KIE and close to unity A$_H$/A$_D$ ratio in P187S suggest that tunnelling does not contribute to the slow HT event in this variant. Such observation agrees with the reduction in k$_{HT2}$ as well as in coenzyme affinity (Fig. 6C–D and Table S5), indicating that the P187S mutation produces a negative impact in the pre-organization motions required to achive a native-like competent conformation for HT at the active site that carries out the slow step. In conclusion, the S82D and P187S mutations hardly affect the organization and dynamics of the active site for the faster HT from NADH/D to the coenzyme, but have an important negative impact in the preorganization motions to achive the reactive active site for the slower HT event. Thus, in the *slow* reduction gating partially overcomes the negative impact on preorganization for S82D, but not for P187S.

3.6. The intracellular stability of NQO1 is differentially modulated by ligand binding in S82D and P187S through effects on the CTD stability

The stability of the CTD is key to understand the intracellular degradation of NQO1 through the ubiquitin-dependent proteasomal activity [7,12,26]. In WT NQO1, under standard riboflavin supplementation, most of the protein likely populates the FAD bound state, whereas riboflavin starvation promotes the accumulation of newly synthesized NQO1$_{apo}$ that is efficiently targeted to degradation through its CTD [7]. In P187S, destabilization of the CTD even in the NQO1$_{holo}$ state has been associated with its low intracellular stability due to enhanced ubiquitination of this domain that accelerates degradation [7, 21]. Consistently, Dic binding stabilizes the CTD and protects P187S towards ubiquitin-dependent proteasomal degradation in cells [7,21].

Our HDX analyses revealed that the CTD of S82D is moderately destabilized in NQO1$_{apo}$, whereas this effect is smaller as NQO1$_{holo}$ or NQO1$_{dic}$ (Fig. 8A). On the contrary, the CTD of P187S is largely destabilized in all ligation states, although some reduction of this destabilizing effect is observed upon Dic binding (Fig. 8A).

To further analyze these effects, we first used proteolysis by trypsin, that provides a partial proteolysis pattern consistent with initial cleavage at the CTD rendering an intermediate containing the N-terminal 240 residues [41](Figure S17). For all three variants in different ligation states, proteolysis kinetics was carried out at conditions (i.e. low protease concentration) in which the proteolysis step was rate-limiting (Figure S18). Thus, comparison of the proteolysis rate constants allows to obtain changes in thermodynamic stability of the CTD upon mutation or ligand binding (Fig. 8B, Figure S18 and Table S8) [12,49]. As NQO1$_{apo}$, the CTD in all variants is very sensitive to degradation by trypsin. FAD binding to WT NQO1 increases the stability of the CTD by ∼2.9 kcal·mol^{-1}, and this cofactor-induced stabilization is severely reduced for the mutants, with S82D and P187S showing a stabilization of the CTD of ∼1.6 and ∼0 kcal·mol^{-1} upon FAD binding. Dic binding slightly increased the stability of the CTD in WT NQO1 (by ∼ 0.4
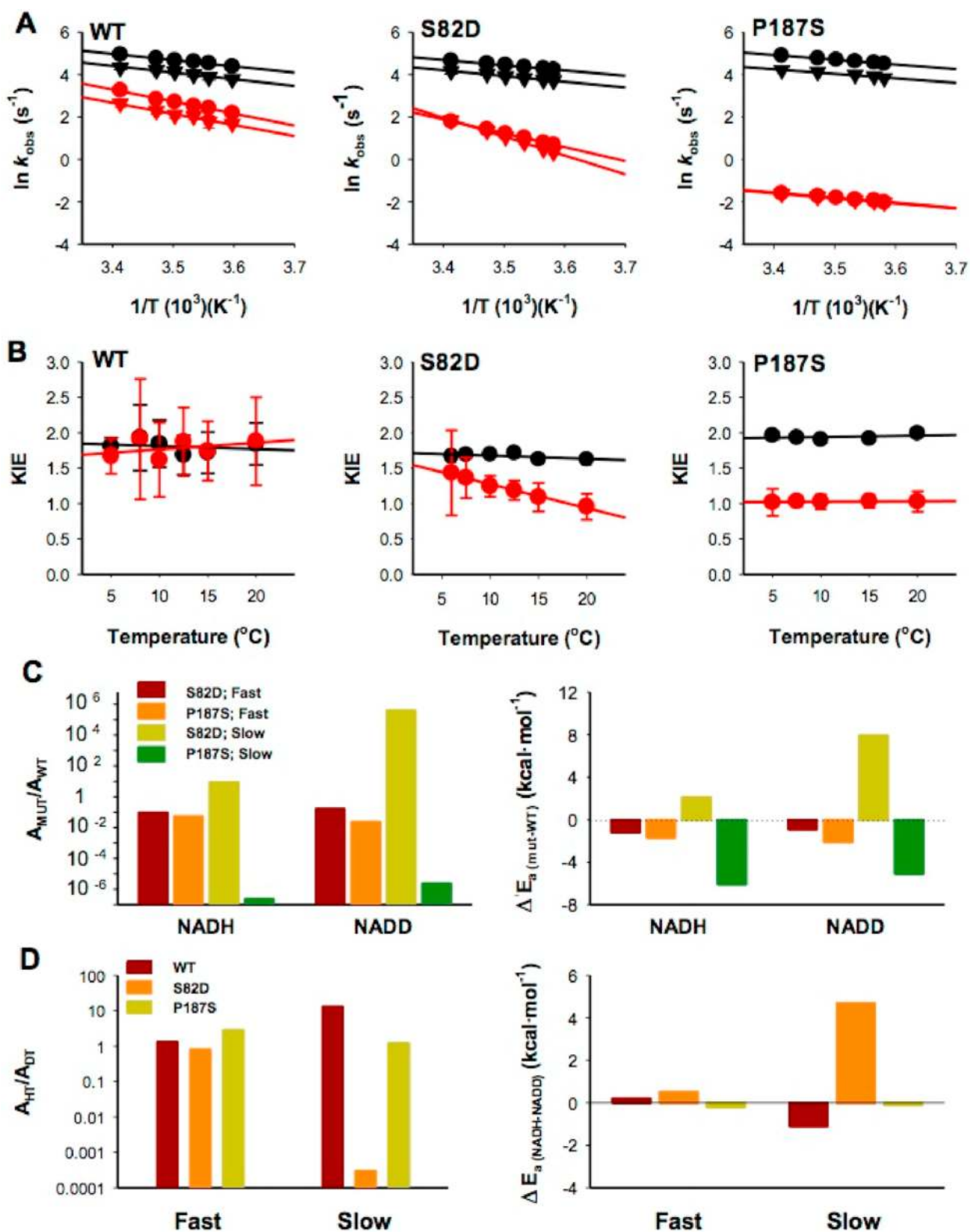
Fig. 7. Temperature-dependent enzyme kinetics and kinetic isotope effects (KIEs). A) Arrhenius plots for the reductive half-reaction carried out using NADH (circles) or NADD (triangles). Black and red symbols indicate the kinetics of fast and slow FAD reduction processes, respectively. B) Temperature dependence of KIEs. Experiments were carried out using 7.5 μM NQO1 and NADH/NADD (n ≥ 3; mean ± s.d). C) Effect of the S82D and P187S mutations (vs. WT) on the Arrhenius activation parameters (left panel, frequency factors; right panel, Activation energies). D) Arrhenius activation parameters for each variant using NADH and NADD (left panel, Frequency factors; right panel, Activation energies). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
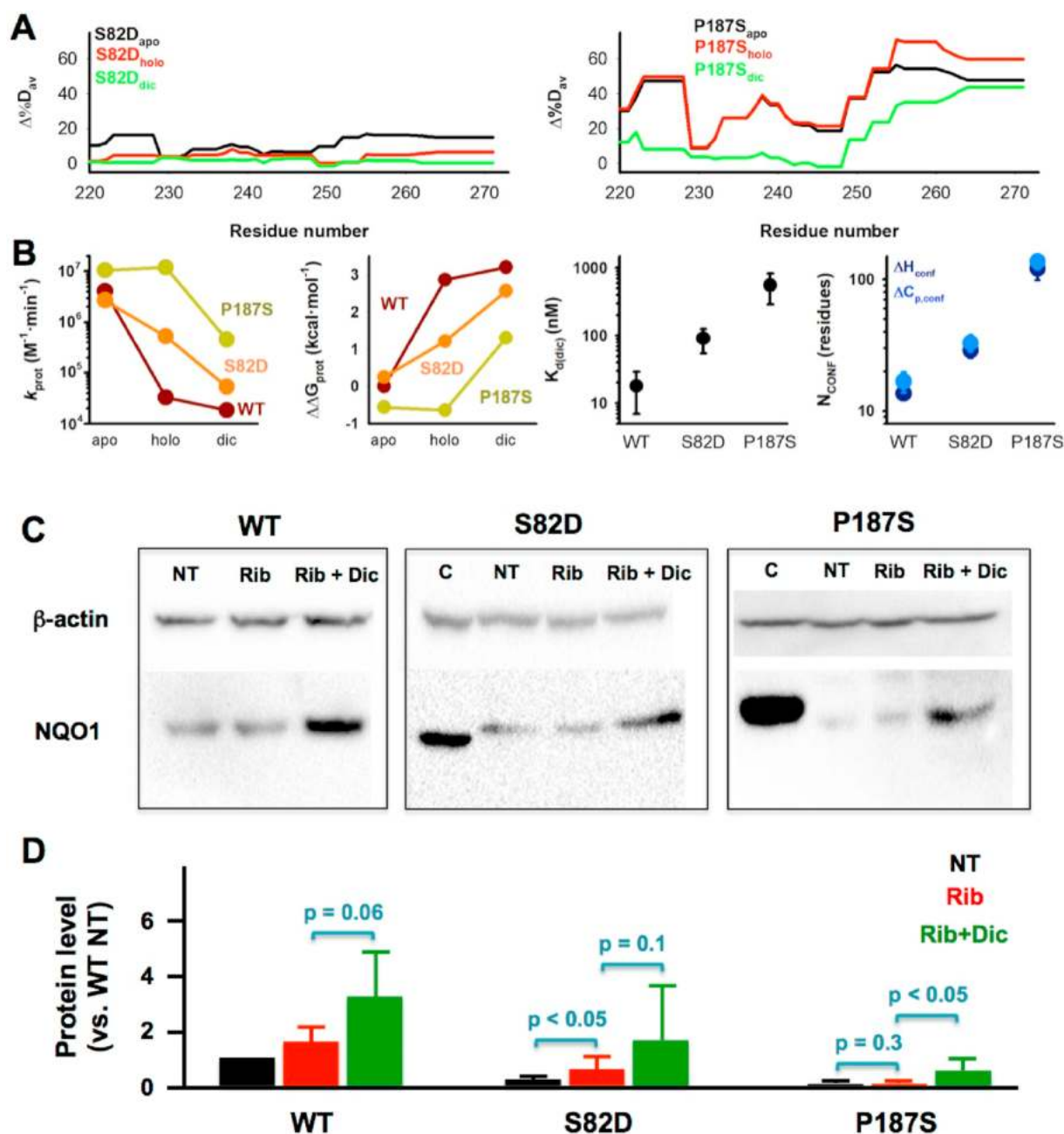
Fig. 8. The stability of the CTD determines the intracellular protein levels of NQO1 variants and their response to Dic supplementation. A) Changes in local stability of the CTD for S82D and P187S variants in different ligation states (vs. WT); B) Thermodynamic stability of the CTD for S82D and P187S derived from proteolysis kinetics with trypsin (left panels, see Figures S17 and S18) and structure-energetic analysis of Dic binding (right panels, see Figure S19). C) Representative western-blot analysis of HAP-1 NQO1-KO cells transfected with NQO1 variants. NT indicates no treatment, Rib, treatment with riboflavin; Rib + Dic, treatment with riboflavin and Dic. C indicates a control from cells transfected with WT NQO1 without treatment; D) Densitometric analysis of western-blots. Data are from at least 4 technical replicates (using samples from two independent transfections) and are normalized using WT NQO1 without treatment. Statistical analyses were carried out from one-tailed t-tests and significance reported as p values.

kcal·mol^{-1}) and this stabilization was larger for the mutants (~1.4 kcal·mol^{-1} and ~1.9 kcal·mol^{-1}, for S82D and P187S respectively). Overall, we can conclude from these experiments that the CTD of S82D is moderately destabilized in the NQO1$_{holo}$ state (by ~ 1.6 kcal·mol^{-1} vs. WT NQO1), whereas in P187S, this destabilization is larger as NQO1$_{holo}$ (~3.5 kcal·mol^{-1}) than for NQO1$_{dic}$ (~1.9 kcal·mol^{-1}).

As a complementary approach to investigate the folding/stability of the CTD, we carried out calorimetric titrations with Dic using NQO1$_{holo}$ variants (Figure S19 and Table S9). The apparent binding affinity (K_d), as well as the apparent binding enthalpies (ΔH) and heat capacities (ΔC_p) can be used to estimate the extent of the conformational change

associated with Dic binding (Equations (5) and (6)) [13]. Dic binds to WT NQO1 tightly, and binding is enthalpy-driven and displays a moderate binding heat capacity (Table S9). Using well-known structure-energetic relationships, a value for the parameter N$_{conf}$ (related to the magnitude of the *conformational change* upon binding) of about 10–20 residues can be estimated from these thermodynamic variables in WT NQO1 (Fig. 8B), consistent with the minimal conformational rearrangements observed crystallographically [25]. Similar analyses carried out with the mutants S82D and P187S reveal clearly different behaviours (Fig. 8B and Table S9). S82D shows about 5-fold lower affinity than the WT protein, and binding proceeds with a more unfavourable

entropic contribution ($\sim$4 kcal·mol^{-1}) that is partially compensated by a more favourable enthalpy change. Structure-energetics analyses on this mutant yield values of N$_{conf}$ of $\sim$30–40 residues (Fig. 8B), supporting a mild but noticeable increase in the magnitude of the conformational change associated with Dic binding vs. WT. These results, combined with our HDX analyses, suggest that S82D mildly destabilizes the CTD in the NQO1$_{holo}$ state. The behaviour of P187S regarding Dic binding is vastly different (Fig. 8B and Table S9). Its affinity for Dic is reduced by $\sim$30-fold, and the thermodynamic signature is similar to the folding of a small protein upon Dic binding: a 25 kcal·mol^{-1} entropic penalization that is largely compensated by a favourable enthalpic contribution (Table S9). Structure-energetics calculations on this variant support that about 100 residues must fold to bind Dic, consistent with an extreme destabilization of the CTD in this variant, particularly as NQO1$_{holo}$ (Fig. 8A–B).

We observed, at least qualitatively, good agreement on the effects of S82D and P187S on the stability of the CTD in different ligation states and using different approaches (HDX, proteolysis and ITC). Basically, S82D causes mild to moderate effects on the stability of the CTD, although these can be larger as NQO1$_{holo}$, whereas P187S strongly destabilizes the CTD in all ligation states, although this effect is smaller in the NQO1$_{dic}$ state. Assuming that the thermodynamic stability of the CTD mainly dictates the intracellular stability of NQO1 variants, these analyses allow us to hypothesize that: i) in the absence of Dic supplementation, the intracellular levels of S82D (as a proxy for their proteasomal degradation) should be intermediate between those of WT and P187S; ii) Dic supplementation should boost P187S protein levels (i.e. intracellular stability) to a higher extent than those of the mutant S82D.

To test these hypotheses, we stably transfected HAP-1 cells (that are knocked-out in NQO1) with NQO1 WT and the mutants P187S and S82D, and we grow them in a standard medium (IMDM) with or without supplementation with riboflavin (Rib) or Rib + Dic, and determined their protein levels by western-blot (Fig. 8C–D and Figure S20). The first point to note was that with or without riboflavin supplementation, protein levels of S82D were reduced to 15–20% of those of WT NQO1, whereas the effect of P187S was much larger ($\leq$5% of WT NQO1). Treatment with the proteasomal inhibitor MG-132 led to a 4.5- (for P187S) and 1.8-fold (for S82D) increase in protein levels, whereas this treatment had no effect on WT NQO1 (Figure S20). These results are consistent with enhanced proteasomal degradation of P187S compared to S82D, which is very slow for WT NQO1. Secondly, supplementation with Dic notably increased the protein levels of P187S (about 7-fold vs. Rib treatment), and this effect was weaker for S82D and WT NQO1 (about 3-fold and 1.8-fold vs. Rib treatment, respectively). Therefore, these results supported our hypothesis linking the difference in protein abundance (as a proxy for intracellular degradation) and the response to Dic treatment to the specific effects of these mutations on the stability of the CTD in different ligation states.

4. Conclusions

The multiple functional features displayed by human flavoproteins (involved in metabolic reactions, regulatory biomacromolecular interactions, transport to different subcellular locations, intracellular stability) can be modulated by ligand binding, disease-associated mutations and post-translational modifications. However, these complex structure-function relationships are generally not well undestood, in part due to the lack of high-resolution structures in most cases. In this work, we investigated these relationships using the human NQO1 as a model of multifunctional protein [16], and evaluated the effect of two mutations, the cancer-associated polymorphism P187S and the phosphomimetic mutation S82D on several functional traits. Our results revealed that the local destabilization of the protein structure caused by these two mutations can be transmitted to distant functional sites, and thus, affect different functional features to different extents, including enzyme catalysis, functional cooperativity, intracellular abundance (i.e.

stability) and pharmacological response to natural ligands. Since most of the human flavoproteins are associated with inherited diseases [2] and are potential targets of multiple post-translational events [10], the approach used in our work can be of general application to unravel these complex structure-function relationships in the human flavoproteome, in particular when high resolution structural information is unavailable or hard-to-get. In addition, the ability of mutations and post-translational modifications to affect multiple functional features in human flavoproteins also stress the importance of understanding the allosteric response to ligand binding, disease-associated and post-translational modifications in order to decipher their roles in physiological and pathological conditions.

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.redox.2021.102112.

References

[1] R. Martin-Vabulas, Ferroptosis-related flavoproteins: their function and stability, Int. J. Mol. Sci. 22 (2021) 430.

[2] W.D. Lienhart, V. Gudipati, P. Macheroux, The human flavoproteome, Arch. Biochem. Biophys. 535 (2013) 150–162.

[3] O. Schmidt, N. Pfanner, C. Meisinger, Mitochondrial protein import: from proteomics to functional mechanisms, Nat. Rev. Mol. Cell Biol. 11 (2010).

[4] B. Cautain, R. Hill, N. De Pedro, W. Link, Components and regulation of nuclear transport processes, FEBS J. 282 (2014) 445–462.

[5] A. Baker, T. Lanyon-Hogg, S.L. Warriner, Peroxisome protein import: a complex journey, Biochem. Soc. Trans. 44 (2016) 783–789.

[6] T.A. Rapoport, L. Li, E. Park, Structural and mechanistic insights into protein translocation, Annu. Rev. Cell Dev. Biol. 33 (2017) 369–390.

[7] A. Martinez-Limon, M. Alriquet, W.H. Lang, G. Calloni, I. Wittig, R.M. Vabulas, Recognition of enzymes lacking bound cofactor by protein quality control, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 12156–12161.

[8] A. Martínez-Limon, G. Calloni, R. Ernst, R. Martin-Vabulas, Flavin dependency undermines proteome stability, lipid metabolism and cellular proliferation during vitamin B2 deficiency, Cell Death Dis. 11 (2020) 725.

[9] D. Siegel, S. Bersie, P. Harris, A. Di Francesco, M. Armstrong, N. Reisdorph, M. Bernier, R. de Cabo, K. Fritz, D. Ross, A redox-mediated conformational change in NQO1 controls binding to microtubules and α-tubulin acetylation, Redox Biol 39 (2021).

[10] E. Medina-Carmona, B. Rizzuti, R. Martin-Escolano, J.L. Pacheco-Garcia, N. Mesa-Torres, J.L. Neira, R. Guzzi, A.L. Pey, Phosphorylation compromises FAD binding

and intracellular stability of wild-type and cancer-associated NQO1: insights into flavo-proteome stability, Int. J. Biol. Macromol. 125 (2019) 1275–1288.

[11] A.L. Pey, C.F. Megarity, D.J. Timson, FAD binding overcomes defects in activity and stability displayed by cancer-associated variants of human NQO1, Biochim. Biophys. Acta 1842 (2014) 2163–2173.

[12] E. Medina-Carmona, R.J. Palomino-Morales, J.E. Fuchs, E. Padín-Gonzalez, N. Mesa-Torres, E. Salido, D.J. Timson, A.L. Pey, Conformational dynamics is key to understanding loss-of-function of NQO1 cancer-associated polymorphisms and its correction by pharmacological ligands, Sci. Rep. 6 (2016) 20331.

[13] A.L. Pey, Biophysical and functional perturbation analyses at cancer-associated P187 and K240 sites of the multifunctional NADP(H):quinone oxidoreductase 1, Int. J. Biol. Macromol. 118 (2018) 1912–1923.

[14] J.L. Pacheco-García, , M. Cano-Muñoz, I. Sánchez-Ramos, E. Salido, A.L. Pey, Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein, J. Personalized Med. 10 (2020) 207.

[15] S.K. Beaver, N. Mesa-Torres, A.L. Pey, D.J. Timson, NQO1: a target for the treatment of cancer and neurological diseases, and a model to understand loss of function disease mechanisms, Biochim. Biophys. Acta Protein Proteonomics 1867 (2019) 663–676.

[16] D. Ross, D. Siegel, The diverse functionality of NQO1 and its roles in redox control, Redox Biology 41 (2021) 101950.

[17] E.T. Oh, J.W. Kim, J.M. Kim, S.J. Kim, J.S. Lee, S.S. Hong, J. Goodwin, R. J. Ruthenborg, M.G. Jung, H.J. Lee, C.H. Lee, E.S. Park, C. Kim, H.J. Park, NQO1 inhibits proteasome-mediated degradation of HIF-1alpha, Nat. Commun. 7 (2016) 13593.

[18] G. Asher, P. Tsvetkov, C. Kahana, Y. Shaul, A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73, Genes Dev. 19 (2005) 316–321.

[19] M. Faig, M.A. Bianchet, P. Talalay, S. Chen, S. Winski, D. Ross, L.M. Amzel, Structures of recombinant human and mouse NAD(P)H:quinone oxidoreductases: species comparison and structural changes with substrate binding and release, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 3177–3182.

[20] R. Li, M.A. Bianchet, P. Talalay, L.M. Amzel, The three-dimensional structure of NAD(P)H:quinone reductase, a flavoprotein involved in cancer chemoprotection and chemotherapy: mechanism of the two-electron reduction, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 8846–8850.

[21] E. Medina-Carmona, J.L. Neira, E. Salido, J.E. Fuchs, R. Palomino-Morales, D. J. Timson, A.L. Pey, Site-to-site interdomain communication may mediate different loss-of-function mechanisms in a cancer-associated NQO1 polymorphism, Sci. Rep. 7 (2017) 44352.

[22] W.D. Lienhart, V. Gudipati, M.K. Uhl, A. Binter, S.A. Pulido, R. Saf, K. Zangger, K. Gruber, P. Macheroux, Collapse of the native structure caused by a single amino acid exchange in human NAD(P)H:quinone oxidoreductase(1.), FEBS J. 281 (2014) 4691–4704.

[23] P. Vankova, E. Salido, D.J. Timson, P. Man, A.L. Pey, A dynamic core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer, Biomolecules 9 (2019) 728.

[24] E. Anoz-Carbonell, D.J. Timson, A.L. Pey, M. Medina, The catalytic cycle of the antioxidant and cancer-associated human NQO1 enzyme: hydride transfer, conformational dynamics and functional cooperativity, Antioxidants 9 (2020).

[25] G. Asher, O. Dym, P. Tsvetkov, J. Adler, Y. Shaul, The crystal structure of NAD(P)H quinone oxidoreductase 1 in complex with its potent inhibitor dicoumarol, Biochemistry 45 (2006) 6372–6378.

[26] D. Siegel, A. Anwar, S.L. Winski, J.K. Kepa, K.L. Zolman, D. Ross, Rapid polyubiquitination and proteasomal degradation of a mutant form of NAD(P)H: quinone oxidoreductase 1, Mol. Pharmacol. 59 (2001) 263–268.

[27] C.F. Megarity, D.J. Timson, Cancer-associated variants of human NQO1: impacts on inhibitor binding and cooperativity, Biosci. Rep. 39 (2019).

[28] B. Lajin, A. Alachkar, The NQO1 polymorphism C609T (Pro187Ser) and cancer susceptibility: a comprehensive meta-analysis, Br. J. Canc. 109 (2013) 1325–1337.

[29] Z. Zhang, D.L. Smith, Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation, Protein Sci. 2 (1993) 522–531.

[30] H. Wako, N. Saitô, Statistical mechanical theory of the protein conformation. II. Folding pathway for protein, J. Phys. Soc. Jpn. 44 (44) (1978) 1939–1945.

[31] V. Muñoz, W.A. Eaton, A simple model for calculating the kinetics of protein folding from three-dimensional structures, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 11311–11316.

[32] S. Gopi, A. Aranganathan, A.N. Naganathan, Thermodynamics and folding landscapes of large proteins from a statistical mechanical model, Current Research in Structural Biology 1 (2019) 6–12.

[33] A.N. Naganathan, Predictions from an ising-like statistical mechanical model on the dynamic and thermodynamic effects of protein surface electrostatics, J. Chem. Theor. Comput. 8 (2012) 4646–4656.

[34] A. Sanchez-Azqueta, D.L. Catalano-Dupuy, A. Lopez-Rivero, M.L. Tondo, E. G. Orellano, E.A. Ceccarelli, M. Medina, Dynamics of the active site architecture in plant-type ferredoxin-NADP(+) reductases catalytic complexes, Biochim. Biophys. Acta 1837 (2014) 1730–1738.

[35] T.P. Schrank, D.W. Bolen, V.J. Hilser, Rational modulation of conformational fluctuations in adenylate kinase reveals a local unfolding mechanism for allostery and functional adaptation in proteins, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 16984–16989.

[36] I. Luque, E. Freire, Structural parameterization of the binding enthalpy of small ligands, Proteins 49 (2002) 181–190.

[37] E. Medina-Carmona, J.E. Fuchs, J.A. Gavira, N. Mesa-Torres, J.L. Neira, E. Salido, R. Palomino-Morales, M. Burgos, D.J. Timson, A.L. Pey, Enhanced vulnerability of human proteins towards disease-associated inactivation through divergent evolution, Hum. Mol. Genet. 26 (2017) 3531–3544.

[38] S. Vega, O. Abian, A. Velazquez-Campoy, On the link between conformational changes, ligand binding and heat capacity, Biochim. Biophys. Acta 1860 (2016) 868–878.

[39] A.D. Robertson, K.P. Murphy, Protein structure and the energetics of protein stability, Chem. Rev. 97 (1997) 1251–1268.

[40] R. Claveria-Gimeno, A. Velazquez-Campoy, A.L. Pey, Thermodynamics of cooperative binding of FAD to human NQO1: implications to understanding cofactor-dependent function and stability of the flavoproteome, Arch. Biochem. Biophys. 636 (2017) 17–27.

[41] S. Chen, P.S. Deng, J.M. Bailey, K.M. Swiderek, A two-domain structure for the two subunits of NAD(P)H:quinone acceptor oxidoreductase, Protein Sci. 3 (1994) 51–57.

[42] O. Moscovitz, P. Tsvetkov, N. Hazan, I. Michaelevski, H. Keisar, G. Ben-Nissan, Y. Shaul, M. Sharon, A mutually inhibitory feedback loop between the 20S proteasome and its regulator, NQO1, Mol. Cell. 47 (2012) 76–86.

[43] W.D. Lienhart, E. Strandback, V. Gudipati, K. Koch, A. Binter, M.K. Uhl, D. M. Rantasa, B. Bourgeois, T. Madl, K. Zangger, K. Gruber, P. Macheroux, Catalytic competence, structure and stability of the cancer-associated R139W variant of the human NAD(P)H:quinone oxidoreductase 1 (NQO1), FEBS J. 284 (2017) 1233–1245.

[44] I.G. Munoz, B. Morel, E. Medina-Carmona, A.L. Pey, A mechanism for cancer-associated inactivation of NQO1 due to P187S and its reactivation by the consensus mutation H80R, FEBS Lett. 591 (2017) 2826–2835.

[45] C.F. Megarity, H. Abdel-Bettley, M.C. Caraher, K.A. Scott, W. Ra, T.A. Jowitt, A. Gutierrez, R.A. Bryce, K.A. Nolan, I.J. Stratford, D. Timson, Negative cooperativity in NADP(H) quinone oxidoreductase 1 (NQO1), Chembiochem 20 (2019) 2841–2849, 2841–2849.

[46] F.K. Yoshimoto, F. Zhou, H.M. Peng, D. Stidd, J.A. Yoshimoto, K.K. Sharma, S. Matthew, R.J. Auchus, Minor activities and transition state properties of the human steroid hydroxylases cytochromes P450c17 and P450c21, from reactions observed with deuterium-labeled substrates, Biochemistry 51 (2012) 7064–7077.

[47] J.P. Klinman, Moving through barriers in science and life, Annu. Rev. Biochem. 88 (2019) 1–24.

[48] Z.D. Nagel, J.P. Klinman, Update 1 of: tunneling and dynamics in enzymatic hydride transfer, Chem. Rev. 110 (2010) PR41–67.

[49] C. Park, S. Marqusee, Probing the high energy states in proteins by proteolysis, J. Mol. Biol. 343 (2004) 1467–1476.

MATERIAL SUPLEMENTARIO DE LA PUBLICACIÓN 3

Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue.

Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue.

Juan Luis Pacheco-Garcia, Ernesto Anoz-Carbonell, Pavla Vankova, Adithi Kannan, Rogelio Palomino-Morales, Noel Mesa-Torres, Eduardo Salido, Petr Man, Milagros Medina, Athi N. Naganathan and Angel L. Pey.

Structure-based protein stability analyses

Structure-based protein stability analyses were performed using the FoldX 4 force field [1]. Protein structures were first relaxed to reduce the global energy by using the *RepairPDB* command. Mutations or phosphorylated residues were introduced in the structures (see Table S1 for PDB codes and list of mutated sites) by using the *BuildModel* command with parameters set as 298 K, pH 7, ionic strength 50 mM and VdW design 2.

The free energy changes ($\Delta\Delta G$) obtained represent the difference in folding free energy between the mutant and the reference state (i.e. the WT protein) and were normalized by protein monomer. The accessible surface area (ASA) was evaluated at the residue level using these crystal structures and the software GetArea [2] (http://curie.utmb.edu/getarea.html). Statistical analyses of stability changes between different groups of variants were performed using the language and environment for statistical computing R (V3.3.2) and graphs were generated using the GGPLOT2 package [3].

A recent report has evaluated the effect of flavin starvation in the proteasomal degradation on 66 flavoproteins in cultured cells, thus accounting the behavior of 2/3 of the entire flavoproteome [4]. From this set, we selected 17 proteins to carry out structure-based stability calculations due to disease-associated mutations and phosphorylation events since they all fulfilled all the following criteria: i) at least one suitable crystal structure was available (preferably, and in most of the cases, of the full-length protein; these were retrieved from the Protein Data Bank; https://www.rcsb.org/); ii) phosphorylation sites (at S, T and/or Y) were already reported by large-scale proteomic studies (these were mined from PhosphoSitePlus®; https://www.phosphosite.org); iii) disease-associated mutations were reported (either in OMIM, https://www.omim.org/, or ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/databases) and we selected them since they cover a wide range of mutations (particularly at the level of the physicochemical

1

changes caused in the side-chain upon mutation and solvent exposure in the structure)(Table S2). Using this set, we analyzed the effects of 75 disease-associated mutations and phosphorylation at 152 sites (including phosphorylation of S, T or Y) on the conformational stability of seventeen human flavoproteins (Table S2). Disease-associated mutations destabilized the protein (in terms of folding free energy) on average by 4.0 kcal·mol^{-1} and phosphorylation events by 3.0 kcal·mol^{-1} (with median values of 2.5 and 1.6 kcal·mol^{-1}, respectively) (Figure S1A). The apparently lower destabilizing effect of phosphorylation *vs.* disease-associated mutations seemed to particularly arise from the lower penalty associated with phospho-serine sites. Among the different sets of phosphorylation sites, phospho-S showed an average penalty of 1.8 kcal·mol^{-1} (median of 0.7), whereas phospho-T and phospho-Y had averages of 4.1 and 3.7 kcal·mol^{-1}, respectively (medians of 3.1 and 2.1 kcal·mol^{-1}). These sets of sites displayed a multi-modal profile consistent with the global mutational-effects on protein thermodynamic stability [5], and thus likely arose from the different solvent accessibility of the mutated sites (Figure S1B). Accordingly, when the stability effects due to phosphorylation and disease-associated mutations were clustered in two groups depending on site solvent accessibility (buried *vs.* solvent-exposed), it was evident that modification of buried sites was generally associated with a much larger destabilizing effect (Figure S1C). Overall, these analyses supported that destabilization due to disease-associated mutations and phosphorylation followed similar patterns and might be of similar magnitude at the flavoproteome scale.

Figure S1. Structure-based stability analyses on the effects of disease-causing mutations and phosphorylation events in human flavoproteins. A) Box plots for the stability effects of 75 disease-causing point mutations and 152 phosphorylation events in seventeen human flavoproteins (see Table S2 for a detailed list). The plot also showed the effect of phosphorylation events distinguishing between serine, threonine and tyrosine sites (phospho-S, phospho-T and phospho-Y). Boxes show inter-quartile ranges and thick horizontal lines show median values. B) Distribution of stability effects displayed with violin plots. Data shown are those in panel A without outliers. C) Box plots for the stability effects of data shown in panel B and grouped according to the solvent accessibility of the sites (buried, ≤10% of ASA, white; solvent-exposed, >10% of ASA, grey). Closed circles in panels A and C show outlier values beyond 1.5 times the inter-quartile ranges.
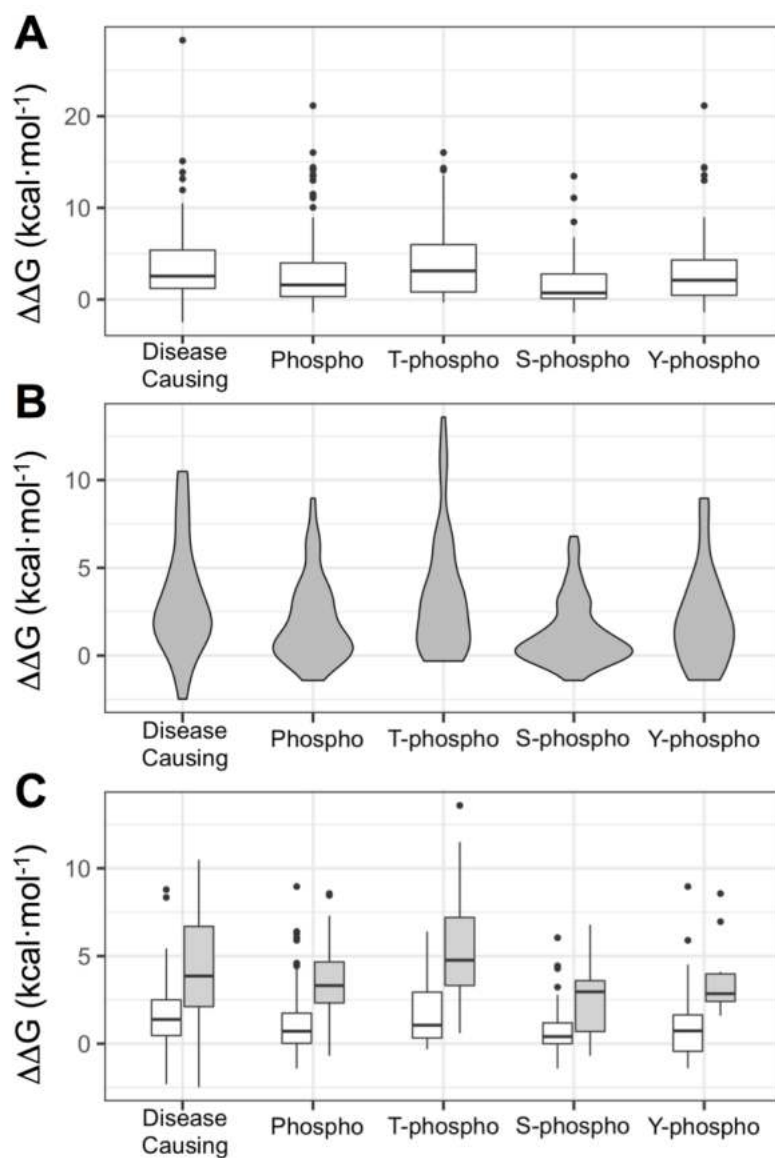
Figure S2. Coverage map for all three studied NQO1 forms - wild-type (grey), **S82D (red)** and **P187S (green)** variants. Sites of mutations are highlighted above the sequence. Only peptides providing HDX data are shown. Peptide sets are virtually identical, small differences are just between WT/S82D and P187S in the region covering residue 187.
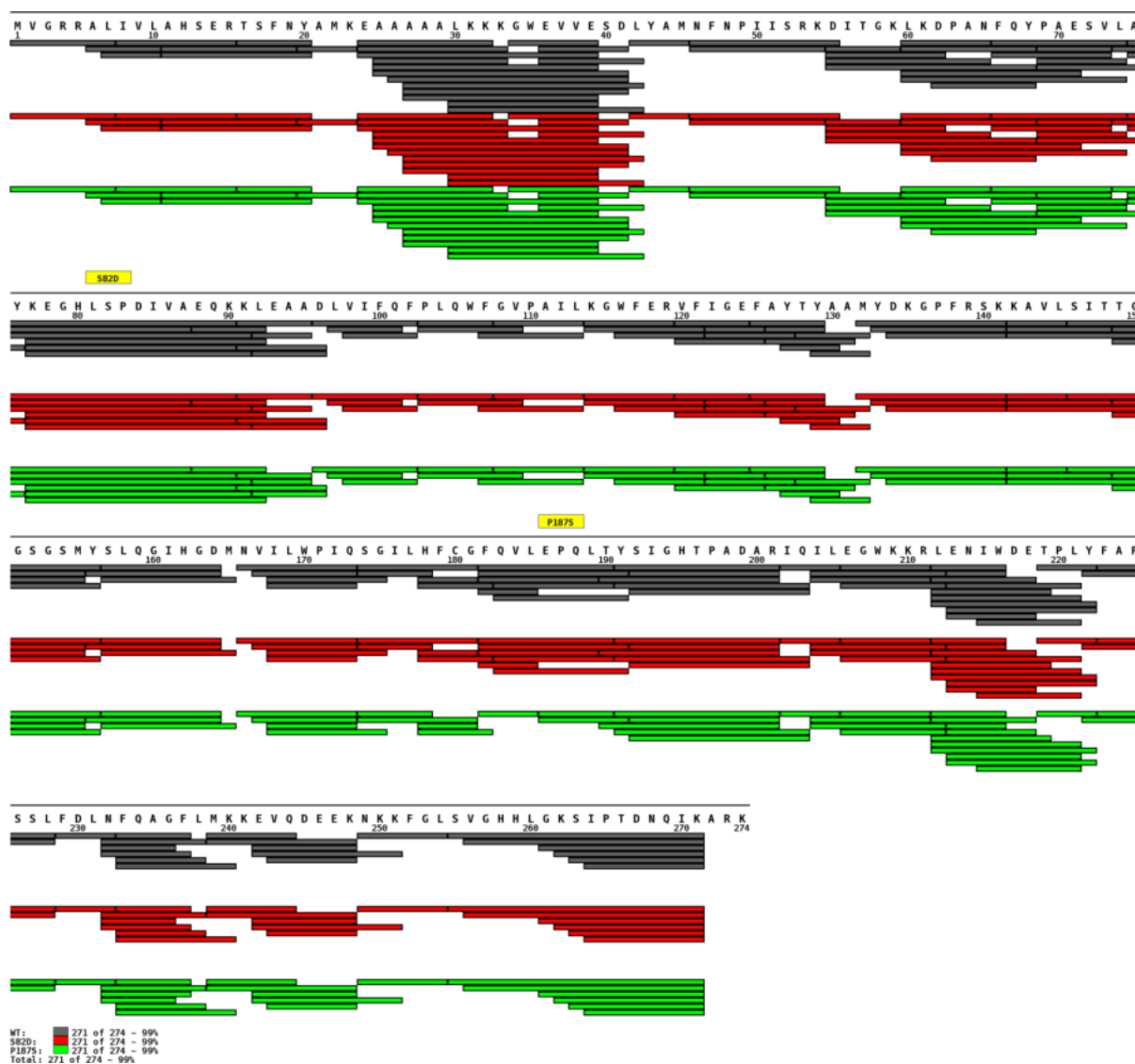
Figure S3. Deuterium uptake plots for all detected peptides of NQO1$_{apo}$ state. Wild-type (black), S82D (red) and P187S (green).

[92-95 (1)]   [92-96 (1)]   [96-102 (1)]   [97-101 (1)]   [98-102 (1)]

[103-107 (1)]   [103-109 (1-2)]   [107-113 (1)]   [108-113 (1)]   [114-119 (1)]

[114-121 (1-3)]   [116-121 (1-2)]   [120-124 (1)]   [120-125 (1)]   [122-125 (1)]

[122-127 (1)]   [125-129 (1)]   [126-129 (1)]   [126-131 (1)]   [127-130 (1)]

[128-132 (1)]   [129-132 (1)]   [132-141 (2-3)]   [133-141 (1-3)]   [134-141 (1-3)]

[142-145 (1)]   [142-148 (1)]   [142-155 (1-2)]   [146-156 (1-2)]   [149-155 (1)]

[149-156 (1)]   [156-164 (1-2)]   [157-164 (1-2)]   [157-165 (1-2)]   [166-173 (1-2)]

[167-173 (1)]   [168-173 (1)]   [168-175 (1-2)]   [174-178 (1)]   [174-181 (1-2)]

[178-181 (1)]   [178-182 (1)]   [182-185 (1)]   [182-189 (1-2)]   [182-190 (1-2)]

6

Figure showing a grid of small plots of Exchanged D's versus log time [s] for various peptide fragments: [92-95 (1)], [92-96 (1)], [96-102 (1)], [97-101 (1)], [98-102 (1)], [103-107 (1)], [103-109 (1-2)], [107-113 (1)], [108-113 (1)], [114-119 (1)], [114-121 (1-3)], [116-121 (1-2)], [120-124 (1)], [120-125 (1)], [122-125 (1)], [122-127 (1)], [125-129 (1)], [126-129 (1)], [126-131 (1)], [127-130 (1)], [128-132 (1)], [129-132 (1)], [132-141 (2-3)], [133-141 (1-3)], [134-141 (1-3)], [142-145 (1)], [142-148 (1)], [142-155 (1-2)], [146-156 (1-2)], [149-155 (1)], [149-156 (1)], [156-164 (1-2)], [157-164 (1-2)], [157-165 (1-2)], [166-173 (1-2)], [167-173 (1)], [168-173 (1)], [168-175 (1-2)], [174-178 (1)], [174-181 (1-2)], [178-181 (1)], [178-182 (1)], [182-185 (1)], [182-189 (1-2)], [182-190 (1-2)].

Figure S4. Isotopic profiles for selected peptides derived from NQO1$_{apo}$ (16-20, 42-45, 69- 73, 108-113, 157-164, 242-248 – deuterium uptake curves highlighted by a blue box in the set of plots above) corresponding to regions shown in the Figure 2C ofthe main text. Isotopic profiles for the individual time points of HDX kinetics for peptides: **(A)** 16-20 (TSFNY, theoretical *m/z* 631.2722 – non-deuterated, 1+) derived from the NQO1$_{apo}$ state; **(B)** 42-45 (LYAM, theoretical *m/z* 497.2428 – non-deuterated, 1+) derived from the NQO1$_{apo}$ state; **(C)** 69-73 (PAESV, theoretical *m/z* 502.2508 – non-deuterated, 1+) derived from the NQO1$_{apo}$ state; **(D)** 108-113 (GVPAIL, theoretical *m/z* 569.3657 – non-deuterated, 1+) derived from the NQO1$_{apo}$ state; **(E)** 157-164 (SLQGIHGD, theoretical *m/z* 826.4054 – non-deuterated, 1+) derived from the NQO1$_{apo}$ state; **(F)** 242-248 (EVQDEEK, theoretical *m/z* 876.3945 – non-deuterated, 1+) derived from the NQO1$_{apo}$ state.

In all cases the graphs are ordered from top to bottom – ND (non-deuterated), 10s, 30s, 2min, 5min, 20min, 1h, 3h and FD (fully-deuterated). Left column – wild-type (black), middle column – S82D (red), right column - P187S (green).

S4A

S4B

S4C

S4D

S4E

157-164

S4F

242-248
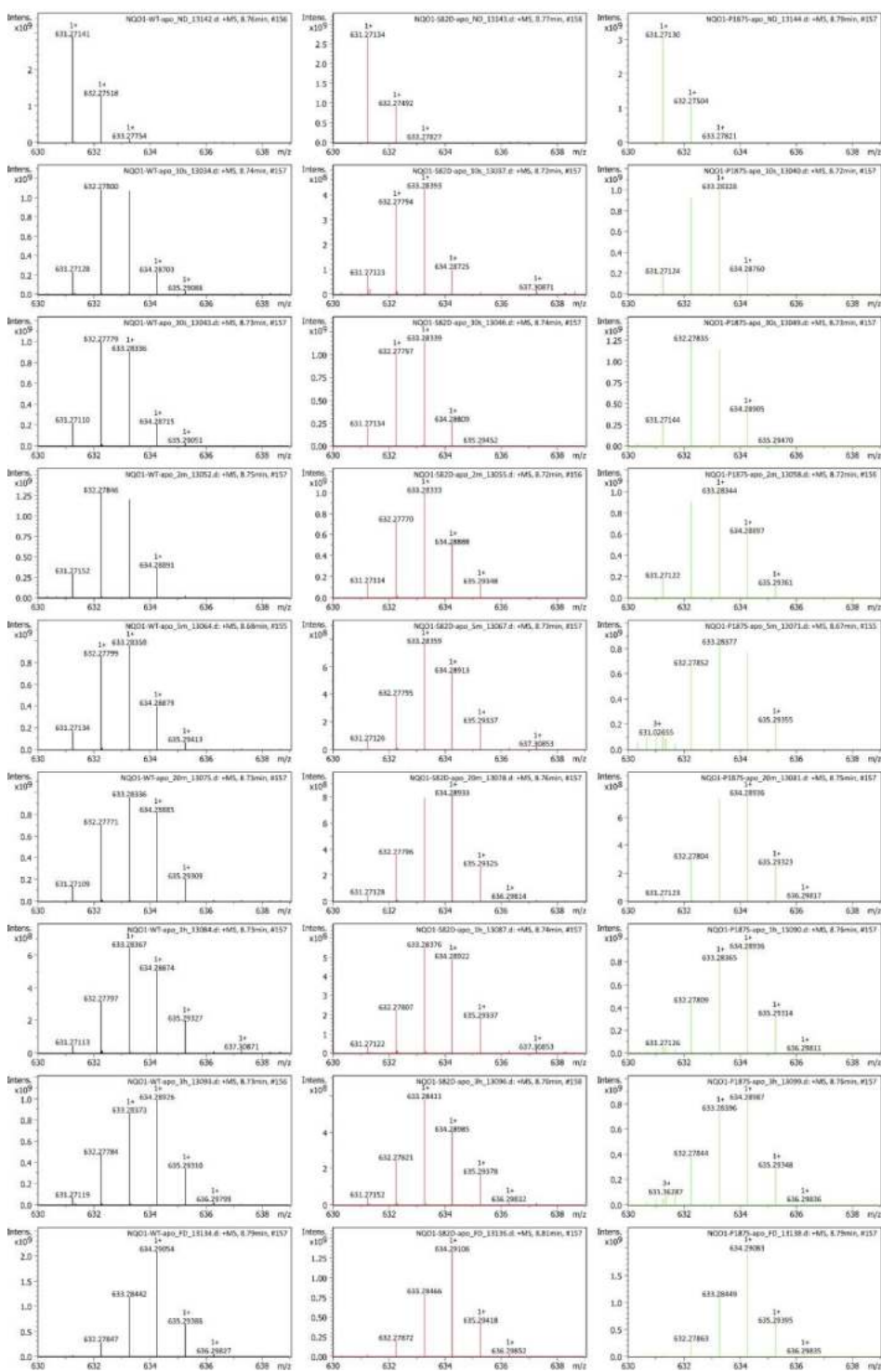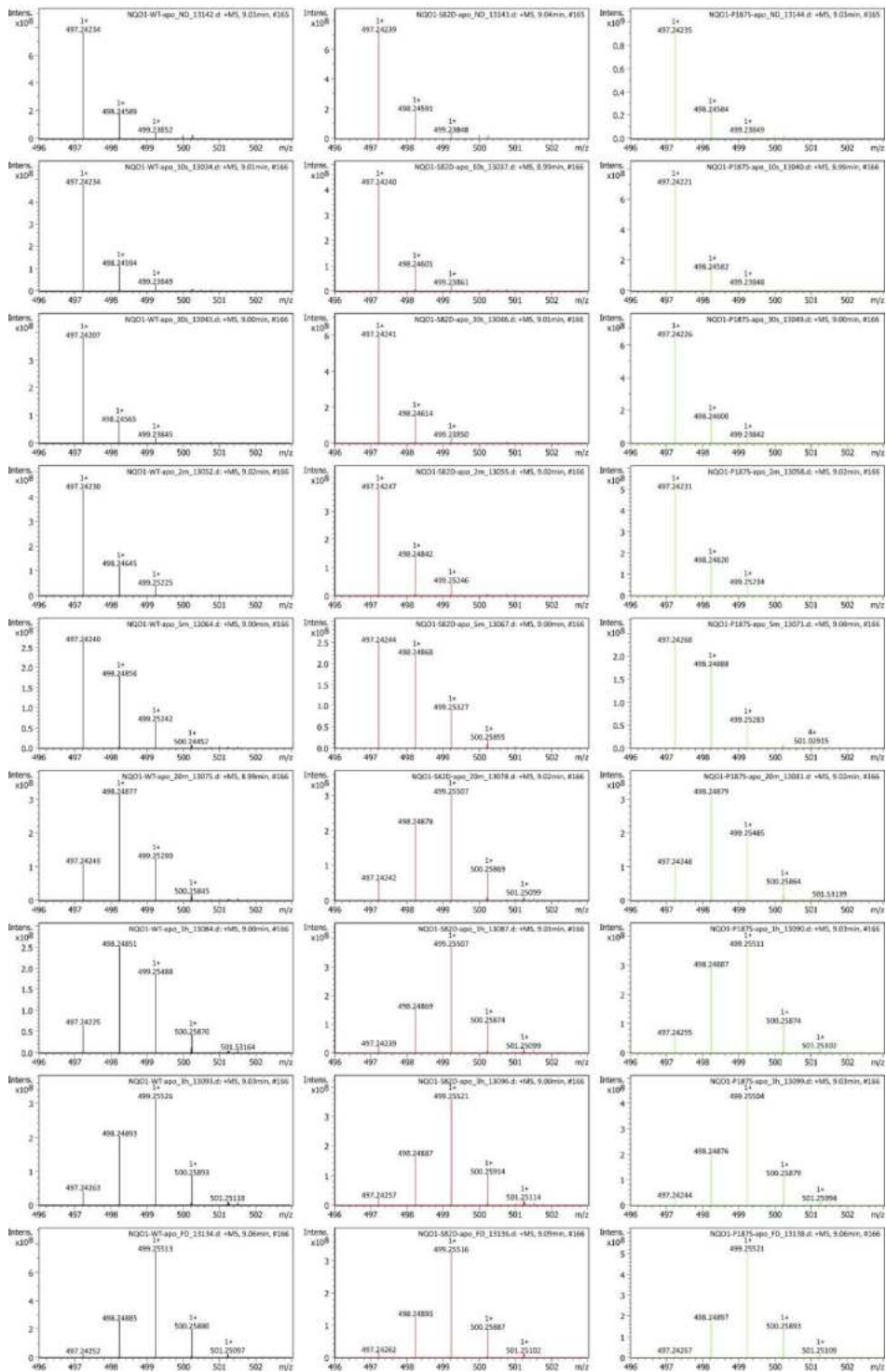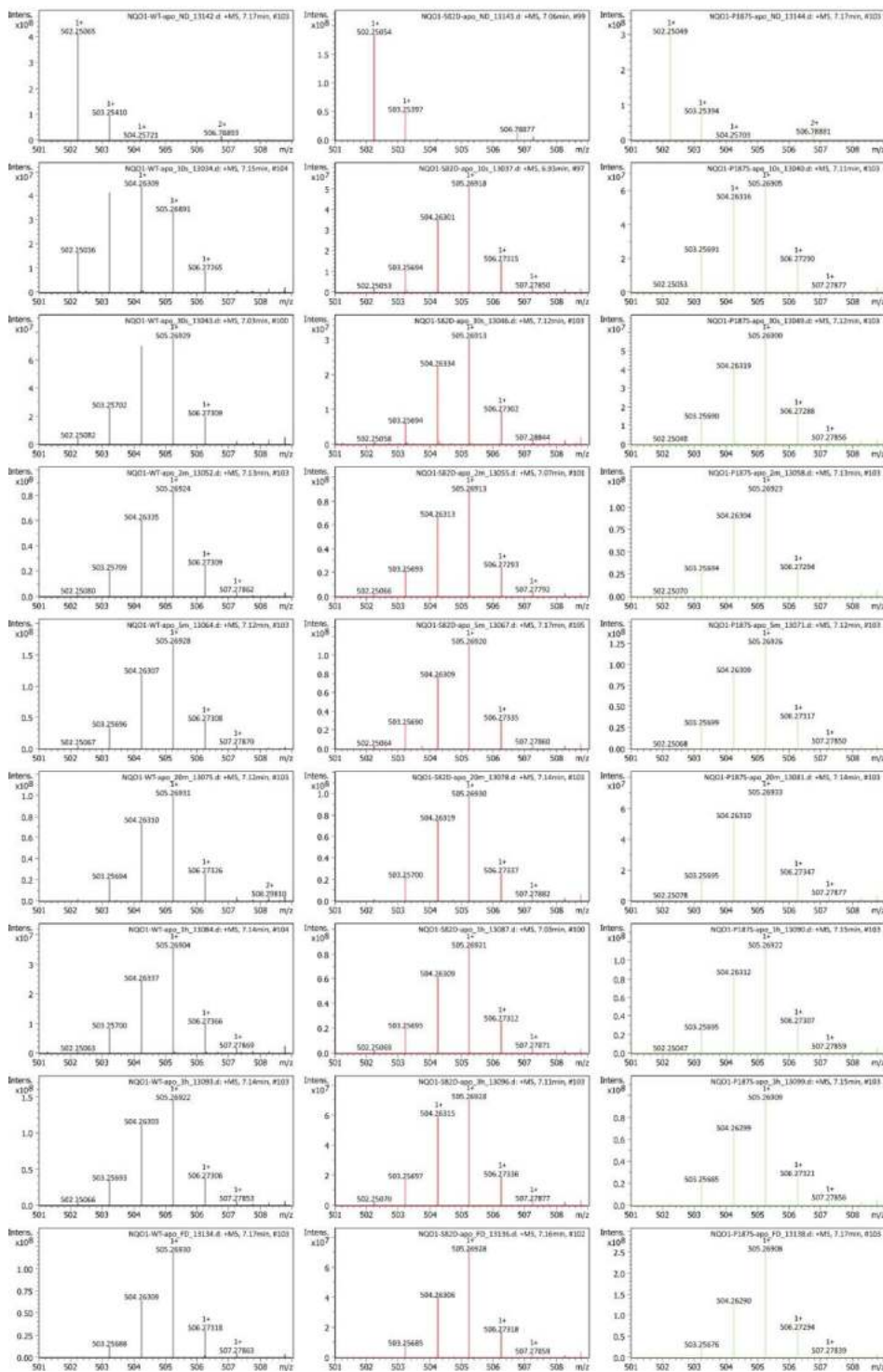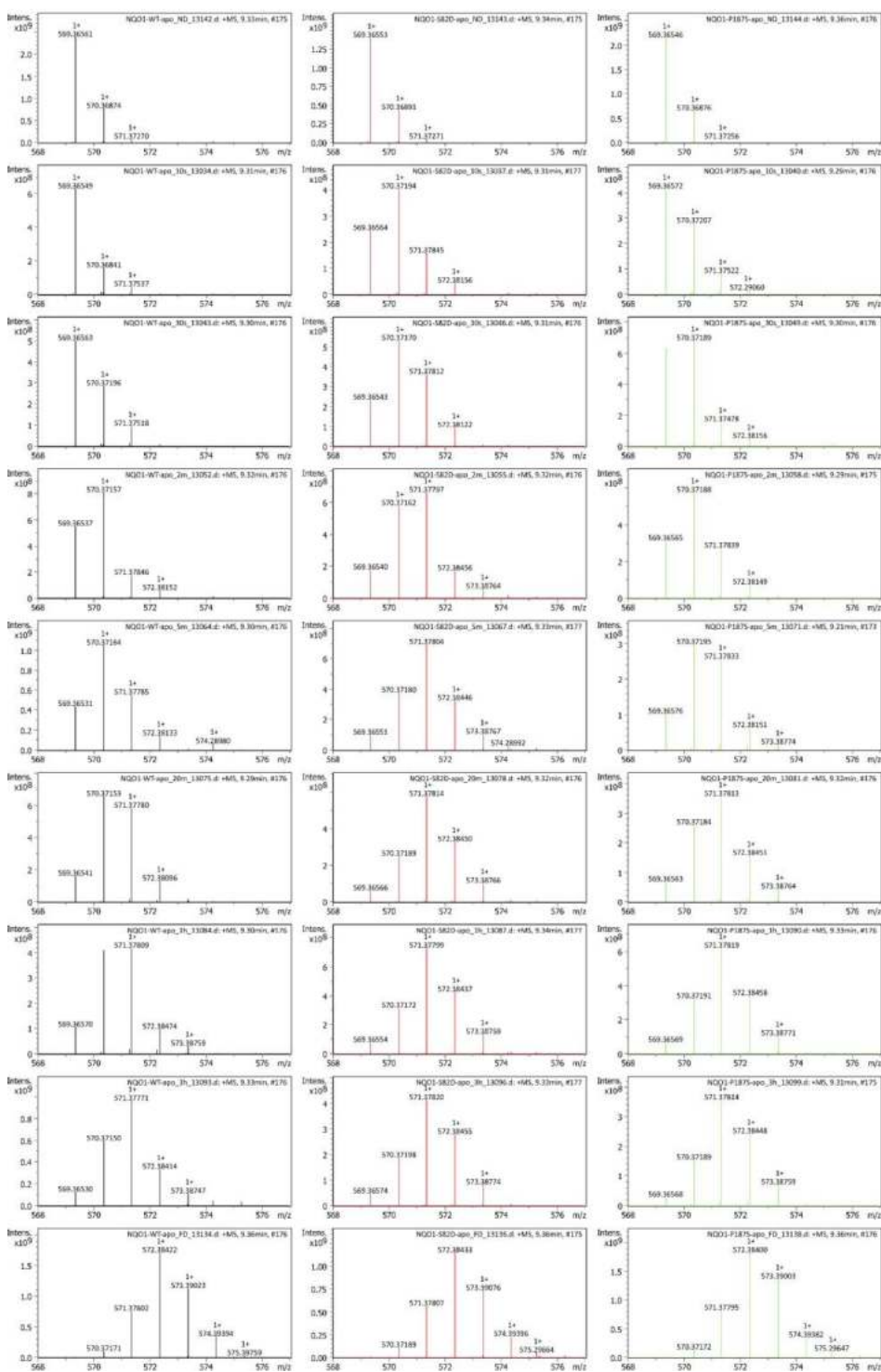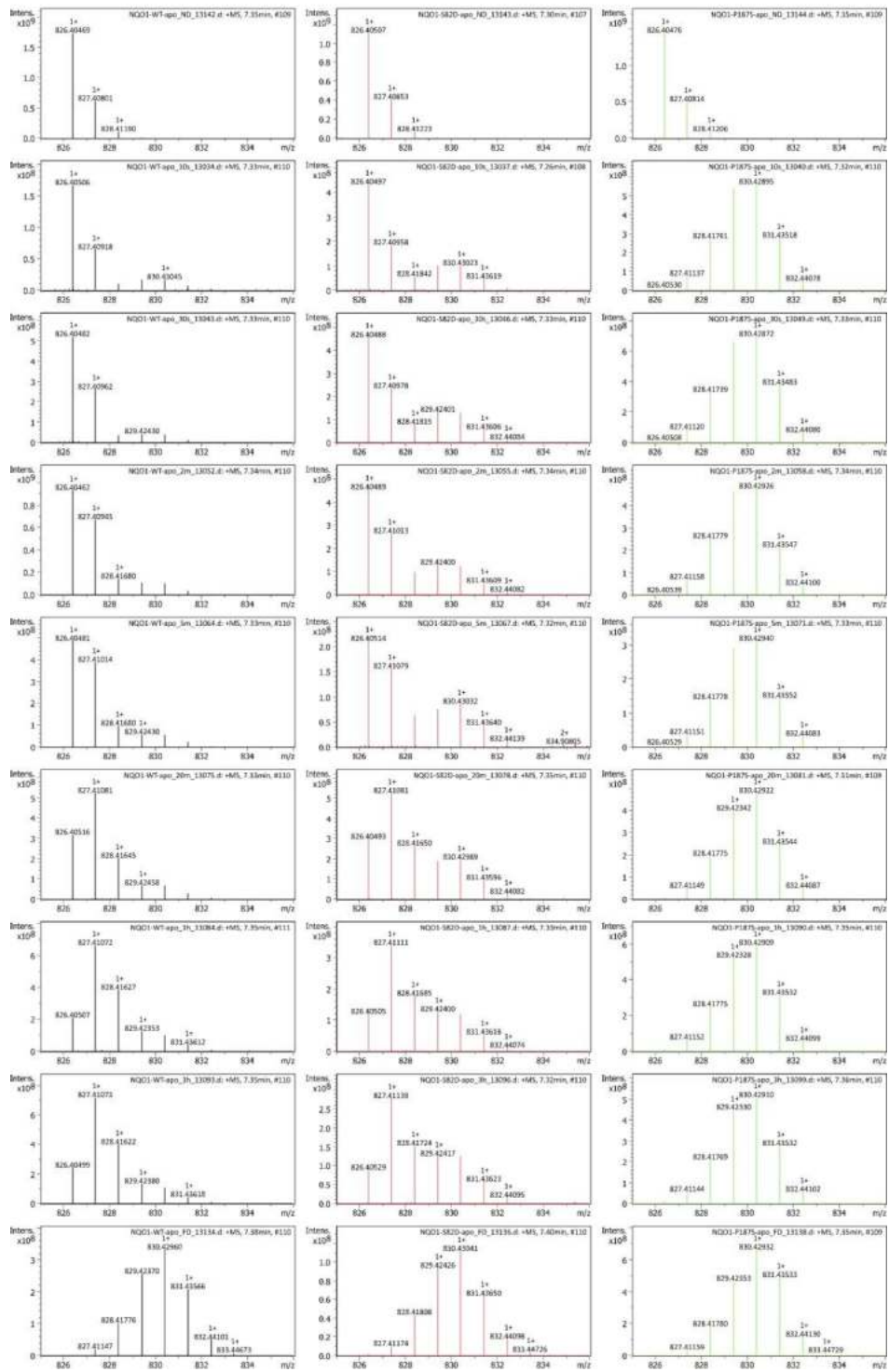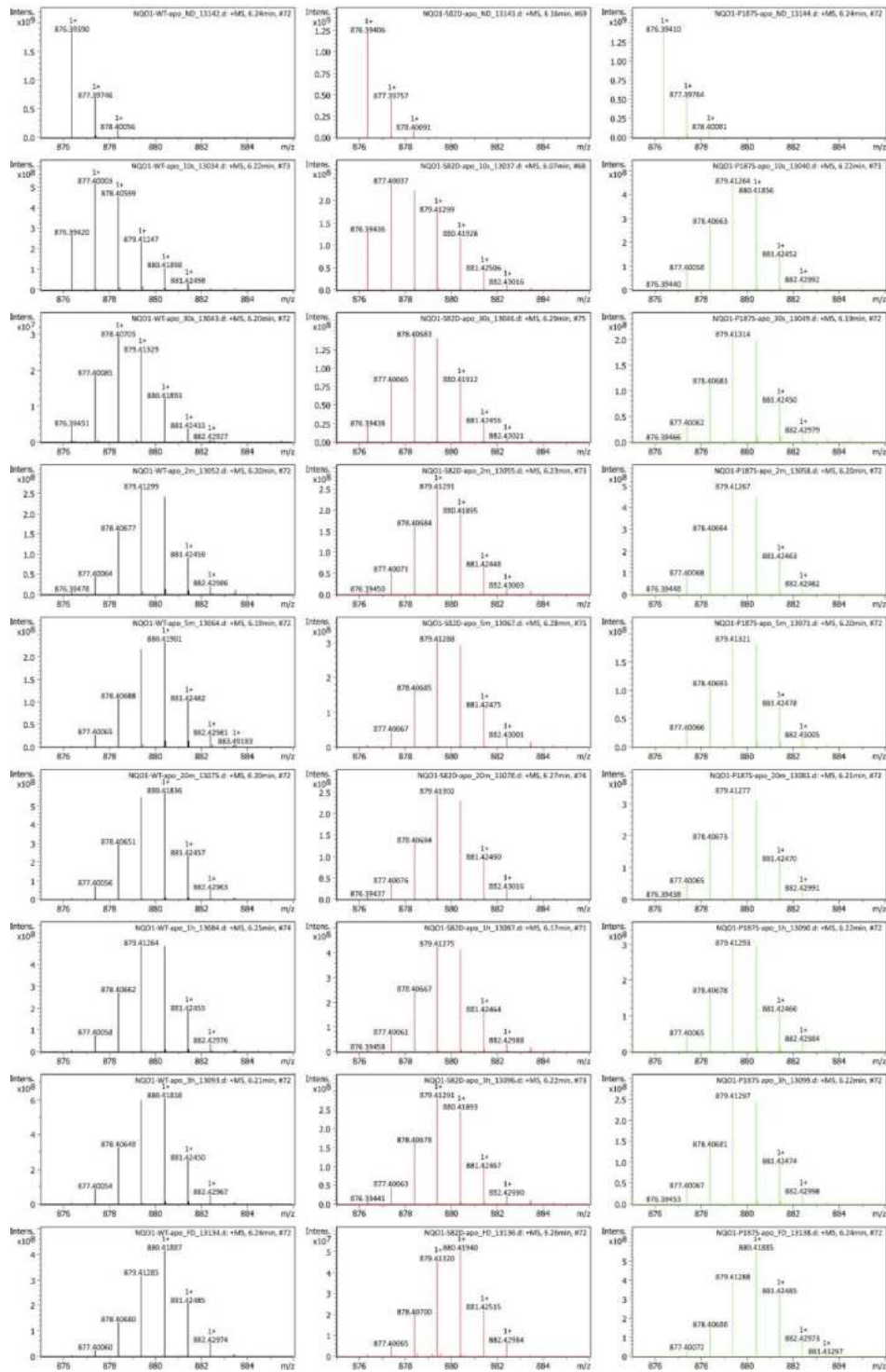
Figure S5. Deuterium uptake plots for all detected peptides of **NQO1holo** state. Wild-type (**black**), S82D (**red**) and P187S (**green**).
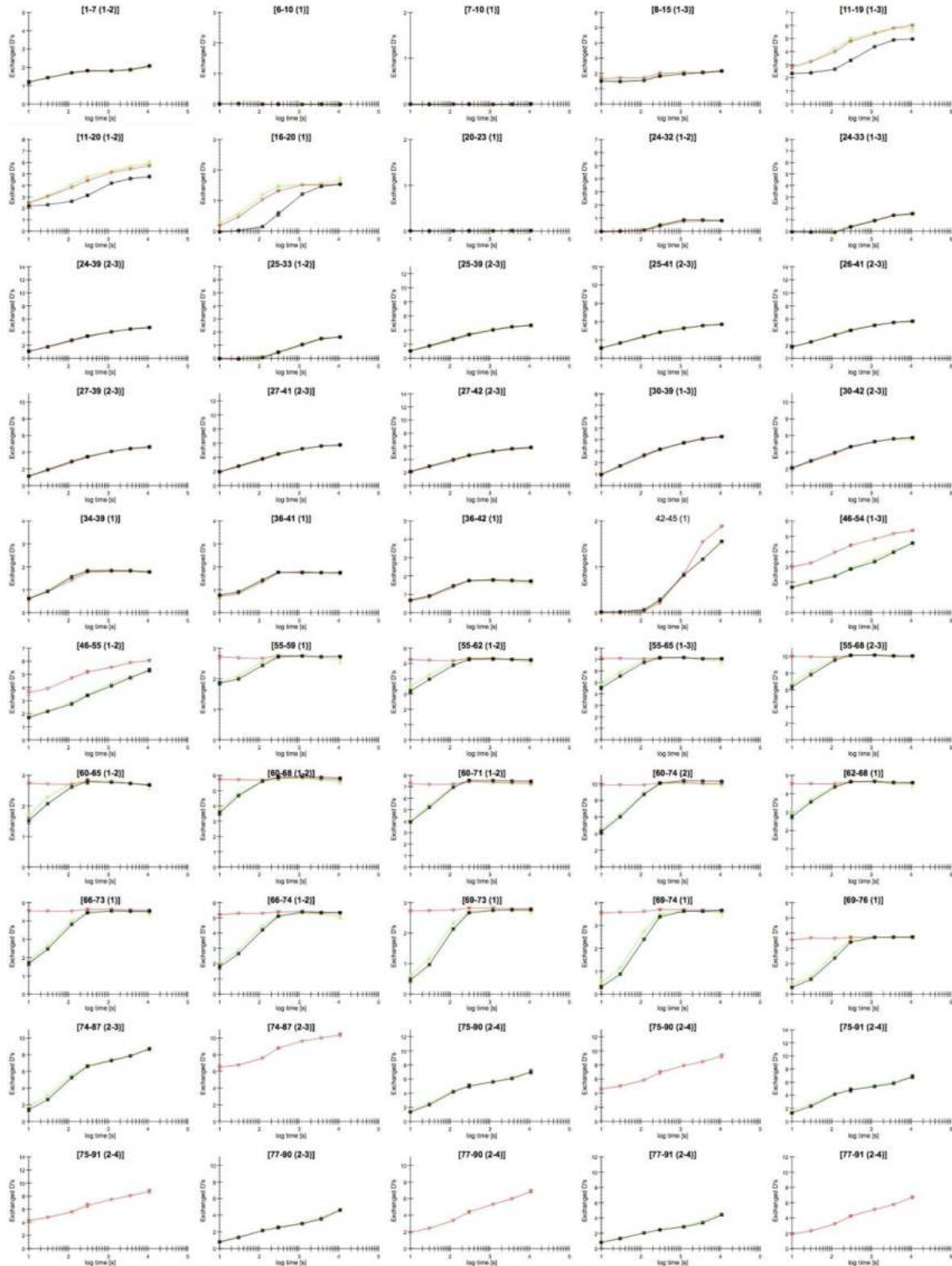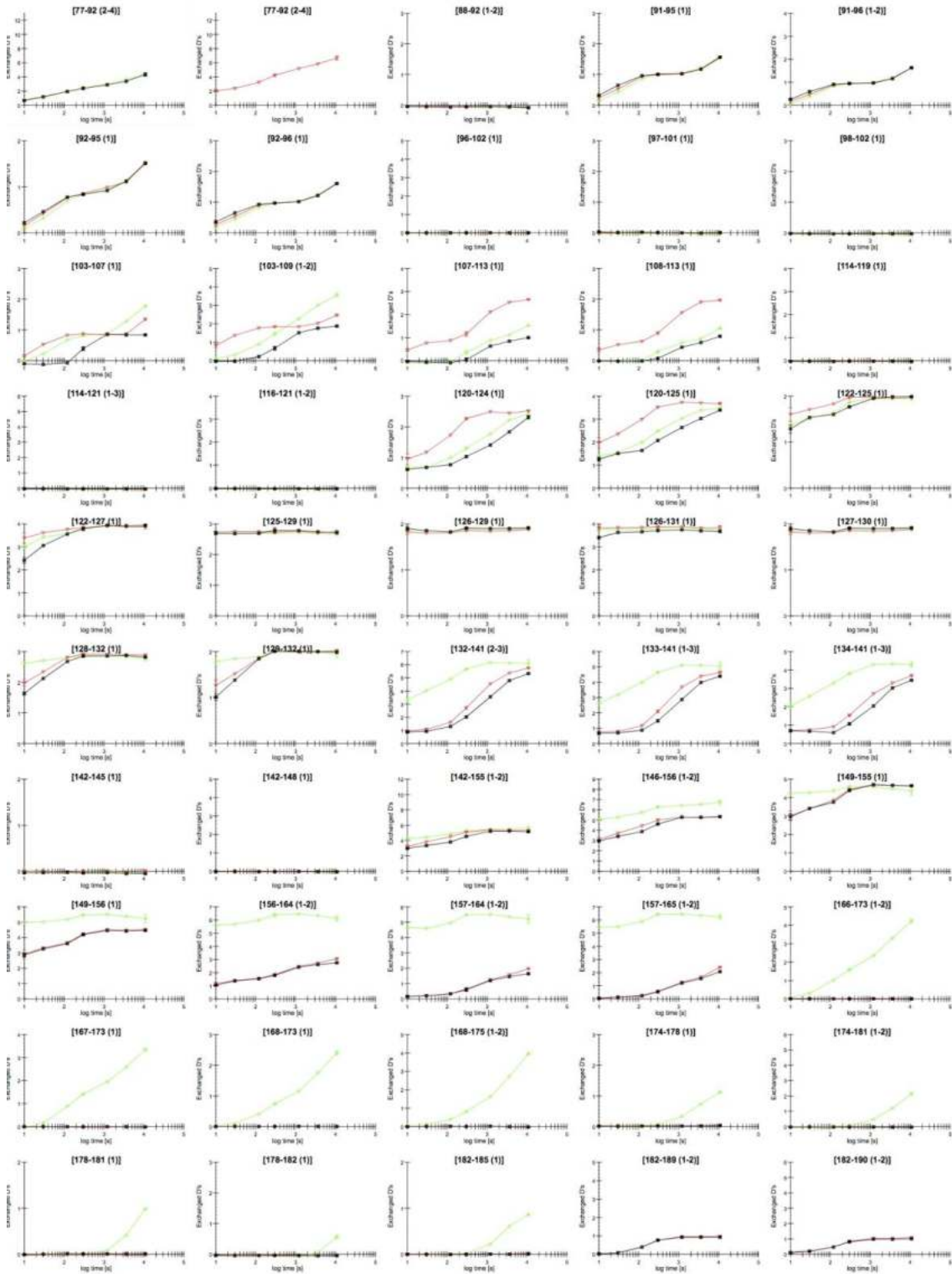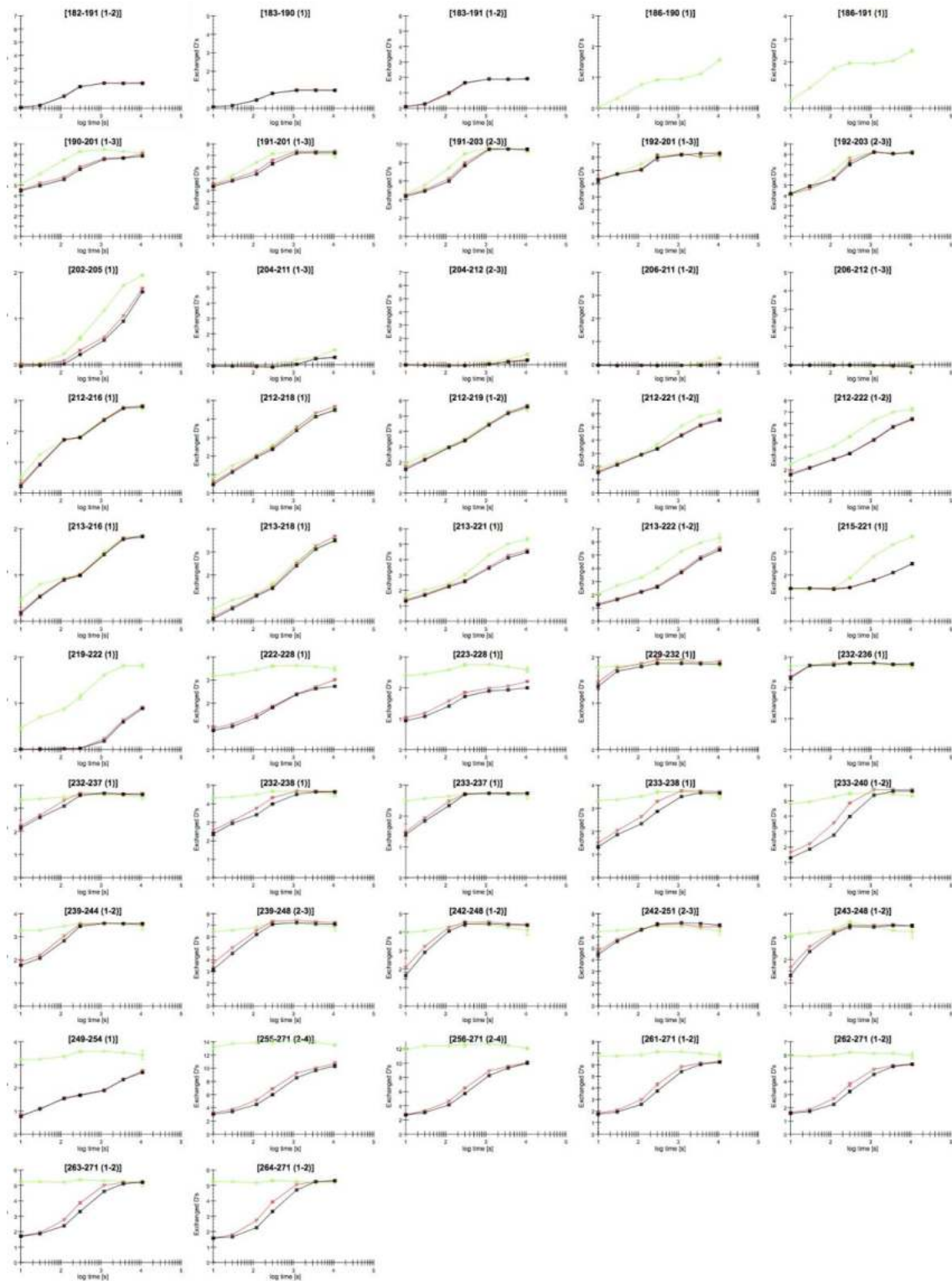
Figure S6. Deuterium uptake plots for all detected peptides of NQO1$_{dic}$ state. Wild-type(black), S82D (red) and P187S (green).
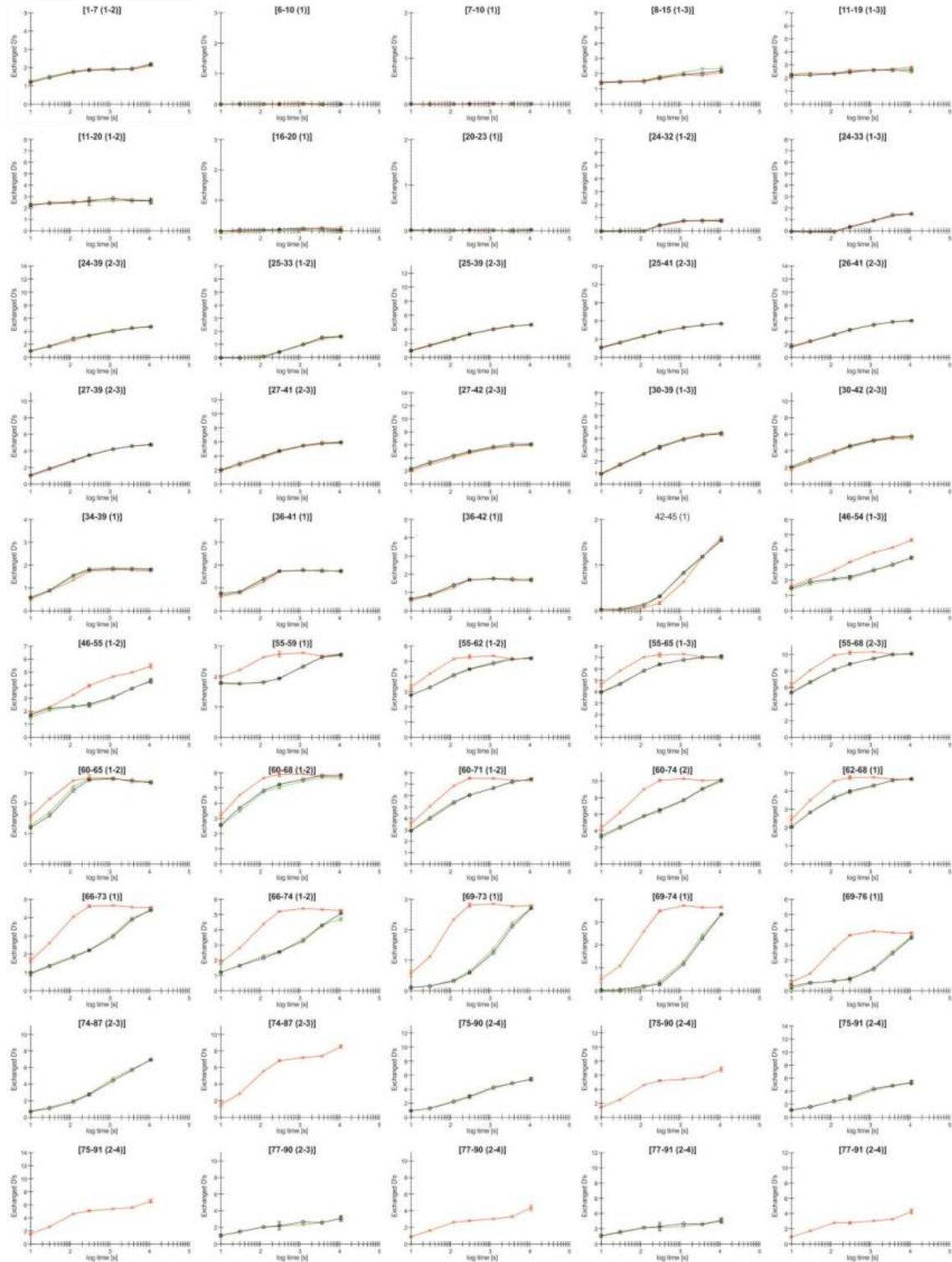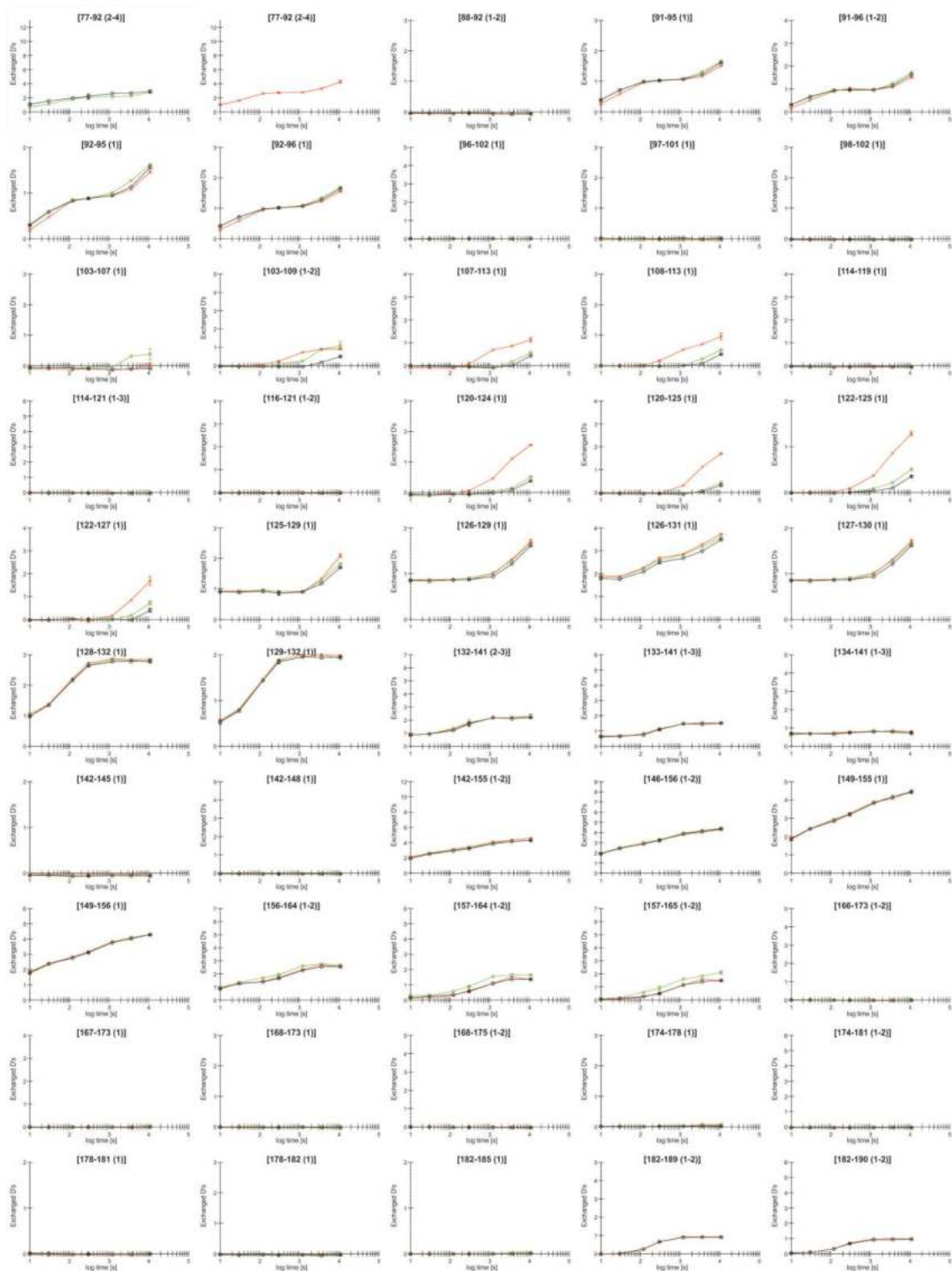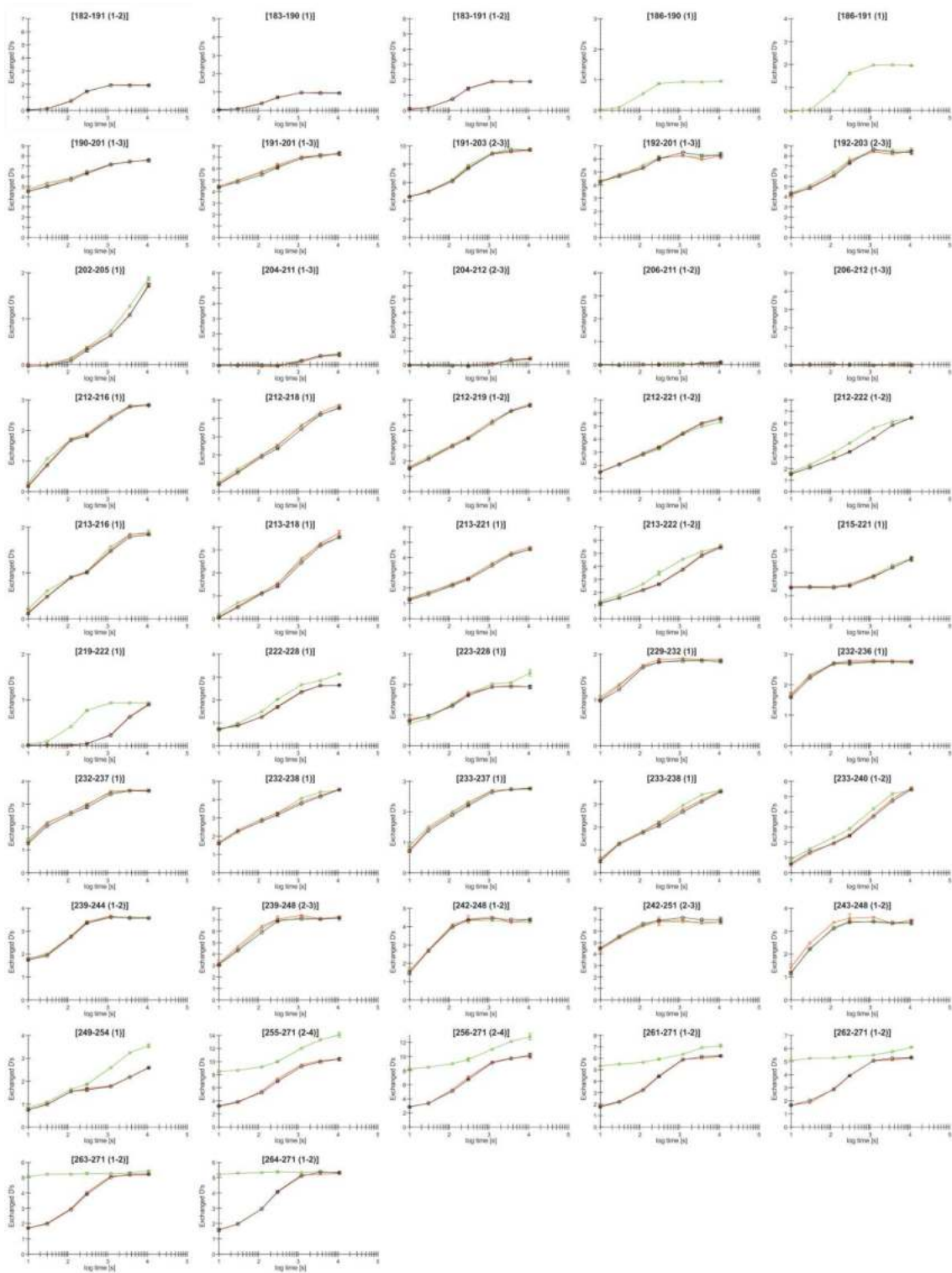
Figure S7. Free energy calculations of the conformational ensemble of NQO1$_{apo}$ WT.
A) Plot of the one-dimensional free energy profile as a function of the number of structured blocks (the reaction coordinate, RC) calculated at pH 7.0 and 298 K. At least two intermediate states (I$_1$ and I$_2$) apart from the native state (N) were identified. B) Representation of different combination of blocks (i.e. segments) structured at different points on the folding free energy profile (colored from 0 to 1 and in blue to red, with red representing fully folded). Intermediates can be seen by the vertical lines spanning the two panels.
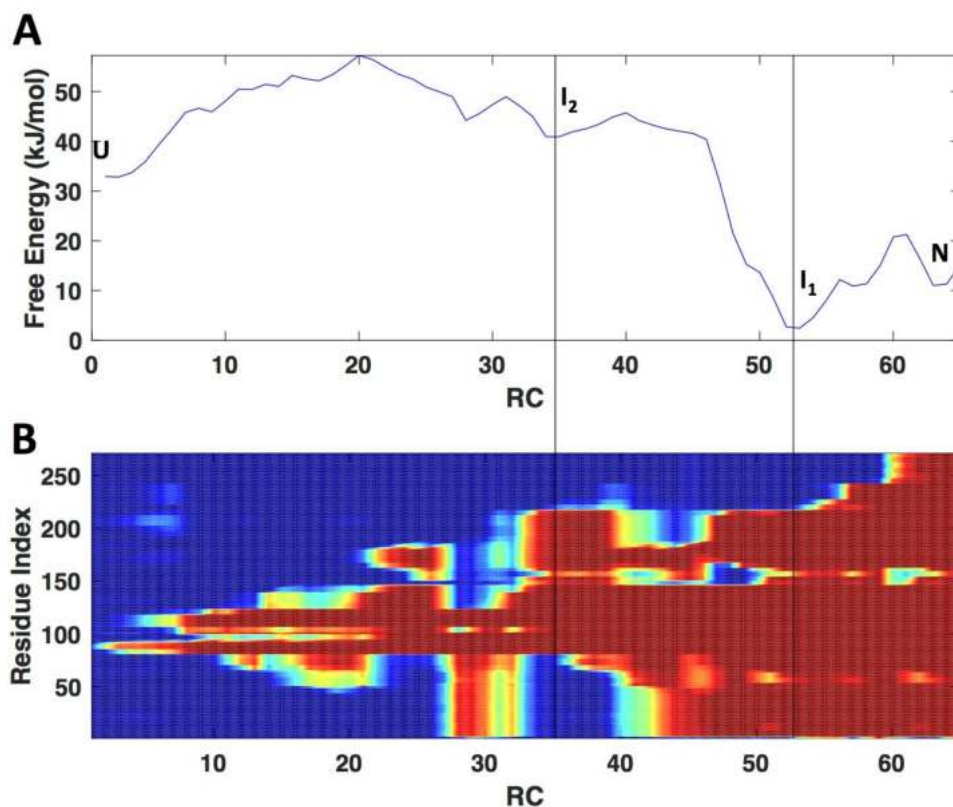
Figure S8. Overall HDX kinetics for 105 segments of NQO1$_{apo}$ variants. The color scale indicates different times of exchange. Data for WT$_{apo}$ are from [6].
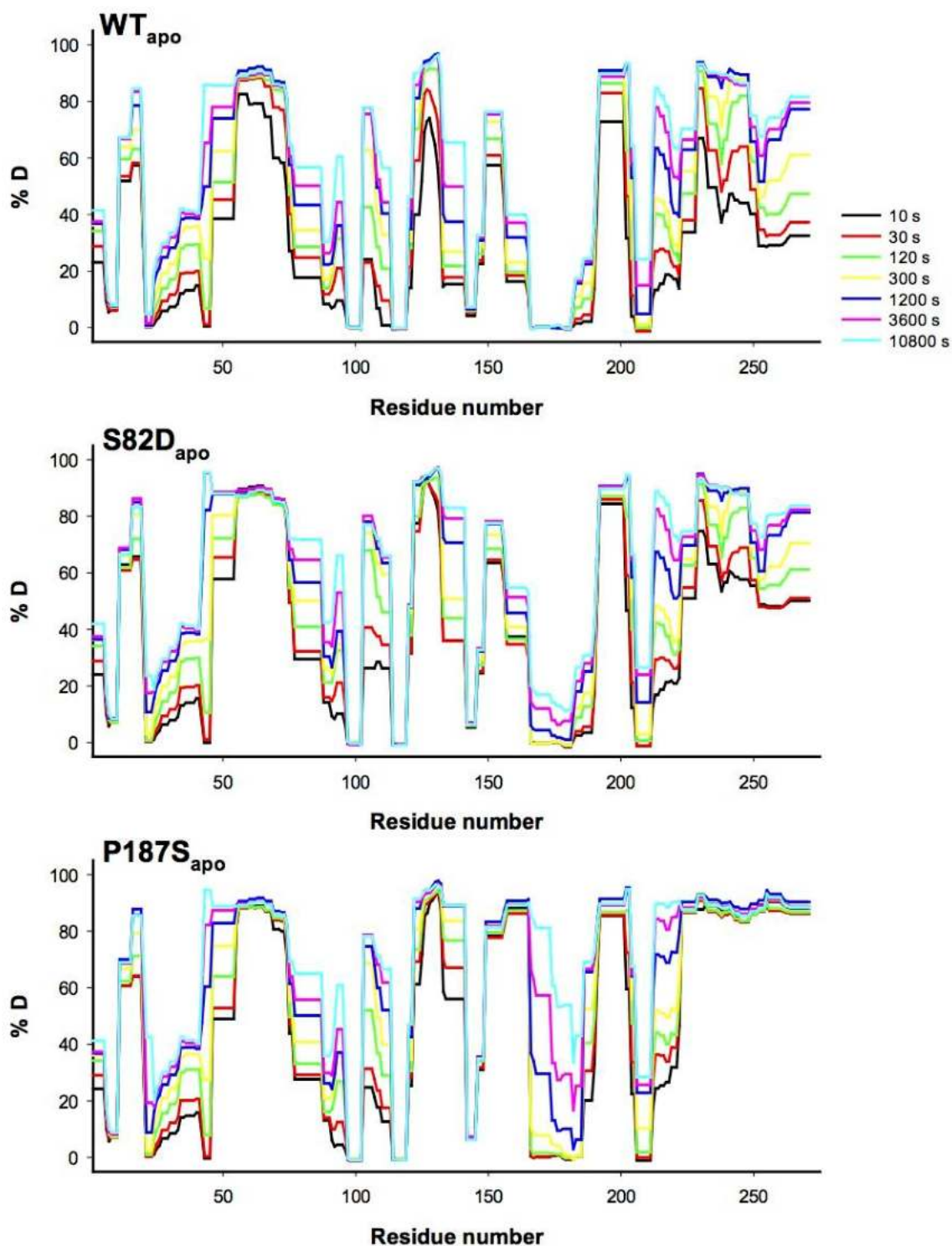
Figure S9. Overall HDX kinetics for 105 segments of NQO1$_{holo}$variants. The color scale indicates different times of exchange. Data for WT$_{holo}$ are from [6].
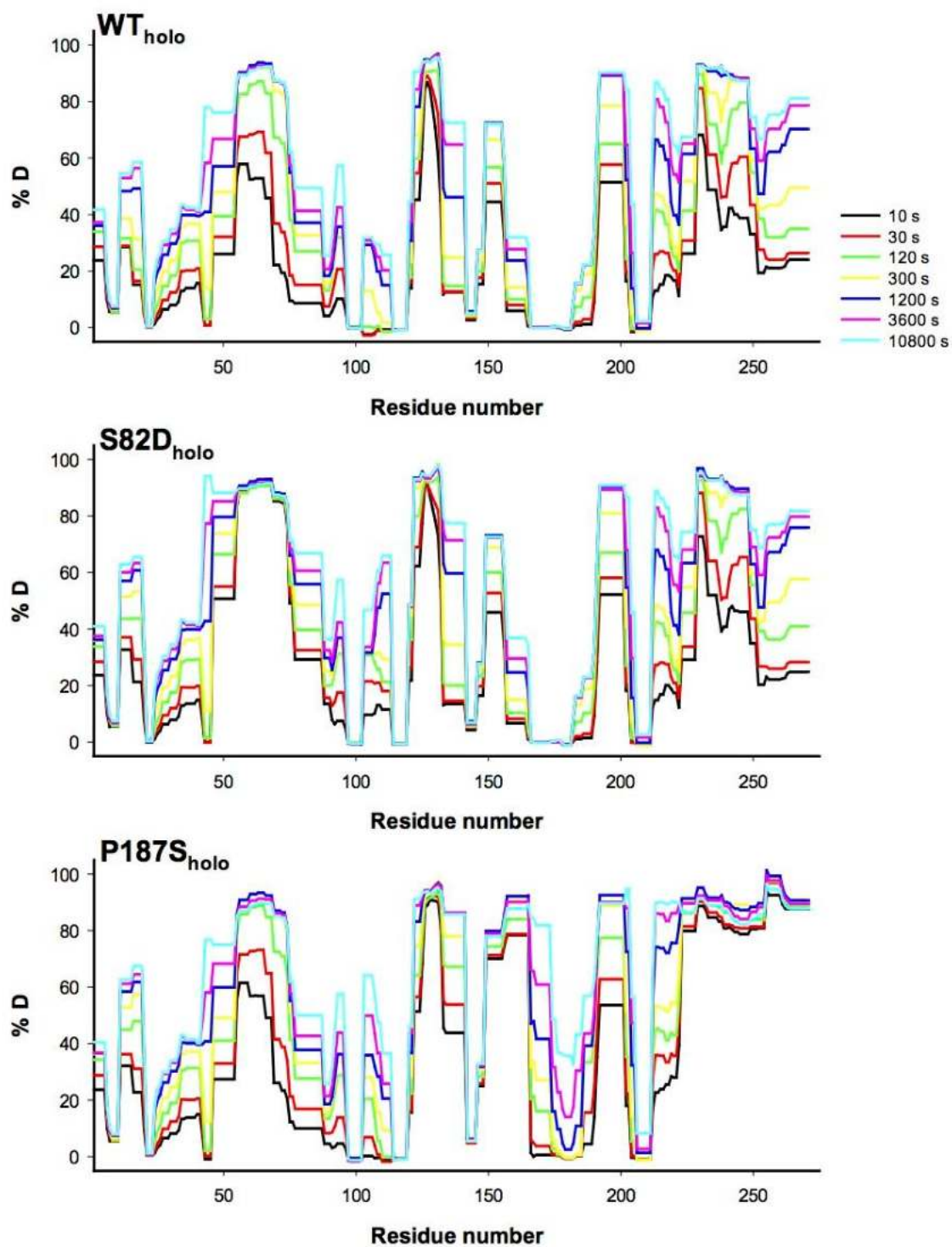
Figure S10. Changes in HDX kinetics due to the S82D and P187S mutations on different ligation states. Data are shown as Δ%D$_{av}$ for different protein segments using the WT protein as reference [6].
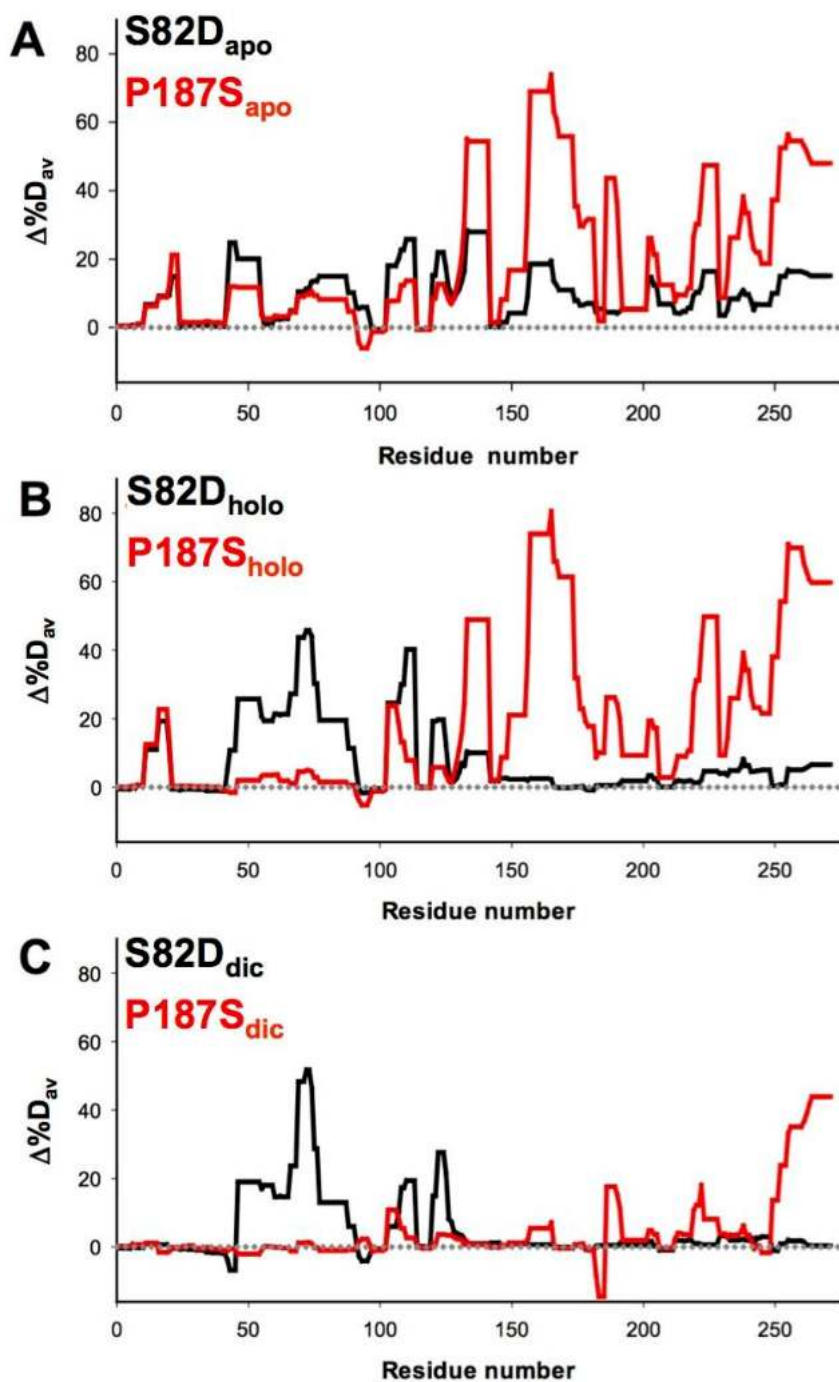
Figure S11. Effects of S82D and P187S on the structural stability of the MMI in NQO1_holo and NQO1_apo. The plots in grey scale show the location of S82 and P187 (as black sphere representation) and the FAD molecule (as black stick representation), whereas those in color scale show the residues with values of $\Delta\%D_{av} \geq 10\%$. Two different views rotated by 90º are shown.
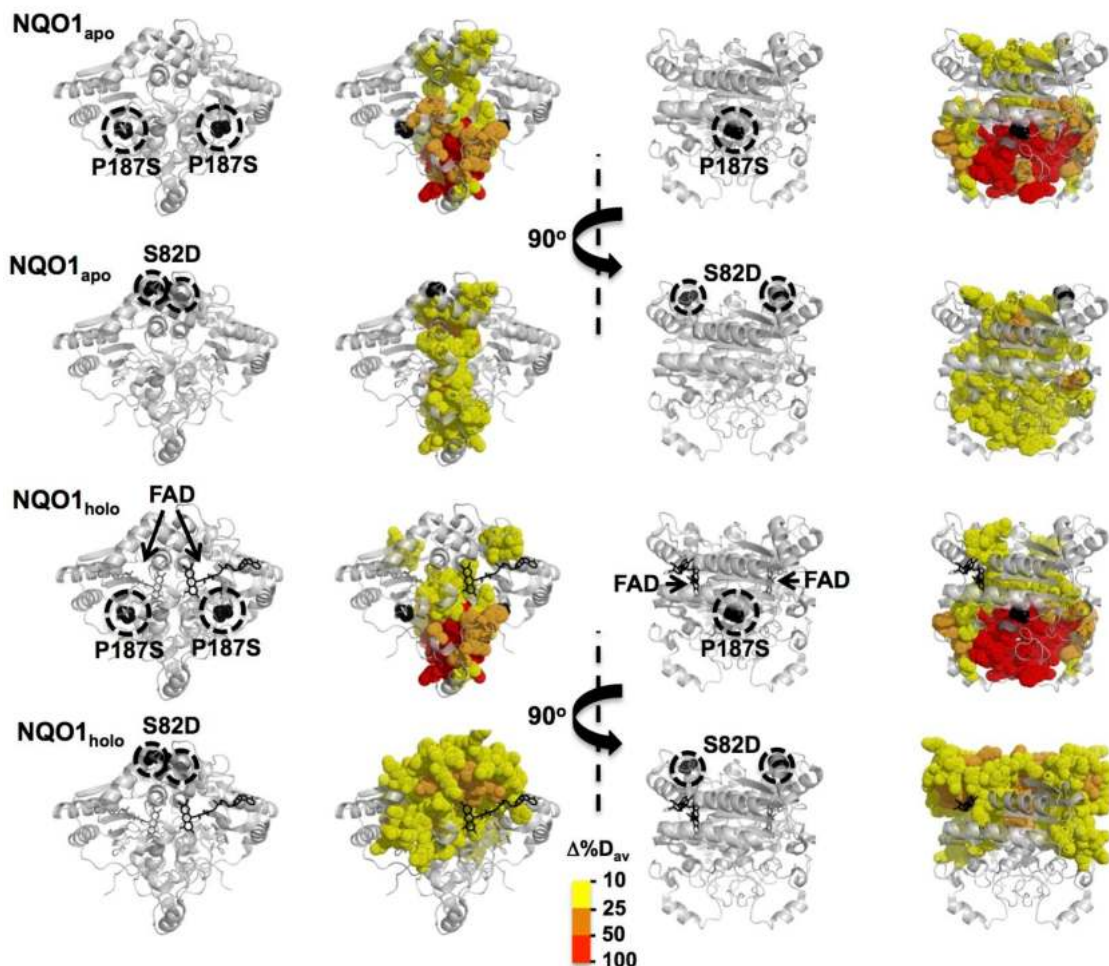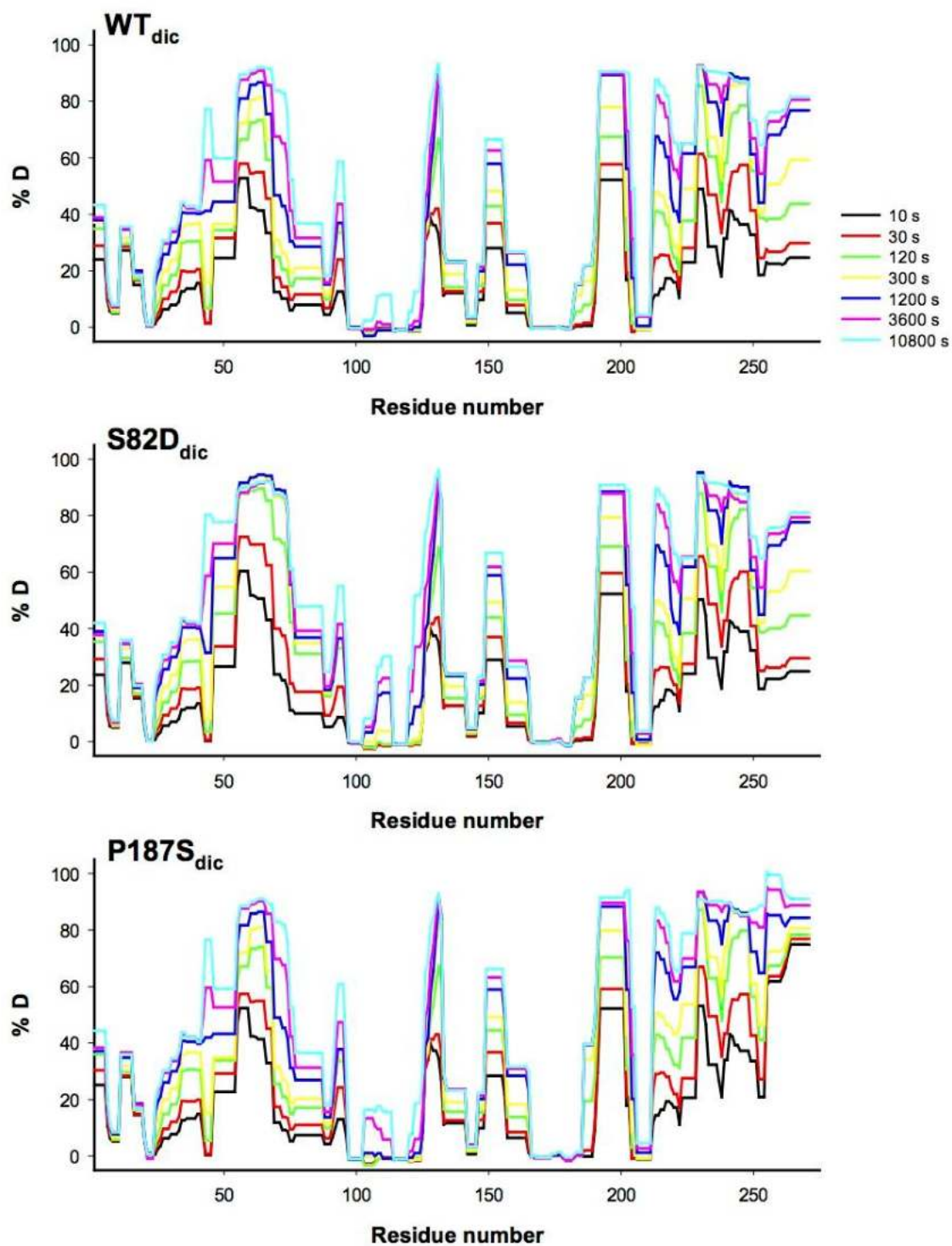
Figure S13. Effects of S82D and P187S on the structural stability of the MMI in NQO1$_{holo}$ and NQO1$_{dic}$. The plots in grey scale show the location of S82 and P187 (as black sphere representation) and the FAD and Dic molecules (as black stick representation), whereas those in color scale show the residues with values of $\Delta\%D_{av} \geq$ 10%. Two different views rotated by 90º are shown.
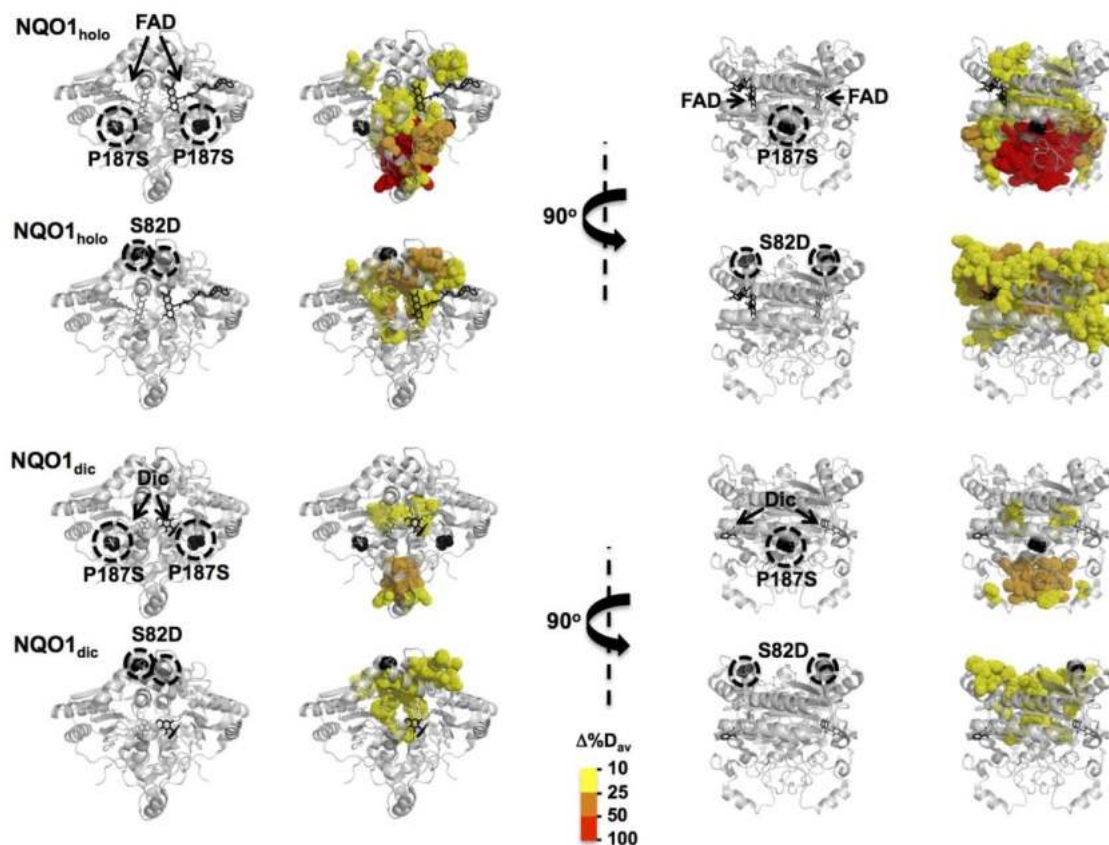
Figure S14. Reductive half-reaction of FAD bound to NQO1 variants with NADH.
A) Time-dependent spectral evolution. B) Spectral deconvolution from a three-step model (A→B→C→D). C) Kinetic traces for the absorption at 450 and 475 nm (thin dashed lines) and fittings to a three-step model (thick solid lines). The upper panels show the residuals corresponding to fittings using data at 450 nm. Data correspond to NQO1 and NADH both at 7.5 μM. Temperature was 6°C. Data for WT NQO1 are from [7].
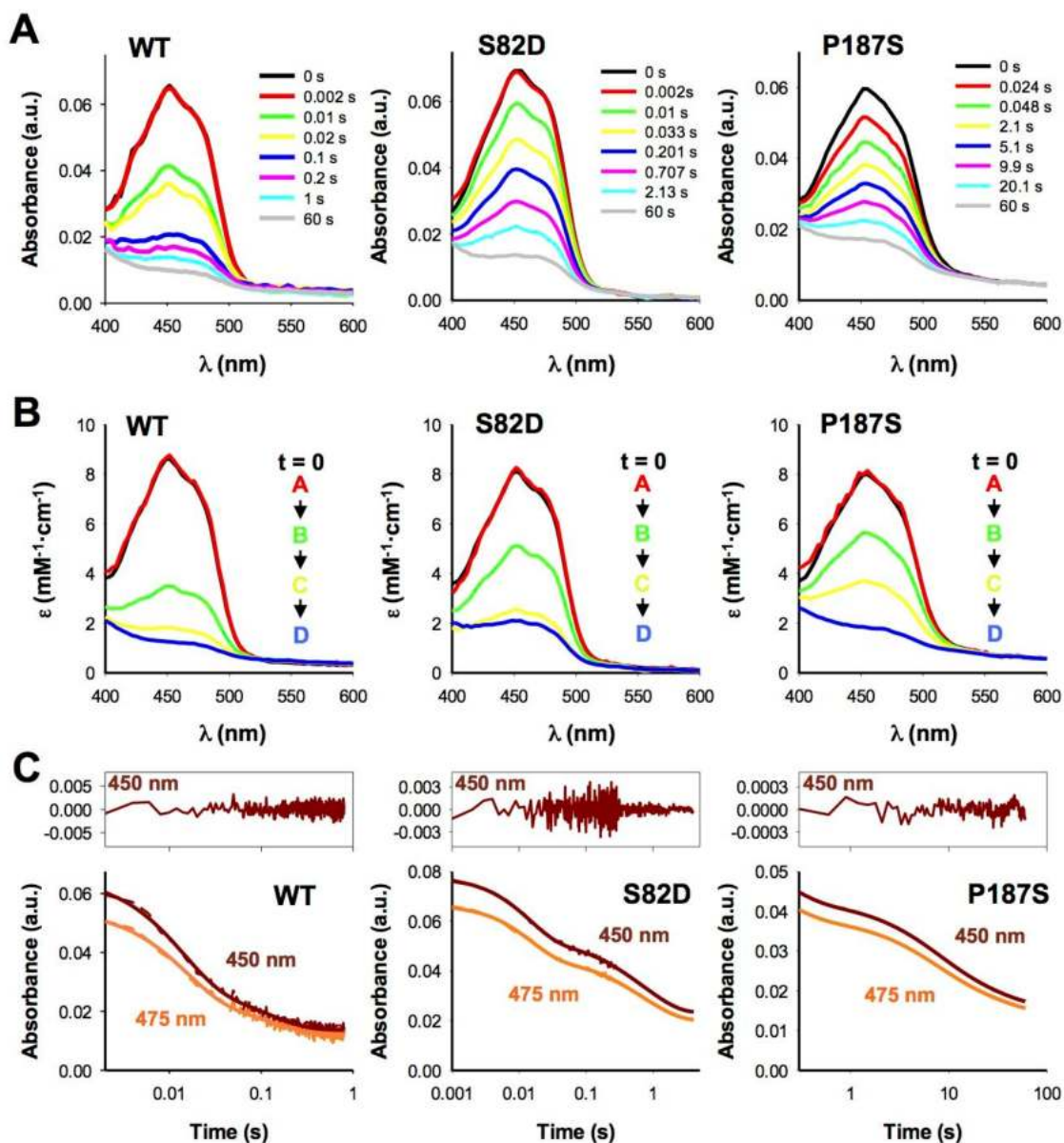
Figure S15. Reductive half-reaction of FAD bound to NQO1 variants with NADPH.
A) Time-dependent spectral evolution. B) Spectral deconvolution from a two-step
(A→B→C) or three-step model (A→B→C→D). When observed, the C→D step
accounted for a very slow spectroscopic change with small amplitude, likely reflecting a
process of little catalytic relevance. C) Amplitudes and observed rate constants. The right
panel shows the fraction corresponding to the main processes (fast, A→B, and slow,
B→C) determined as the fractional change in absorbance (average of 445-455 nm)
corresponding to the spectral deconvolution (see panel B). The left panel shows the
observed rate constant (k_{obs}) obtained from kinetic analysis (see Panel D). Kinetic traces
for the absorption at 450 and 475 nm (thin dashed lines) and fittings to a two-step or three-
step model (thick solid lines). The upper panels show the residuals corresponding to
fittings using data at 450 nm. Data correspond to NQO1 and NADPH both at 7.5 μM.
Temperature was 6°C.

29
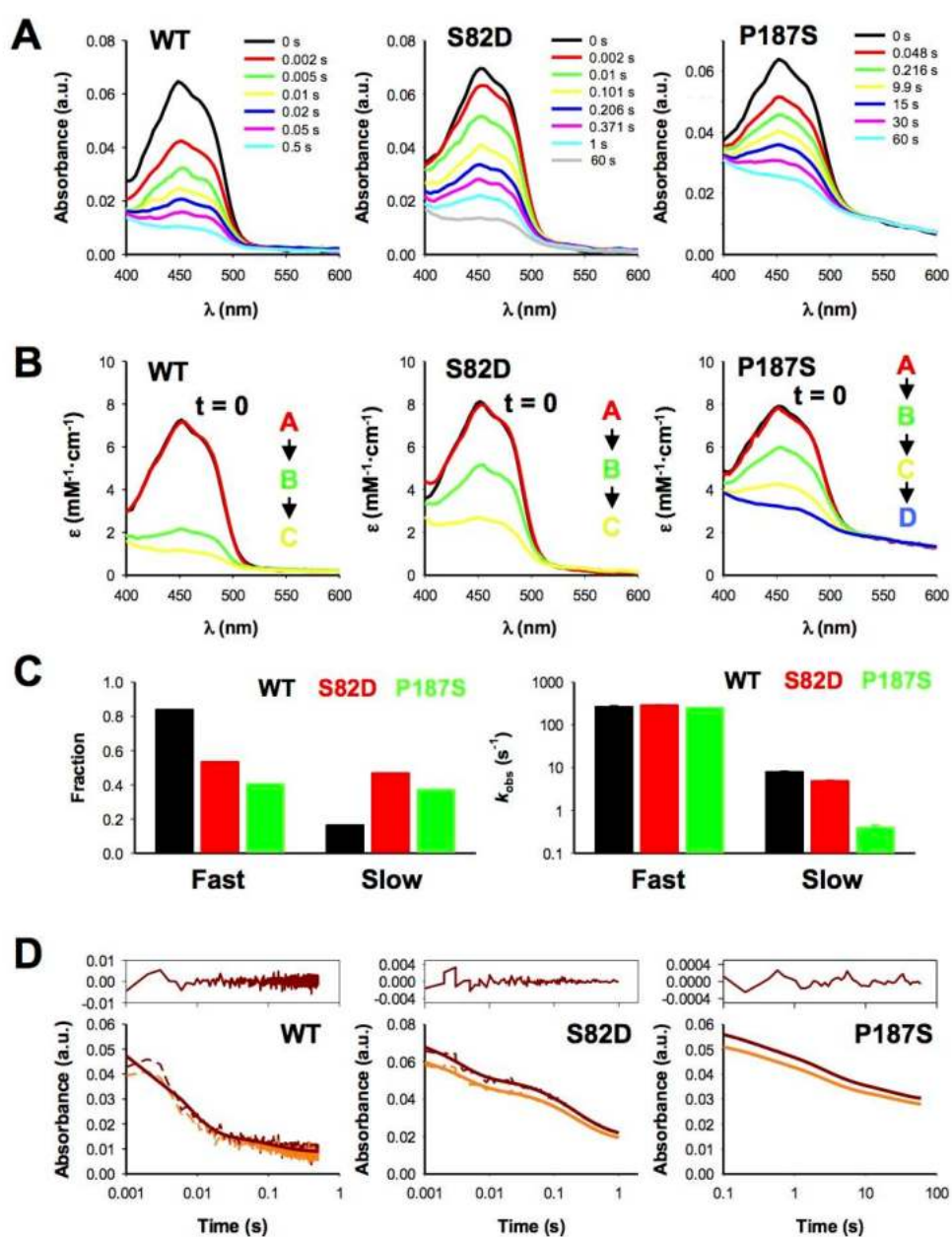
Figure S16. Oxidative half-reaction of NQO1$_{hq}$ variants with DCPIP. A) Time-dependent spectral evolution. B) Spectral deconvolution from a two-step model (A→B→C). C) Kinetic traces for the absorption at 450, 475 and 600 nm (thin dashed lines) and fittings to a two-step model (thick solid lines). The upper panels show the residuals corresponding to fittings using data at 450 nm. Data correspond to NQO1$_{hq}$ and DCPIP both at 7.5 μM. Temperature was 6°C.
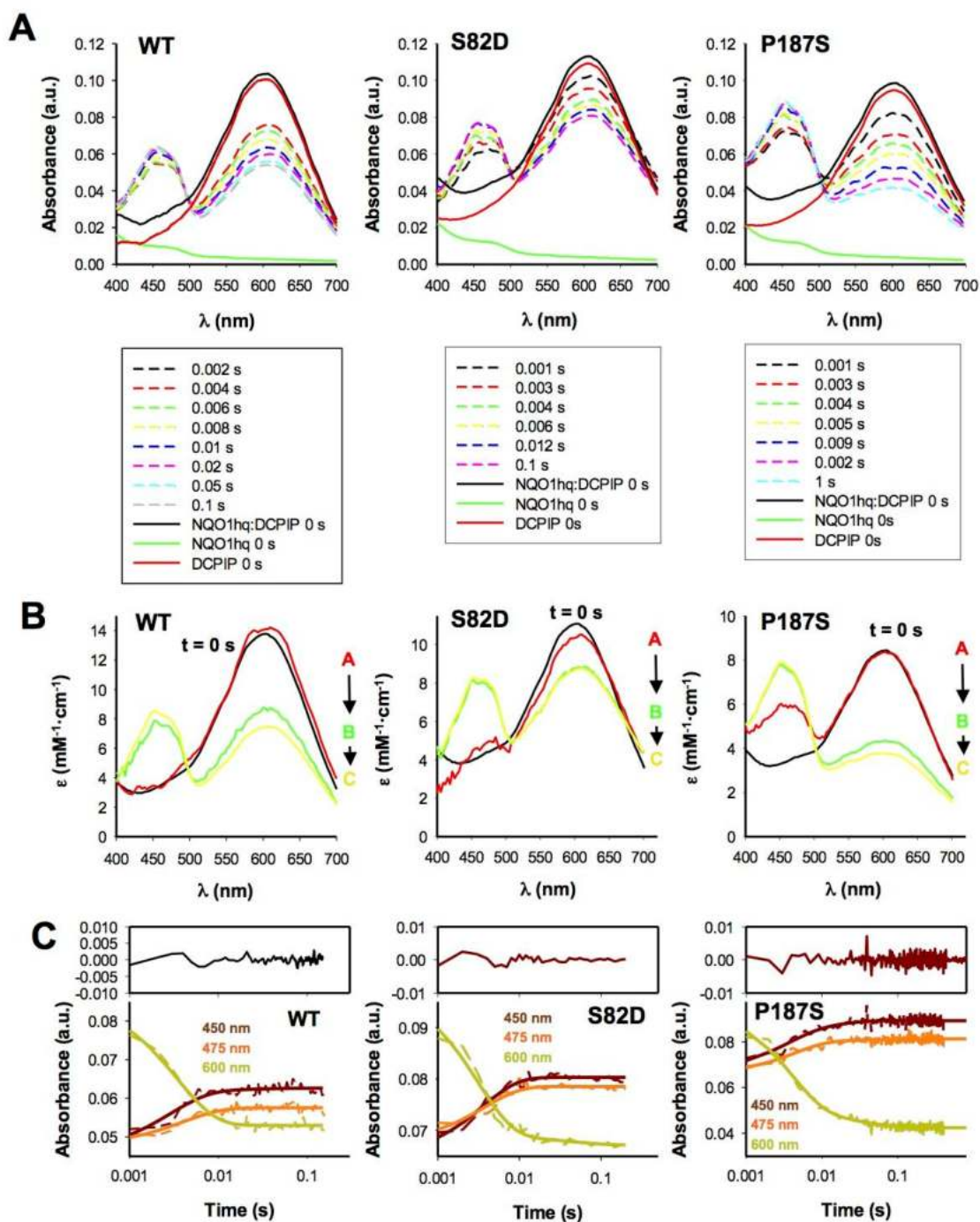
Figure S17. Representative SDS-PAGE gels for proteolysis kinetics with trypsin. Each variant and ligation state is indicated, as well as the time point of experiments and protease concentration.
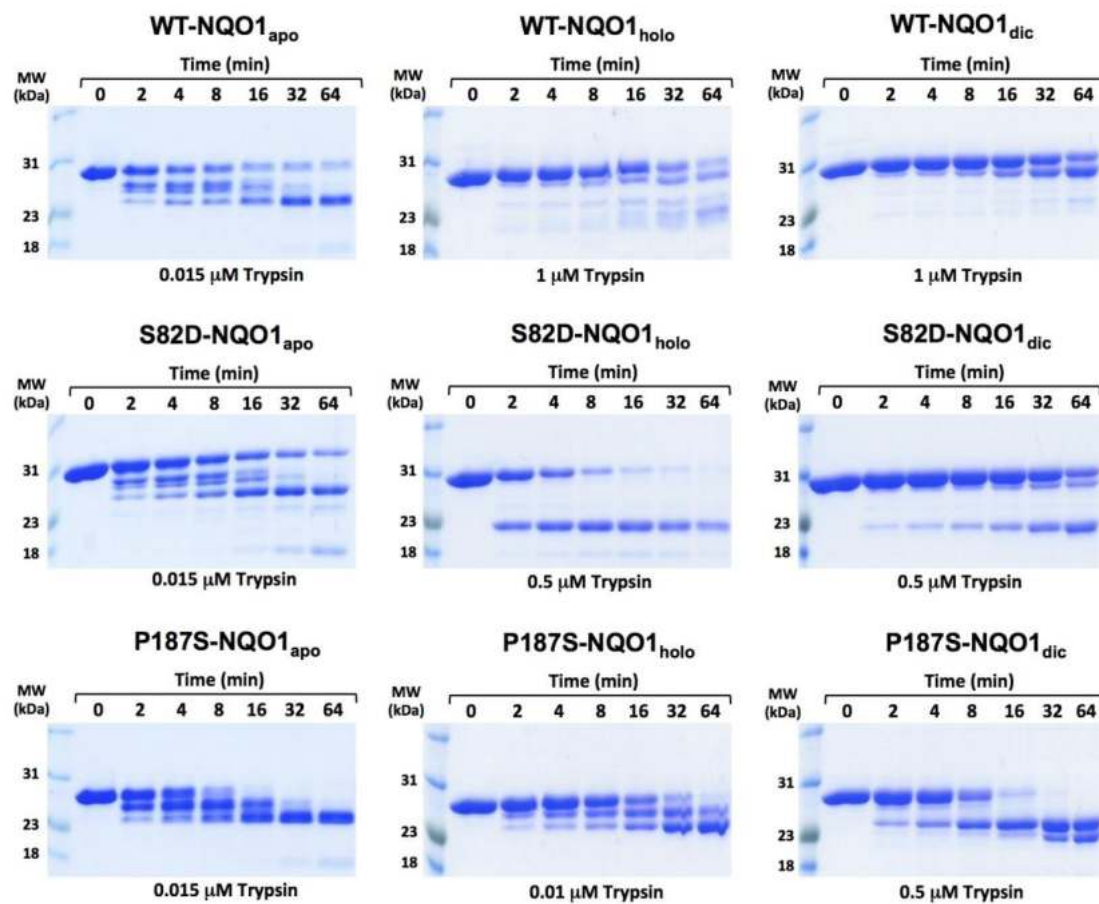
Figure S18. Proteolysis kinetics by trypsin of NQO1 variants in different ligation states. In different panels, the apparent first-order decays are shown at different protease concentrations for different variants and ligation states to provide k_{obs}. From the linear dependence of k_{obs} on protease concentration (left row), the second-order rate constant is calculated.
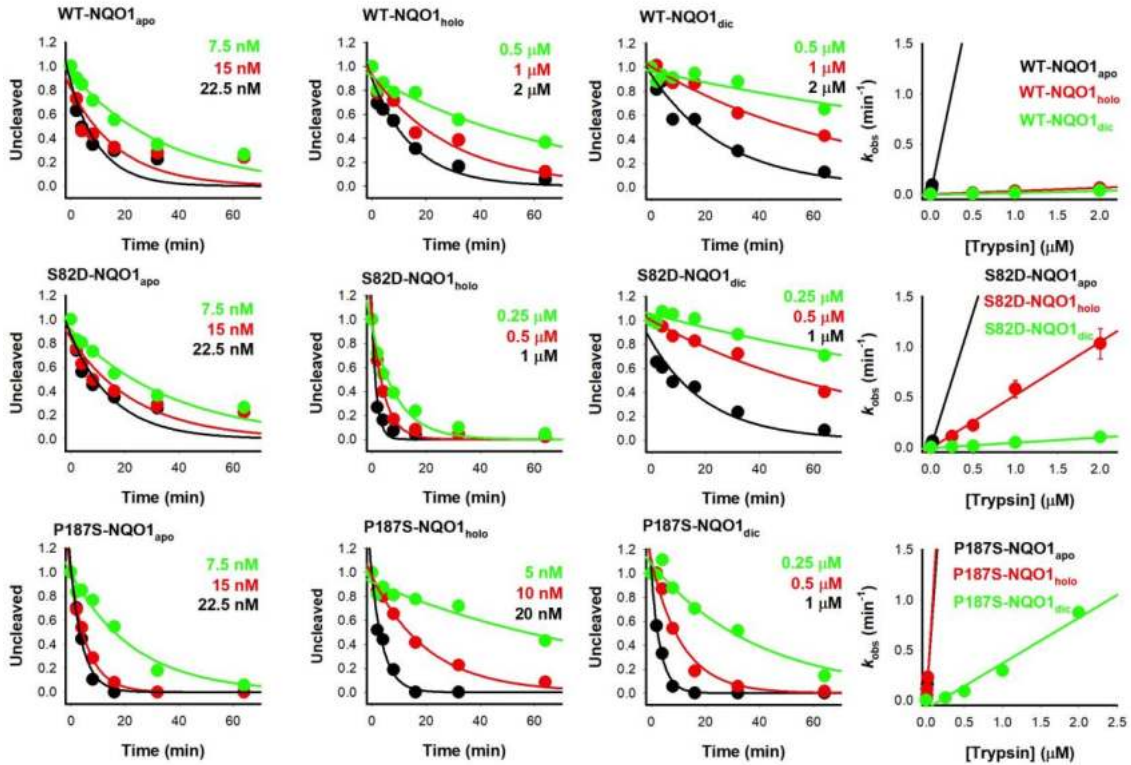
Figure S19. Dic binding thermodynamics by ITC. In A-C, left panels show representative titrations of NQO1 variants (15-17 μM in NQO1, 180 μM Dic; initial injection 0.5 μL, then 23 x 1.5 μL for WT and S82D or 18 x 2 μL in P187S; both protein and ligand with 100 μM FAD), whereas the right panels display representative binding isotherms at different temperatures. Panel D shows the temperature dependence of binding free energy, enthalpy and entropy. Data at 25°C are from 3-4 different titrations. Two protein batches from different purifications were used for each variant.
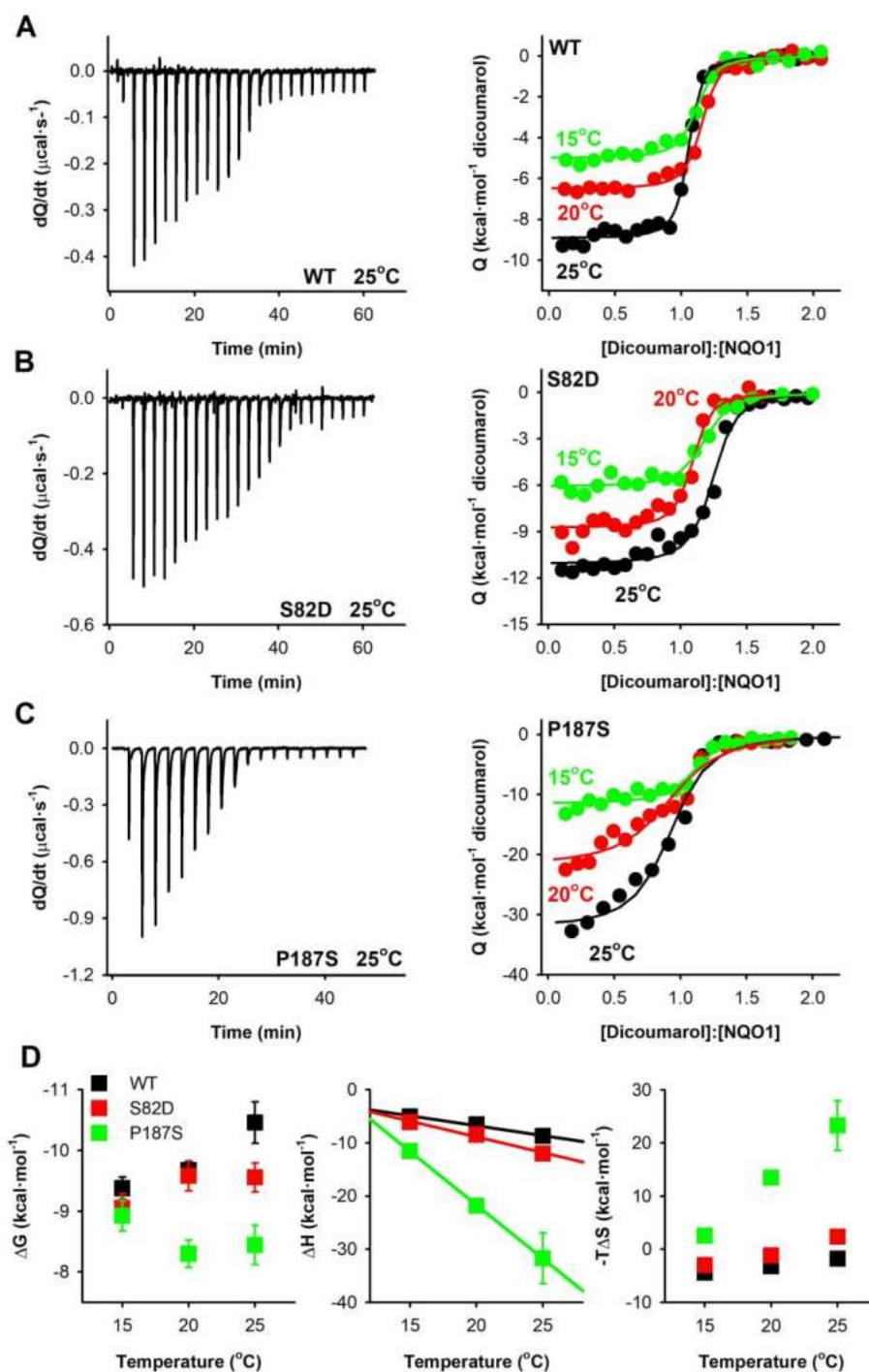
Figure S20. The stability of S82D and P187S mutants is more sensitive to proteasomal degradation than that of WT NQO1. Representative western-blot analysis of HAP-1 NQO1-KO cells transfected with NQO1 variants. Cells were treated for 4 h with the proteasomal inhibitor MG-132 and protein levels were determined by western-blot. Densitometric analysis of western-blots are indicated as mean ±s.d. from three 3 technical replicates and were normalized using WT NQO1 without treatment. Statistical analyses were carried out from one-tailed t-tests and significance reported as p values.
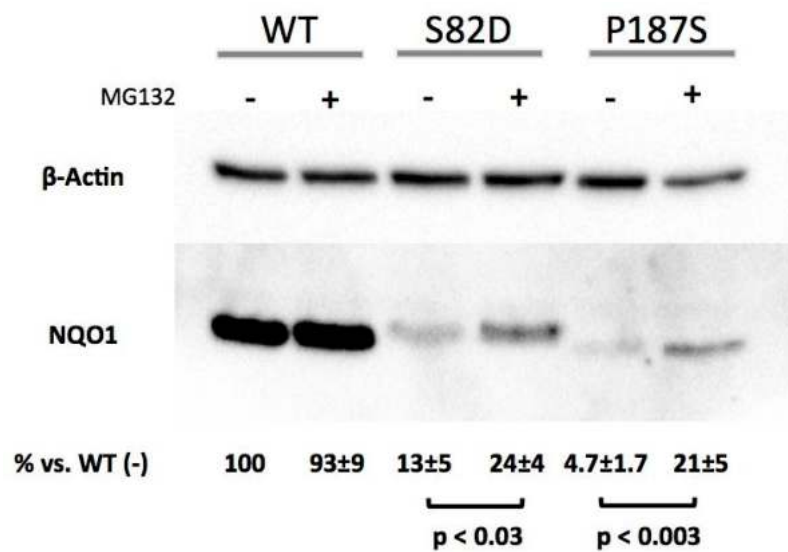
Table S1. Human flavoproteins display multiple functional features. Human flavoproteins (95) were searched according to the list provided by {Lienhart, 2013 #3986}. ***Function and catalytic activity*** was retrieved from the Uniprot server (https://www.uniprot.org/). For sake of simplicity, reactions are indicated only the forward direction. ***Protein:protein interactios (PPI)*** were retrieved from the BioGrid server (https://thebiogrid.org/). Only those interactions with at least two independent reports are included (note that the number of reports is indicated in parenthesis). 65% of the flavoproteins compiled in this table develop PPI. ***Subcellular locations*** were retrieved from the GeneCards server (https://www.genecards.org/). The confidence on the subcellular location for each protein is provided by this server a 1-5 scale (from the lowest to the highest confidence) and only those cases with a confidence ≥ 3 are included (note that the level of confidence is indicated in parenthesis. 78 % of the flavoproteins compiled are located in at least two different subcellular compartments. ***Structures for apo-proteins*** were retrieved (with their PDB code) from the Protein Data Bank in Europe (PDBe) server (https://www.ebi.ac.uk/pdbe/). Only those structures describing a full-length protein model without any bound ligand were compiled. Only 2% of the flavoproteins compiled in this table were found to have high-resolution crystal structure following this criteria . All data were retrieved between November 2020 and January 2021.

Name (Gene Name; UniProt code)	Function/Catalytic activity (UniProt)	PPI (BioGrid)	Subcellular location (GeneCards)	Apo-structures (PDBe)
1. D-lactate dehydrogenase (LDHD; Q86WU2)	Involved in D-lactate, but not L-lactate catabolic process. / (R)-lactate + 2 [Fe(III)cytochrome c] $\rightarrow$ 2 [Fe(II)cytochrome c] + 2 H$^+$ + pyruvate **EC:1.1.2.4**	None	Mitochondria (5)	None
2. Xanthine dehydrogenase/oxidase (XDH; P47989)	Key enzyme in purine degradation. Catalyzes the oxidation of hypoxanthine to xanthine. Catalyzes the oxidation of xanthine to uric acid. Contributes to the generation of reactive oxygen species. Has also low oxidase activity towards aldehydes (in vitro)./ H$_2$O + NAD$^+$ + xanthine $\rightarrow$ H$^+$ + NADH + urate **EC:1.17.1.4;** H$_2$O + hypoxanthine + NAD$^+$ $\rightarrow$ H$^+$ + NADH + xanthine; **EC:1.17.1.4**; H$_2$O + O$_2$ + xanthine $\rightarrow$ H$_2$O$_2$ + urate **EC:1.17.3.2**	None	Extracellular (5), peroxisome (4), cytosol (4), nucleus (3), endoplasmic reticulum (3)	None
3. Hydroxyacid oxidase 1 (HAO1; Q9UJM8)	Has 2-hydroxyacid oxidase activity. Most active on the 2-carbon substrate glycolate, but is also active on 2-hydroxy fatty acids, with high activity towards 2-hydroxy palmitate and 2-hydroxy octanoate./	None	Peroxisome (5), cytosol (5), nucleus (3)	None

	(2S)-2-hydroxycarboxylate + O_2 → 2-oxocarboxylate + H_2O_2 **EC:1.1.3.15**			
4. Hydroxyacid oxidase 2 (HAO2; Q9NYQ3)	Catalyzes the oxidation of L-alpha-hydroxy acids as well as, more slowly, that of L-alpha-amino acids./ (2S)-2-hydroxycarboxylate + O_2 → 2-oxocarboxylate + H_2O_2 **EC:1.1.3.15**	None	Peroxisome (5), cytosol (4)	None
5. Glycerol-3-phosphate dehydrogenase, mitochondrial (GPD2; P43304)	Calcium-responsive mitochondrial glycerol-3-phosphate dehydrogenase which seems to be a key component of the pancreatic beta-cell glucose-sensing device./ Quinone + sn-glycerol 3-phosphate → quinol + dihydroxyacetone phosphate **EC:1.1.5.3**	None	Mitochondria (5)	None
6. Choline dehydrogenase, mitochondrial (CHDH; Q8NE62)	A + choline → AH_2 + betaine aldehyde; **EC:1.1.99.1**	None	Mitochondria (5)	None
7. L-2-hydroxyglutarate dehydrogenase, mitochondrial (L2HGDH; Q9H9P8)	(S)-2-hydroxyglutarate + A → 2-oxoglutarate + AH_2 **EC:1.1.99.2**	None	Mitochondria (5)	None
8. D-2-hydroxyglutarate dehydrogenase (D2HGDH; Q8N465)	Catalyzes the oxidation of D-2-hydroxyglutarate to alpha-ketoglutarate./ (R)-2-hydroxyglutarate + A → 2-oxoglutarate + AH_2 **EC:1.1.99.-**	DHRS4L2 (2); ZNF462 (2)	Mitochondria (5)	None

# Enzyme	Function / Reaction	Interactors	Localization	
9. Aldehyde oxidase (AOX1; Q06278)	Oxidase with broad substrate specificity, oxidizing aromatic azaheterocycles, such as N^1-methylnicotinamide, N-methylphthalazinium and phthalazine, as well as aldehydes, such as benzaldehyde, retinal, pyridoxal, and vanillin. Plays a key role in the metabolism of xenobiotics and drugs containing aromatic azaheterocyclic substituents. Participates in the bioactivation of prodrugs such as famciclovir, catalyzing the oxidation step from 6-deoxypenciclovir to penciclovir, which is a potent antiviral agent. Is probably involved in the regulation of reactive oxygen species homeostasis. May be a prominent source of superoxide generation via the one-electron reduction of molecular oxygen. Also may catalyze nitric oxide (NO) production via the reduction of nitrite to NO with NADH or aldehyde as electron donor. May play a role in adipogenesis./ Aldehyde + H_2O + $O_2 \rightarrow$ Carboxylate + H^+ + H_2O_2 **EC:1.2.3.1**; H_2O + O_2 + retinal $\rightarrow H^+$ + H_2O_2 + retinoate	None	Cytosol (5), extracelular (4)	None
10. Dihydropyrimidine dehydrogenase (DPYD; Q12882)	5,6-dihydrouracil + $NADP^+ \rightarrow H^+$ + NADPH + uracil **EC:1.3.1.2**	GOPC (2)	Cytosol (5), nucleus (3)	None
11. Delta(24)-sterol reductase (DHCR24; Q15392)	Catalyzes the reduction of the delta-24 double bond of sterol intermediates during cholesterol biosynthesis. In addition to its cholesterol-synthesizing activity, can protect cells from oxidative stress by reducing caspase 3 activity during apoptosis induced by oxidative stress. Also protects against amyloid-beta peptide-induced	ACTA2 (2), APP (2), ATP12A (2), ERAP1 (2), FUNDC2 (2), HRAS (2), LUZP1 (2)	Nucleus (5), endoplasmic reticulum (5), Golgi apparatus(4), cytoskeleton (3), cytosol (3).	None

apoptosis. / Cholesterol + NADP$^+$ $\rightarrow$ cholesta-5,24-dien-3β-ol + H$^+$ + NADPH **EC:1.3.1.723**; H$^+$ + lanosterol + NADPH $\rightarrow$ 24,25-dihydrolanosterol + NADP$^+$; 5α-cholest-8-en-3β-ol + NADP$^+$ = 5α-cholesta-8,24-dien-3β-ol + H$^+$ + NADPH; **EC:1.3.1.72**

12. Dihydroorotate dehydrogenase (DHODH, Q02127)	(S)-dihydroorotate + a quinone $\rightarrow$ a quinol + orotate **EC:1.3.5.2**	None	Mitochondrion (5), nucleus (5), cytosol (5).	None
13. Peroxisomal acyl-coenzyme A oxidase 1 (ACOX1, Q15067)	Catalyzes the desaturation of acyl-CoAs to 2-trans-enoyl-CoAs. First enzyme of the fatty acid beta-oxidation pathway./ 2,3-saturated acyl-CoA + O$_2$ $\rightarrow$ (2E)-enoyl-CoA + H$_2$O$_2$ **EC:1.3.3.6**; hexadecanoyl-CoA + O$_2$ $\rightarrow$ (2E)-hexadecenoyl-CoA + H$_2$O$_2$; dodecanoyl-CoA + O$_2$ $\rightarrow$ (2E)-dodecenoyl-CoA + H$_2$O$_2$; O$_2$ + octanoyl-CoA $\rightarrow$ (2E)-octenoyl-CoA + H$_2$O$_2$; decanoyl-CoA + O$_2$ = (2E)-decenoyl-CoA + H$_2$O$_2$; O$_2$ + tetradecanoyl-CoA $\rightarrow$ (2E)-tetradecenoyl-CoA + H$_2$O$_2$; hexadecanedioyl-CoA + O$_2$ $\rightarrow$ (2E)-hexadecenedioyl-CoA + H$_2$O$_2$; O$_2$ + tetracosanoyl-CoA $\rightarrow$ (2E)-tetracosenoyl-CoA + H$_2$O$_2$; glutaryl-CoA + O$_2$ $\rightarrow$ (2E)-glutaconyl-CoA + H$_2$O$_2$; hexanoyl-CoA + O$_2$ $\rightarrow$ (2E)-hexenoyl-CoA + H$_2$O$_2$; O$_2$ + octadecanoyl-CoA $\rightarrow$ (2E)-octadecenoyl-CoA + H$_2$O$_2$.	PEX5 (4), CAT (3), FOXS1 (2), MBNL2 (2), UBE2D2 (2).	Peroxisome (5), cytosol (4), plasma membrane (3), mitochondrion (3), nucleus (3).	None
14. Peroxisomal acyl-coenzyme A oxidase 2 (ACOX2; Q99424)	Oxidizes the CoA esters of the bile acid intermediates di- and tri-hydroxycholestanoic acids. Capable of oxidizing short as well as long chain 2-methyl branched fatty acids./(25R)-3α,7α,12α-trihydroxy-5β-cholestan-26-oyl-CoA + A + H2O $\rightarrow$ (24R,25R)-	None	Peroxisome (5), cytosol (5)	None

38

	$3\alpha,7\alpha,12\alpha,24$-tetrahydroxy-$5\beta$-cholestan-26-oyl-CoA + AH_2 **EC:1.17.99.31**; (25S)-$3\alpha,7\alpha,12\alpha$-trihydroxy-5β-cholestan-26-oyl-CoA + $O_2 \rightarrow$ (24*E*)-$3\alpha,7\alpha,12\alpha$-trihydroxy-5β-cholest-24-en-26-oyl-CoA + H_2O_2			
15. Peroxisomal acyl-coenzyme A oxidase 3 (ACOX3; O15254)	Oxidizes the CoA-esters of 2-methyl-branched fatty acids./ 2,3-saturated acyl-CoA + $O_2 \rightarrow$ a (2*E*)-enoyl-CoA + H_2O_2 **EC:1.3.3.6**; (2*S*)-pristanoyl-CoA + $O_2 \rightarrow$ (2*E*)-pristenoyl-CoA + H_2O_2; O_2 + tetracosanoyl-CoA $\rightarrow$ (2*E*)-tetracosenoyl-CoA + H_2O_2; hexadecanoyl-CoA + $O_2 \rightarrow$ (2*E*)-hexadecenoyl-CoA + H_2O_2; hexadecanedioyl-CoA + $O_2 \rightarrow$ (2*E*)-hexadecenedioyl-CoA + H_2O_2	FKBP6 (2), MICU2 (2), TMEM150A (2)	Peroxisome (5), cytosol (4)	None
16. Succinate-hydroxymethylglutarate CoA-transferase (SUGCT; Q9HAC7)	Catalyzes the succinyl-CoA-dependent conversion of glutarate to glutaryl-CoA. Can use different dicarboxylic acids as CoA acceptors, the preferred ones are glutarate, succinate, adipate, and 3-hydroxymethylglutarate./ 3-hydroxy-3-methylglutarate + succinyl-CoA $\rightarrow$ (3*S*)-hydroxy-3-methylglutaryl-CoA + succinate **EC:2.8.3.13**	SKAP1 (2)	Mitochondrion (5)	None
17. Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial (SDHA; P31040)	Flavoprotein (FP) subunit of succinate dehydrogenase (SDH) that is involved in complex II of the mitochondrial electron transport chain and is responsible for transferring electrons from succinate to ubiquinone (coenzyme Q). Can act as a tumor suppressor. / quinone + succinate $\rightarrow$ quinol + fumarate **EC:1.3.5.11**	SDHB (3), ACAD9 (2), ATP5F1 (2), ATXN3 (2), CS (2), CYB5R3 (2), EMC2 (2), MDH2 (2), NDUFV1 (2), PDHA1 (2), PTPN3 (2), SDHAF2 (2), SLC27A2 (2), SUCLG2 (2), UQCRB (2), UQCRFS1 (2)	Mitochondrion (5), nucleus (4)	None
18. Short-chain specific acyl-CoA dehydrogenase, mitochondrial	Short-chain specific acyl-CoA dehydrogenase is one of the acyl-CoA dehydrogenases that catalyze the first step of mitochondrial fatty acid beta-oxidation, an	SEC22C (2)	Cytoskeleton (5), mitochondrion (5),	None

(ACADS; P16219).	aerobic process breaking down fatty acids into acetyl-CoA and allowing the production of energy from fats. The first step of fatty acid beta-oxidation consists in the removal of one hydrogen from C-2 and C-3 of the straight-chain fatty acyl-CoA thioester, resulting in the formation of trans-2-enoyl-CoA. Among the different mitochondrial acyl-CoA dehydrogenases, short-chain specific acyl-CoA dehydrogenase acts specifically on acyl-CoAs with saturated 4 to 6 carbons long primary chains. /Short-chain 2,3-saturated fatty acyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ short-chain (2E)-enoyl-CoA + reduced [electron-transfer flavoprotein] **EC:1.3.8.11**; butanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-butenoyl-CoA + reduced [electron-transfer flavoprotein] **EC:1.3.8.11;** H^+ + oxidized [electron-transfer flavoprotein] + pentanoyl-CoA $\rightarrow$ (2E)-pentenoyl-CoA + reduced [electron-transfer flavoprotein]; H^+ + hexanoyl-CoA + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-hexenoyl-CoA + reduced [electron-transfer flavoprotein]		nucleus (5), cytosol (3)
19. Medium-chain specific acyl-CoA dehydrogenase, mitochondrial (ACADM; P11310)	Medium-chain specific acyl-CoA dehydrogenase is one of the acyl-CoA dehydrogenases that catalyze the first step of mitochondrial fatty acid beta-oxidation, an aerobic process breaking down fatty acids into acetyl-CoA and allowing the production of energy from fats. The first step of fatty acid beta-oxidation consists in the removal of one hydrogen from C-2 and C-3 of the straight-chain fatty acyl-CoA thioester, resulting in the formation of trans-2-enoyl-CoA. Electron transfer	None	Mitochondrion (5), nucleus (5), peroxisome (3), cytosol (3)	None

flavoprotein (ETF) is the electron acceptor that transfers electrons to the main mitochondrial respiratory chain via ETF-ubiquinone oxidoreductase (ETF dehydrogenase). Among the different mitochondrial acyl-CoA dehydrogenases, medium-chain specific acyl-CoA dehydrogenase acts specifically on acyl-CoAs with saturated 6 to 12 carbons long primary chains./ a medium-chain 2,3-saturated fatty acyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ a medium-chain trans-(2E)-enoyl-CoA + reduced [electron-transfer flavoprotein] **EC:1.3.8.7**;

H^+ + oxidized [electron-transfer flavoprotein] + pentanoyl-CoA $\rightarrow$ (2E)-pentenoyl-CoA + reduced [electron-transfer flavoprotein]; H^+ + hexanoyl-CoA + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-hexenoyl-CoA + reduced [electron-transfer flavoprotein]; H^+ + octanoyl-CoA + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-octenoyl-CoA + reduced [electron-transfer flavoprotein];

decanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-decenoyl-CoA + reduced [electron-transfer flavoprotein]; dodecanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-dodecenoyl-CoA + reduced [electron-transfer flavoprotein];

H+ + oxidized [electron-transfer flavoprotein] + tetradecanoyl-CoA $\rightarrow$ (2E)-tetradecenoyl-CoA + reduced [electron-transfer flavoprotein];

H^+ + hexadecanoyl-CoA + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-hexadecenoyl-CoA + reduced [electron-transfer flavoprotein]

20. Glutaryl-CoA dehydrogenase, mitochondrial (GCDH; Q92947)	Catalyzes the oxidative decarboxylation of glutaryl-CoA to crotonyl-CoA and CO2 in the degradative pathway of L-lysine, L-hydroxylysine, and L-tryptophan metabolism. It uses electron transfer flavoprotein as its electron acceptor./ Glutaryl-CoA + 2 H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-butenoyl-CoA + CO_2 + reduced [electron-transfer flavoprotein] **EC:1.3.8.6**	ISCA1 (2), MARS2 (2), NOS3 (2)	Mitochondrion (5)	None
21. Isovaleryl-CoA dehydrogenase, mitochondrial (IVD; P26440)	Catalyzes the conversion of isovaleryl-CoA/3-methylbutanoyl-CoA to 3-methylbut-2-enoyl-CoA as an intermediate step in the leucine (Leu) catabolic pathway. To a lesser extent, is also able to catalyze the oxidation of other saturated short-chain acyl-CoA thioesters as pentanoyl-CoA, hexenoyl-CoA and butenoyl-CoA./ 3-methylbutanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ 3-methylbut-2-enoyl-CoA + reduced [electron-transfer flavoprotein] **EC:1.3.8.4**; H^+ + oxidized [electron-transfer flavoprotein] + pentanoyl-CoA $\rightarrow$ (2E)-pentenoyl-CoA + reduced [electron-transfer flavoprotein]; H^+ + hexanoyl-CoA + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-hexenoyl-CoA + reduced [electron-transfer flavoprotein]; butanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-butenoyl-CoA + reduced [electron-transfer flavoprotein] **EC:1.3.8.1**	ETFA (2), PMPCA (2)	Mitochondrion (5), nucleus (3), cytosol (3)	None
22. Short/branched chain specific acyl-CoA dehydrogenase, mitochondrial (ACADSB; P45954).	Short and branched chain specific acyl-CoA dehydrogenase that catalyzes the removal of one hydrogen from C-2 and C-3 of the fatty acyl-CoA thioester, resulting in the formation of trans-2-enoyl-CoA. Among the different mitochondrial acyl-CoA dehydrogenases, acts specifically on short and branched chain acyl-CoA derivatives such as (*S*)-2-	None	Mitochondrion (5)	None

methylbutyryl-CoA as well as short straight chain acyl-CoAs such as butyryl-CoA. Plays an important role in the metabolism of L-isoleucine by catalyzing the dehydrogenation of 2-methylbutyryl-CoA, one of the steps of the L-isoleucine catabolic pathway. Can also act on valproyl-CoA, a metabolite of valproic acid, an antiepileptic drug./

2-methylbutanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-2-methylbut-2-enoyl-CoA + reduced [electron-transfer flavoprotein];

(2S)-2-methylbutanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-2-methylbut-2-enoyl-CoA + reduced [electron-transfer flavoprotein]; (2R)-2-methylbutanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ ethylacryloyl-CoA + reduced [electron-transfer flavoprotein]; butanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-butenoyl-CoA + reduced [electron-transfer flavoprotein]; 2-methylpropanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] = (2E)-2-methylpropenoyl-CoA + reduced [electron-transfer flavoprotein]; H^+ + hexanoyl-CoA + oxidized [electron-transfer flavoprotein] $\rightarrow$ (2E)-hexenoyl-CoA + reduced [electron-transfer flavoprotein]; 2-methylhexanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ 2-methylhexenoyl-CoA + reduced [electron-transfer flavoprotein];

H^+ + oxidized [electron-transfer flavoprotein] + valproyl-CoA $\rightarrow$ (2E)-2-propylpent-2-enoyl-CoA + reduced [electron-transfer flavoprotein]

| 23. Long-chain specific acyl-CoA dehydrogenase, mitochondrial (ACADL; P28330) | a long-chain 2,3-saturated fatty acyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → a long-chain (2E)-enoyl-CoA + reduced [electron-transfer flavoprotein] **EC:1.3.8.8**; H$^+$ + hexanoyl-CoA + oxidized [electron-transfer flavoprotein] → (2E)-hexenoyl-CoA + reduced [electron-transfer flavoprotein]; H+ + octanoyl-CoA + oxidized [electron-transfer flavoprotein] → (2E)-octenoyl-CoA + reduced [electron-transfer flavoprotein]; decanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E)-decenoyl-CoA + reduced [electron-transfer flavoprotein]; dodecanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E)-dodecenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + oxidized [electron-transfer flavoprotein] + tetradecanoyl-CoA → (2E)-tetradecenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + hexadecanoyl-CoA + oxidized [electron-transfer flavoprotein] → (2E)-hexadecenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + octadecanoyl-CoA + oxidized [electron-transfer flavoprotein] → (2E)-octadecenoyl-CoA + reduced [electron-transfer flavoprotein]; eicosanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E)-eicosenoyl-CoA + reduced [electron-transfer flavoprotein]; docosanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E)-docosenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + oxidized [electron-transfer flavoprotein] + tetracosanoyl-CoA → (2E)-tetracosenoyl-CoA + reduced [electron-transfer flavoprotein]; (5E)-tetradecenoyl-CoA + H+ + oxidized [electron-transfer | SYT2 (2), TMPRSS11A (2) | Mitochondrion (5), cytosol (3) | None |

flavoprotein] → (2*E*,5*E*)-tetradecadienoyl-CoA + reduced [electron-transfer flavoprotein]; (5*Z*)-tetradecenoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2*E*,5*Z*)-tetradecadienoyl-CoA + reduced [electron-transfer flavoprotein]; (9*Z*)-octadecenoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E,9Z)-octadecadienoyl-CoA + reduced [electron-transfer flavoprotein]

24. Very long-chain specific acyl-CoA dehydrogenase, mitochondrial (ACADVL; P49748).	Very long-chain specific acyl-CoA dehydrogenase is one of the acyl-CoA dehydrogenases that catalyze the first step of mitochondrial fatty acid beta-oxidation, an aerobic process breaking down fatty acids into acetyl-CoA and allowing the production of energy from fats. The first step of fatty acid beta-oxidation consists in the removal of one hydrogen from C-2 and C-3 of the straight-chain fatty acyl-CoA thioester, resulting in the formation of trans-2-enoyl-CoA. Among the different mitochondrial acyl-CoA dehydrogenases, very long-chain specific acyl-CoA dehydrogenase acts specifically on acyl-CoAs with saturated 12 to 24 carbons long primary chains./ a very-long-chain 2,3-saturated fatty acyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → a very-long-chain (2*E*)-enoyl-CoA + reduced [electron-transfer flavoprotein] **EC:1.3.8.9**; decanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] = (2*E*)-decenoyl-CoA + reduced [electron-transfer flavoprotein]; dodecanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2*E*)-dodecenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + oxidized [electron-transfer flavoprotein] + tetradecanoyl-CoA → (2*E*)-tetradecenoyl-CoA + reduced [electron-	SURF1 (3), ECSIT (2), HIBCH (2), HMGCL (2), MRM1 (2), SIRT3 (2), SOCS3 (2)	Mitochondrion (5), nucleus (5), cytosol (5), peroxisome (3)	None

transfer flavoprotein]; H$^+$ + hexadecanoyl-CoA + oxidized [electron-transfer flavoprotein] → (2E)-hexadecenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + octadecanoyl-CoA + oxidized [electron-transfer flavoprotein] → (2E)-octadecenoyl-CoA + reduced [electron-transfer flavoprotein]; eicosanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E)-eicosenoyl-CoA + reduced [electron-transfer flavoprotein];
docosanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E)-docosenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + oxidized [electron-transfer flavoprotein] + tetracosanoyl-CoA → (2E)-tetracosenoyl-CoA + reduced [electron-transfer flavoprotein]; (9Z)-hexadecenoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E,9Z)-hexadecadienoyl-CoA + reduced [electron-transfer flavoprotein];
(9Z)-octadecenoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E,9Z)-octadecadienoyl-CoA + reduced [electron-transfer flavoprotein]

25. Isobutyryl-CoA dehydrogenase, mitochondrial (ACAD8; Q9UKU7)	Isobutyryl-CoA dehydrogenase which catalyzes one of the steps of the valine catabolic pathway. To a lesser extent, is also able to catalyze the oxidation of (2S)-2-methylbutanoyl-CoA./2-methylpropanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E)-2-methylpropenoyl-CoA + reduced [electron-transfer flavoprotein]; (2S)-2-methylbutanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E)-2-methylbutenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + oxidized [electron-transfer	SRGAP3 (2), VSTM2A (2)	Mitochondrion (5)	None

flavoprotein] + propanoyl-CoA → acryloyl-CoA + reduced [electron-transfer flavoprotein]

| 26. Complex I assembly factor ACAD9, mitochondrial (ACAD9; Q9H845) | As part of the MCIA complex, primarily participates in the assembly of the mitochondrial complex I and therefore plays a role in oxidative phosphorylation. This moonlighting protein has also a dehydrogenase activity toward a broad range of substrates with greater specificity for long-chain unsaturated acyl-CoAs. However, in vivo, it does not seem to play a primary role in fatty acid oxidation. In addition, the function in complex I assembly is independent of the dehydrogenase activity of the protein. /eicosanoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E)-eicosenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + octadecanoyl-CoA + oxidized [electron-transfer flavoprotein] → (2E)-octadecenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + hexadecanoyl-CoA + oxidized [electron-transfer flavoprotein] → (2E)-hexadecenoyl-CoA + reduced [electron-transfer flavoprotein]; decanoyl-CoA + H+ + oxidized [electron-transfer flavoprotein] → (2E)-decenoyl-CoA + reduced [electron-transfer flavoprotein]; H+ + nonanoyl-CoA + oxidized [electron-transfer flavoprotein] → (2E)-nonenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + oxidized [electron-transfer flavoprotein] + pentadecanoyl-CoA → (2E)-pentadecenoyl-CoA + reduced [electron-transfer flavoprotein]; H+ + oxidized [electron-transfer flavoprotein] + undecanoyl-CoA → reduced [electron-transfer flavoprotein] + trans-2-undecenoyl-CoA; | ECSIT (5), NDUFS3 (5), FAF2 (4), NDUFAF1 (4), TIMMDC1 (4), DBT (3), EYA2 (3), NDUFA13 (3), NDUFAF4 (3), NDUFS7 (3), PDK1 (3), COA1 (2), DLAT (2), ERAL1 (2), GLS (2), GRSF1 (2), HSPD1 (2), ICT1 (2), MRPL1 (2), MTRF1L (2), NDUFA2 (2), NDUFA8 (2), NDUFAF3 (2), NDUFS2 (2), NDUFS5 (2), NFS1 (2), P2RY12 (2), PDK2 (2), PDK3 (2), PDSS1 (2), PNPT1 (2), SDHA (2), SIRT3 (2), TRMT61B (2) | mitochondrion (5), nucleus (5) | None |

(9Z)-hexadecenoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E,9Z)-hexadecadienoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + heptadecanoyl-CoA + oxidized [electron-transfer flavoprotein] → reduced [electron-transfer flavoprotein] + trans-2-heptadecenoyl-CoA; (9E)-octadecenoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E,9E)-octadecadienoyl-CoA + reduced [electron-transfer flavoprotein]; (9Z)-octadecenoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E,9Z)-octadecadienoyl-CoA + reduced [electron-transfer flavoprotein]; (9Z,12Z)-octadecadienoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E,9Z,12Z)-octadecatrienoyl-CoA + reduced [electron-transfer flavoprotein]; (4Z,7Z,10Z,13Z,16Z,19Z)-docosahexaenoyl-CoA + H$^+$ + oxidized [electron-transfer flavoprotein] → (2E,4Z,7Z,10Z,13Z,16Z,19Z)-docosaheptaenoyl-CoA + reduced [electron-transfer flavoprotein]; H$^+$ + oxidized [electron-transfer flavoprotein] + tetradecanoyl-CoA → (2E)-tetradecenoyl-CoA + reduced [electron-transfer flavoprotein]

| 27. Acyl-CoA dehydrogenase family member 10 (ACAD10; Q6JQN1) | Acyl-CoA dehydrogenase only active with R- and S-2-methyl-C15-CoA./ A + a 2,3-saturated acyl-CoA → a 2,3-dehydroacyl-CoA + AH$_2$ | BSG (2), NDUFA10 (2), P2RY8 (2), PTH1R (2) | Mitochondrion (5) | None. |
| 28. Acyl-CoA dehydrogenase family member 11 (ACAD11; Q709F0) | Acyl-CoA dehydrogenase, that exhibits maximal activity towards saturated C22-CoA. Probably participates in beta-oxydation and energy production but could also play a role in the metabolism of specific fatty acids to control fatty acids composition of cellular lipids in brain./ a 2,3-saturated acyl-CoA + H$^+$ | CD2BP2 (3), DUSP19 (2), FAM9A (2), GTSE1 (2), KCTD17 (2), MIER2 (2), MTMR4 (2), PDYN (2), PEX5 (2), PTPN14 (2), SDF4 (2), TCL1B (2), TRAK2 (2), TTC4 (2), UBXN7 (2) | Mitochondrion (5), nucleus (5), peroxisome (4) | None. |

+ oxidized [electron-transfer flavoprotein] $\rightarrow$ a ($2E$)-enoyl-CoA + reduced [electron-transfer flavoprotein]; docosanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ ($2E$)-docosenoyl-CoA + reduced [electron-transfer flavoprotein]; H^+ + oxidized [electron-transfer flavoprotein] + tetracosanoyl-CoA $\rightarrow$ ($2E$)-tetracosenoyl-CoA + reduced [electron-transfer flavoprotein]; eicosanoyl-CoA + H^+ + oxidized [electron-transfer flavoprotein] $\rightarrow$ ($2E$)-eicosenoyl-CoA + reduced [electron-transfer flavoprotein]; H^+ + hexacosanoyl-CoA + oxidized [electron-transfer flavoprotein] $\rightarrow$ ($2E$)-hexacosenoyl-CoA + reduced [electron-transfer flavoprotein]; H+ + oxidized [electron-transfer flavoprotein] + tricosanoyl-CoA $\rightarrow$ ($2E$)-tricosenoyl-CoA + reduced [electron-transfer flavoprotein]

29. D-aspartate oxidase (DDO; Q99489)	Selectively catalyzes the oxidative deamination of D-aspartate and its N-methylated derivative, N-methyl D-aspartate./ D-aspartate + H_2O + O_2 $\rightarrow$ H_2O_2 + NH_4^+ + oxaloacetate **EC:1.4.3.1**.	PEX5 (2)	Peroxisome (5), cytosol (5)	None
30. L-amino-acid oxidase (IL4I1; Q96RQ9).	Secreted L-amino-acid oxidase that acts as a key immunoregulator. Has preference for L-aromatic amino acids: converts phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp) to phenylpyruvic acid (PP), hydroxyphenylpyruvic acid (HPP), and indole-3-pyruvic acid (I3P), respectively. Also has weak L-arginine oxidase activity. Acts as a negative regulator of anti-tumor immunity by mediating Trp degradation via an indole pyruvate pathway that activates the transcription factor AHR. IL4I1-mediated Trp catabolism generates I3P, giving rise to indole metabolites (indole-3-acetic acid (IAA) and indole-3-aldehyde (I3A)) and kynurenic acid, which act as	None	Extracellular (4), lysosome (4), plasma membrane (3)	None

ligands for AHR, a ligand-activated transcription factor that plays important roles in immunity and cancer. AHR activation by indoles following IL4I1-mediated Trp degradation enhances tumor progression by promoting cancer cell motility and suppressing adaptive immunity. Also has an immunoregulatory function in some immune cells, probably by mediating Trp degradation and promoting downstream AHR activation: inhibits T-cell activation and proliferation, promotes the differentiation of naive CD4+ T-cells into FOXP3+ regulatory T-cells (Treg) and regulates the development and function of B-cells. Also regulates M2 macrophage polarization by inhibiting T-cell activation. Also has antibacterial properties by inhibiting growth of Gram negative and Gram positive bacteria through the production of NH_4^+ and H_2O_2./ an L-α-amino acid + H_2O + O_2 → a 2-oxocarboxylate + H_2O_2 + NH_4^+ **EC:1.4.3.2;** H_2O + L-tryptophan + O_2 → H_2O_2 + indole-3-pyruvate + NH_4^+; H_2O + L-phenylalanine + O_2 → 3-phenylpyruvate + H_2O_2 + NH_4^+;

H_2O + L-tyrosine + O_2 → 3-(4-hydroxyphenyl)pyruvate + H_2O_2 + NH_4^+ ;

H_2O + L-arginine + O_2 → 5-guanidino-2-oxopentanoate + H_2O_2 + NH_4^+

31. D-amino-acid oxidase (DAO; P14920)	Regulates the level of the neuromodulator D-serine in the brain. Has high activity towards D-DOPA and contributes to dopamine synthesis. Could act as a detoxifying agent which removes D-amino acids accumulated during aging. Acts on a variety of D-amino acids with a preference for those having small hydrophobic side chains followed by those bearing polar, aromatic, and basic groups. Does not act on	DAOA (3), PRKAB2 (3)	Peroxisome (5), cytosol (5), extracellular (3), nucleus (3)	None

	acidic amino acids./ a D-α-amino acid + H_2O + O_2 $\rightarrow$ a 2-oxocarboxylate + H_2O_2 + NH_4^+ **EC:1.4.3.3**			
32. Amine oxidase [flavin-containing] A (MAOA; P21397)	Catalyzes the oxidative deamination of biogenic and xenobiotic amines and has important functions in the metabolism of neuroactive and vasoactive amines in the central nervous system and peripheral tissues. MAOA preferentially oxidizes biogenic amines such as 5-hydroxytryptamine (5-HT), norepinephrine and epinephrine./ a secondary aliphatic amine + H_2O + O_2 $\rightarrow$ a primary amine + an aldehyde + H_2O_2 **EC:1.4.3.4**	MAOB (2)	mitochondrion (5), cytosol (5), nucleus (3)	None
33. Amine oxidase [flavin-containing] B (MAOB; P27338)	Catalyzes the oxidative deamination of biogenic and xenobiotic amines and has important functions in the metabolism of neuroactive and vasoactive amines in the central nervous system and peripheral tissues. MAOB preferentially degrades benzylamine and phenylethylamine./ a secondary aliphatic amine + H_2O + O_2 $\rightarrow$ a primary amine + an aldehyde + H_2O_2 **EC:1.4.3.4**	MAOA (2)	mitochondrion (5), nucleus (3), cytosol (3)	None
34. Pyridoxine-5'-phosphate oxidase (PNPO; Q9NVS9)	Catalyzes the oxidation of either pyridoxine 5'-phosphate (PNP) or pyridoxamine 5'-phosphate (PMP) into pyridoxal 5'-phosphate (PLP)./ H_2O + O_2 + pyridoxamine 5'-phosphate $\rightarrow$ H_2O_2 + NH_4^+ + pyridoxal 5'-phosphate **EC:1.4.3.5**; O_2 + pyridoxine 5'-phosphate $\rightarrow$ H_2O_2 + pyridoxal 5'-phosphate **EC:1.4.3.5**	AGTRAP (2), ARL6IP6 (2), FBXO25 (2), MOK (2), MTERF1 (2)	nucleus (5), cytosol (5), mitochondrion (3)	None
35. Renalase (RNLS; Q5VYX0)	Catalyzes the oxidation of the less abundant 1,2-dihydro-beta-NAD(P) and 1,6-dihydro-beta-NAD(P) to form beta-NAD(P)+. The enzyme hormone is secreted by the kidney, and circulates in blood and	None	Extracellular (5)	None

modulates cardiac function and systemic blood pressure. Lowers blood pressure in vivo by decreasing cardiac contractility and heart rate and preventing a compensatory increase in peripheral vascular tone, suggesting a causal link to the increased plasma catecholamine and heightened cardiovascular risk. High concentrations of catecholamines activate plasma renalase and promotes its secretion and synthesis./1,2-dihydro-β-NAD + H$^+$ + O$_2$ → H$_2$O$_2$ + NAD$^+$

EC:1.6.3.5; 1,2-dihydro-β-NADP + H$^+$ + O$_2$ → H$_2$O$_2$ + NADP$^+$ **EC:1.6.3.5**; 1,6-dihydro-β-NADP + H$^+$ + O$_2$ → H$_2$O$_2$ + NADP$^+$ **EC:1.6.3.5**; 1,6-dihydro-β-NAD + H$^+$ + O$_2$ → H$_2$O$_2$ + NAD$^+$ **EC:1.6.3.5**

36. Methylenetetrahydrofolate reductase (MTHFR; P42898)	Catalyzes the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, a co-substrate for homocysteine remethylation to methionine. /(6*S*)-5-methyl-5,6,7,8-tetrahydrofolate + NADP$^+$ → (6*R*)-5,10-methylene-5,6,7,8-tetrahydrofolate + H$^+$ + NADPH **EC:1.5.1.20**; (6*S*)-5-methyl-5,6,7,8-tetrahydrofolate + NAD$^+$ → (6*R*)-5,10-methylene-5,6,7,8-tetrahydrofolate + H$^+$ + NADH **EC:1.5.1.20**.	BCL2L11 (2), CDCA8 (2), GNAI1 (2), GNAZ (2), RASSF1 (2), SMPD2 (2)	Cytosol (4), plasma membrane (3), extracellular (3), nucleus (3)	None
37. Peroxisomal sarcosine oxidase (PIPOX; Q9P0Z9)	Metabolizes sarcosine, L-pipecolic acid and L-proline./ H$_2$O + O$_2$ + sarcosine → formaldehyde + glycine + H$_2$O$_2$ **EC:1.5.3.1**; L-pipecolate + O$_2$ → H$^+$ + H$_2$O$_2$ + L-1-piperideine-6-carboxylate **EC:1.5.3.7**.	ELP4 (2), ELP5 (2), ELP (6), FKBP5 (2), NAPG (2), ZNF217 (2)	Peroxisome (5), cytosol (4)	None
38. Spermine oxidase (SMOX; Q9NWM0)	Flavoenzyme which catalyzes the oxidation of spermine to spermidine. Can also use N$_1$-acetylspermine and spermidine as substrates, with different affinity depending on the isoform (isozyme) and on the experimental conditions. Plays an important role in the regulation of polyamine	None	Nucleus (5), cytosol (5), plasma membrane (3), extracellular (3)	None

intracellular concentration and has the potential to act as a determinant of cellular sensitivity to the antitumor polyamine analogs. May contribute to beta-alanine production via aldehyde dehydrogenase conversion of 3-amino-propanal./$H_2O + O_2 +$ spermine $\rightarrow$ 3-aminopropanal $+ H_2O_2 +$ spermidine **EC:1.5.3.16**

39. Electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial (ETFDH; Q16134)	Accepts electrons from ETF and reduces ubiquinone./ a ubiquinone + reduced [electron-transfer flavoprotein] $\rightarrow$ a ubiquinol + H^+ + oxidized [electron-transfer flavoprotein] **EC:1.5.5.1**	None	Mitochondrion (5)	None
40. Electron transfer flavoprotein subunit alpha, mitochondrial (ETFA; P13804)	Heterodimeric electron transfer flavoprotein that accepts electrons from several mitochondrial dehydrogenases, including acyl-CoA dehydrogenases, glutaryl-CoA and sarcosine dehydrogenase. It transfers the electrons to the main mitochondrial respiratory chain via ETF-ubiquinone oxidoreductase (ETF dehydrogenase). Required for normal mitochondrial fatty acid oxidation and normal amino acid metabolism.	ETFB (5), LYRM5 (5), C8ORF82 (3), AGR2 (2), ATP5A1 (2), COX15 (2), CRYZ (2), GRSF1 (2), HINT2 (2), ICT1 (2), IVD (2)	Mitochondrion (5)	None
41. Electron transfer flavoprotein subunit beta (ETFB; P38117)	Heterodimeric electron transfer flavoprotein that accepts electrons from several mitochondrial dehydrogenases, including acyl-CoA dehydrogenases, glutaryl-CoA and sarcosine dehydrogenase. It transfers the electrons to the main mitochondrial respiratory chain via ETF-ubiquinone oxidoreductase. Required for normal mitochondrial fatty acid oxidation and normal amino acid metabolism. ETFB binds an AMP molecule that probably has a purely structural role.	ETFA (5), LYRM5 (4),CRYZ (2), GRSF1 (2), LRPPRC (2)	Mitochondrion (5)	None
42. Sarcosine dehydrogenase, mitochondrial	$H^+ + H_2O +$ oxidized [electron-transfer flavoprotein] + sarcosine $\rightarrow$ formaldehyde + glycine + reduced [electron-transfer flavoprotein]	None	Mitochondrion (5)	None

(SARDH, Q9UL12)	**EC:1.5.8.3**			
43. Dimethylglycine dehydrogenase, mitochondrial (DMGDH, Q9UI17)	Catalyzes the demethylation of N,N-dimethylglycine to sarcosine. Also has activity with sarcosine in vitro. / (6S)-5,6,7,8-tetrahydrofolyl-(γ-L-Glu)n + H$^+$ + N,N-dimethylglycine + oxidized [electron-transfer flavoprotein] $\rightarrow$ (6R)-5,10-methylenetetrahydrofolyl-(γ-L-Glu)(n) + reduced [electron-transfer flavoprotein] + sarcosine **EC:1.5.8.4**	None	Mitochondrion (5)	None
44. Lysine-specific histone demethylase 1A (KDM1A; O60341)	Histone demethylase that can demethylate both 'Lys-4' (H3K4me) and 'Lys-9' (H3K9me) of histone H3, thereby acting as a coactivator or a corepressor, depending on the context. Acts by oxidizing the substrate by FAD to generate the corresponding imine that is subsequently hydrolyzed. Acts as a corepressor by mediating demethylation of H3K4me, a specific tag for epigenetic transcriptional activation. Demethylates both mono- (H3K4me1) and di-methylated (H3K4me2) H3K4me (. May play a role in the repression of neuronal genes. Alone, it is unable to demethylate H3K4me on nucleosomes and requires the presence of RCOR1/CoREST to achieve such activity. Also acts as a coactivator of androgen receptor (AR)-dependent transcription, by being recruited to AR target genes and mediating demethylation of H3K9me, a specific tag for epigenetic transcriptional repression. The presence of PRKCB in AR-containing complexes, which mediates phosphorylation of 'Thr-6' of histone H3 (H3T6ph), a specific tag that prevents demethylation H3K4me, prevents H3K4me demethylase activity of KDM1A. Demethylates di-methylated 'Lys-370' of p53/TP53 which prevents interaction of p53/TP53 with TP53BP1 and represses p53/TP53-mediated	RCOR1 (23), HDAC1 (21), HDAC2 (14), HMG20B (9), SNAI1 (9), CTBP1 (8), HIST1H3A (8), ZNF217 (7), ATG16L1 (6), HMG20A (6), PHF21A (6), TP53 (5), MTA2 (5), USP7 (5), HDAC3 (4), MYC (4), PPP1R12A (4), RCOR3 (4), RIOK1 (4), USP28 (4), ZMYM2 (4), KDM1A (3), BRCA1 (3), CARM1 (3), CSNK2A1 (3), CSNK2A2 (3), DNMT1 (3), ESR1 (3), EZH2 (3), GATA3 (3), GFI1B (3), GSE1 (3), MBD3 (3), MTA1 (3), MTA3 (3), NMI (3), NR2C2 (3), NR2E1 (3), PELP1 (3), PPM1D (3), RNF168 (3), RPRD1A (3), SOCS6 (3), TAL1 (3), TRIM39 (3), USP22 (3), ZMYM3 (3), AKAP9 (2), AR (2), ARRDC1 (2), C15ORF27 (2), C8ORF74 (2), CBFA2T3 (2), CCDC14 (2), CCDC151 (2), CCDC53 (2), CDCA4 (2), CEP70 (2), CHD3 (2), DNAJA3 (2), E2F1	Nucleus (5), extracellular (3), cytosol (3)	None.

54

transcriptional activation. Demethylates and stabilizes the DNA methylase DNMT1. Required for gastrulation during embryogenesis. Component of a RCOR/GFI/KDM1A/HDAC complex that suppresses, via histone deacetylase (HDAC) recruitment, a number of genes implicated in multilineage blood cell development. Effector of SNAI1-mediated transcription repression of E-cadherin/CDH1, CDN7 and KRT8. Required for the maintenance of the silenced state of the SNAI1 target genes E-cadherin/CDH1 and CDN7./ 2 A + 2 H_2O + N^6,N^6-dimethyl-L-lysyl^4-[histone H3] $\rightarrow$ 2 AH_2 + 2 formaldehyde + L-lysyl4-[histone H3] **EC:1.14.99.66**

(2), ECE2 (2), FAM9A (2), FYCO1 (2), GOLGA2 (2), GSK3B (2), GSTCD (2), GTF2I (2), GTPBP2 (2), HAUS1 (2), HAUS6 (2), HESX1 (2), HIST2H3C (2), IMMT (2), INSM1 (2), KDM5B (2), KIFC3 (2), KLHDC4 (2), KRT33B (2), NECAB2 (2), NOTCH1 (2), NR1H2 (2), NR1H3 (2), NR2C1 (2), OFCC1 (2), OIP5 (2), PHF19 (2), PPARD (2), PRDM1 (2), PSMC1 (2), RASSF1 (2), RASSF2 (2), RBBP4 (2), RBPJ (2), SAMD3 (2), SERGEF (2), SF3B2 (2), SMAD9 (2), SMARCC1 (2), SNF8 (2), SOX2 (2), SPICE1 (2), SPSB1 (2), STAT3 (2), STX19 (2), TEX9 (2), TNNT2 (2), TRAF4 (2), TRIM28 (2), YEATS4 (2), ZBED1 (2)

| 45. Proline dehydrogenase 1, mitochondrial (PRODH; O43272) | Converts proline to delta-1-pyrroline-5-carboxylate./ a quinone + L-proline $\rightarrow$ (*S*)-1-pyrroline-5-carboxylate + a quinol + H^+ **EC:1.5.5.2** | None | Mitochondrion (5), cytosol (3) | None |
| 46. NADH-cytochrome b5 reductase 3 (CYB5R3; P00387) | 2 [Fe(III)-cytochrome b_5] + NADH $\rightarrow$ 2 [Fe(II)-cytochrome b_5] + H^+ + NAD^+ **EC:1.6.2.2** | RTN4 (3), ACTL6A (2), CPVL (2), PDHA1 (2), SDHA (2). | Extracellular (5), mitochondrion (5), endoplasmic reticulum (5), cytosol (4), lysosome (4) | None |

55

47. NADPH-cytochrome P450 reductase (POR; P16435)	This enzyme is required for electron transfer from NADP to cytochrome P450 in microsomes. It can also provide electron transfer to heme oxygenase and cytochrome B5./ NADPH + 2 oxidized [cytochrome P450] $\rightarrow$ H$^+$ + NADP$^+$ + 2 reduced [cytochrome P450] **EC:1.6.2.4**	CYP2E1 (3), FANCC (3), UBXN6 (3), CYP1A2 (2), CYP2C19 (2), CYP2C9 (2).	Endoplasmic reticulum (5), cytosol (4), mitochondrion (3), nucleus (3)	None
48. NAD(P)H dehydrogenase [quinone] 1 (NQO1; P15559)	The enzyme apparently serves as a quinone reductase in connection with conjugation reactions of hydroquinons involved in detoxification pathways as well as in biosynthetic processes such as the vitamin K-dependent gamma-carboxylation of glutamate residues in prothrombin synthesis./ a quinone + H$^+$ + NADH $\rightarrow$ a quinol + NAD$^+$ **EC:1.6.5.2;** a quinone + H$^+$ + NADPH $\rightarrow$ a quinol + NADP$^+$ **EC:1.6.5.2**	TP53 (5), FOS (3), HIF1A (3), STUB1 (3), BAG3 (2).	Cytosol (5), mitochondrion (3), nucleus (3)	None
49. NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial (NDUFV1; P49821)	Core subunit of the mitochondrial membrane respiratory chain NADH dehydrogenase (Complex I) that is believed to belong to the minimal assembly required for catalysis. Complex I functions in the transfer of electrons from NADH to the respiratory chain. The immediate electron acceptor for the enzyme is believed to be ubiquinone. / a ubiquinone + 5 H$^+$(in) + NADH = a ubiquinol + 4 H$^+$(out) + NAD$^+$ **EC:7.1.1.2**	NDUFS3 (6), NDUFA9 (4), NDUFS1 (4), NDUFS2 (4), NDUFA8 (3), NDUFS7 (3), NDUFS8 (3), NDUFV2 (3), C6ORF203 (2), COA3 (2), COX5A (2), CPT1A (2), GRSF1 (2), HSCB (2), ICT1 (2), LONP1 (2), LRPPRC (2), NDUFA2 (2), NDUFS4 (2), NDUFS5 (2), NDUFS6 (2), NDUFV3 (2), PDHA1 (2), PHB2 (2), POR (2), SDHA (2), TOMM40 (2), TUFM (2), VDAC2 (2)	Mitochondrion (5), cytosol (4)	None
50. NADPH-dependent diflavin oxidoreductase 1 (NDOR1; Q9UHB4)	Component of the cytosolic iron-sulfur (Fe-S) protein assembly (CIA) machinery. Required for the maturation of extramitochondrial Fe-S proteins. Part of an electron transfer chain functioning in an early	CIAPIN1 (4), MTUS2 (2), TCF4 (2), TCHP (2)	Nucleus (5), cytosol (5), cytoskeleton (4)	None

	step of cytosolic Fe-S biogenesis. Transfers electrons from NADPH to the Fe/S cluster of CIAPIN1.			
51. tRNA-dihydrouridine(20) synthase [NAD(P)$^+$]-like (DUS2; Q9NX74)	Dihydrouridine synthase. Catalyzes the NADPH-dependent synthesis of dihydrouridine, a modified base found in the D-loop of most tRNAs. Negatively regulates the activation of EIF2AK2/PKR./ 5,6-dihydrouridine^{20} in tRNA + NADP$^+$ → H$^+$ + NADPH + uridine^{20} in tRNA **EC:1.3.1.91**	None	Cytosol (5), endoplasmic reticulum (3)	None
52. Dihydrolipoyl dehydrogenase, mitochondrial (DLD; P09622)	Lipoamide dehydrogenase is a component of the glycine cleavage system as well as an E3 component of three alpha-ketoacid dehydrogenase complexes (pyruvate-, alpha-ketoglutarate-, and branched-chain amino acid-dehydrogenase complex). The 2-oxoglutarate dehydrogenase complex is mainly active in the mitochondrion. A fraction of the 2-oxoglutarate dehydrogenase complex also localizes in the nucleus and is required for lysine succinylation of histones: associates with KAT2A on chromatin and provides succinyl-CoA to histone succinyltransferase KAT2A. In monomeric form may have additional moonlighting function as serine protease. Involved in the hyperactivation of spermatazoa during capacitation and in the spermatazoal acrosome reaction (R)-N^6-dihydrolipoyl-L-lysyl-[protein] + NAD$^+$ → (R)-N^6-lipoyl-L-lysyl-[protein] + H$^+$ + NADH **EC:1.8.1.4**	OGDH (4), PDHA1 (4), DBT (3), DLST (3), PDHB (3), PDHX (3), AIFM1 (2), AK2 (2), BTRC (2), C7ORF55 (2), CKB (2), DDX39B (2), DHTKD1 (2), FBXW11 (2), HSD17B10 (2), HSDL2 (2), LDHA (2), MRPS36 (2), NFS1 (2), NUDT21 (2), OAT (2), OGDHL (2), PARS2 (2), PCBP3 (2), PDK3 (2), PSMD4 (2), SIAH1 (2), SLX4IP (2), SOCS6 (2), SORL1 (2), UBE2M (2)	Mitochondrion (5), nucleus (5), cytosol (3)	None
53. Glutathione reductase, mitochondrial (GSR; P00390)	Maintains high levels of reduced glutathione in the cytosol./ 2 glutathione + NADP$^+$ → glutathione disulfide + H$^+$ + NADPH **EC:1.8.1.7**	None	Extracellular (5), mitochondrion (5), cytosol (5), plasma membrane (4), nucleus (3)	None
54. Thioredoxin reductase 1, cytoplasmic	Isoform 1 may possess glutaredoxin activity as well as thioredoxin reductase activity and induces actin and	APEX1 (2), ESR1 (2), ISG15 (2), PDIA6 (2), RP2 (2), TRIM25 (2).	Nucleus (5), cytosol (5),	None

(TXNRD1; Q16881)	tubulin polymerization, leading to formation of cell membrane protrusions. Isoform 4 enhances the transcriptional activity of estrogen receptors alpha and beta while isoform 5 enhances the transcriptional activity of the beta receptor only. Isoform 5 also mediates cell death induced by a combination of interferon-beta and retinoic acid./ [thioredoxin]-dithiol + NADP$^+$ $\rightarrow$ [thioredoxin]-disulfide + H$^+$ + NADPH **EC:1.8.1.9**		extracellular (4), mitochondrion (4)
55. Thioredoxin reductase 2, mitochondrial (TXNRD2; Q9NNW7)	Involved in the control of reactive oxygen species levels and the regulation of mitochondrial redox homeostasis. Maintains thioredoxin in a reduced state. May play a role in redox-regulated cell signaling./ [thioredoxin]-dithiol + NADP$^+$ $\rightarrow$ [thioredoxin]-disulfide + H$^+$ + NADPH **EC:1.8.1.9**	GTF2H1 (3), PSMC3 (3), FOXR2 (2), NCOA6 (2), NCOR1 (2), PPP1R3C (2)	Mitochondrion (5), cytosol (5), peroxisome (3) / None
56. Thioredoxin reductase 3 (TXNRD3; Q86VQ6)	Displays thioredoxin reductase, glutaredoxin and glutathione reductase activities. Catalyzes disulfide bond isomerization. Promotes disulfide bond formation between GPX4 and various sperm proteins and may play a role in sperm maturation by promoting formation of sperm structural components./ [thioredoxin]-dithiol + NADP$^+$ $\rightarrow$ [thioredoxin]-disulfide + H$^+$ + NADPH **EC:1.8.1.9**	None	Mitochondrion (4), nucleus (4), endoplasmic reticulum (4), cytosol (4), extracellular (3) / None
57. FAD-dependent oxidoreductase domain-containing protein 2 (FOXRED2; Q8IWF2).	Probable flavoprotein which may function in endoplasmic reticulum associated degradation (ERAD). May bind non-native proteins in the endoplasmic reticulum and target them to the ubiquitination machinery for subsequent degradation	SEL1L (4), OS9 (3), TXNDC16 (2)	Endoplasmic reticulum (5) / None
58. FAD-linked sulfhydryl oxidase ALR (GFER; P55789)	FAD-dependent sulfhydryl oxidase that regenerates the redox-active disulfide bonds in CHCHD4/MIA40, a chaperone essential for disulfide bond formation and protein folding in the mitochondrial intermembrane space. The reduced form of CHCHD4/MIA40 forms a	COPS5 (6), BNIPL (4), AIFM1 (2), AMMECR1 (2), CDK4 (2), KLHL20 (2), PGPEP1 (2), TUBB2A (2)	Extracellular (5), mitochondrion (5), cytosol (5) / None

	transient intermolecular disulfide bridge with GFER/ERV1, resulting in regeneration of the essential disulfide bonds in CHCHD4/MIA40, while GFER/ERV1 becomes re-oxidized by donating electrons to cytochrome c or molecular oxygen./ O_2 + 2 R'C(R)SH → H_2O_2 + R'C(R)S-S(R)CR' **EC:1.8.3.2**			
59. Prenylcysteine oxidase 1 (PCYOX1;Q9UHG3)	Involved in the degradation of prenylated proteins. Cleaves the thioether bond of prenyl-L-cysteines, such as farnesylcysteine and geranylgeranylcysteine./ an S-prenyl-L-cysteine + H_2O + O_2 → a prenal + H_2O_2 + L-cysteine **EC:1.8.3.5**	ANTXR1 (2), BLVRA (2), DLK1 (2), PLAUR (2), UBAC2 (2)	Lysosome (5), extracellular (4)	None
60.Ribosyldihydronicotina mide dehydrogenase [quinone] (NQO2; P16083)	The enzyme apparently serves as a quinone reductase in connection with conjugation reactions of hydroquinones involved in detoxification pathways as well as in biosynthetic processes such as the vitamin K-dependent gamma-carboxylation of glutamate residues in prothrombin synthesis./ 1-(β-D-ribofuranosyl)-1,4-dihydronicotinamide + a quinone + H^+ → a quinol + β-nicotinamide D-riboside **EC:1.10.5.1**	GORASP2 (2), LRRC7 (2)	Nucleus (5), cytosol (5), extracellular (4)	4qod.
61. Dimethylaniline monooxygenase [N-oxide-forming] 1(FMO1; Q01740)	This protein is involved in the oxidative metabolism of a variety of xenobiotics such as drugs and pesticides. Form I catalyzes the N-oxygenation of secondary and tertiary amines. / H^+ + *N,N*-dimethylaniline + NADPH + O_2 → H_2O + *N,N*-dimethylaniline *N*-oxide + $NADP^+$ **EC:1.14.13.8**	None	Endoplasmic reticulum (4), peroxisome (3)	None
62. Dimethylaniline monooxygenase [N-oxide-forming] 2 (FMO2; Q99518)	Catalyzes the N-oxidation of certain primary alkylamines to their oximes via an N-hydroxylamine intermediate. Inactive toward certain tertiary amines, such as imipramine or chloropromazine. Can catalyze the S-oxidation of methimazole./ H^+ + *N,N*-dimethylaniline + NADPH + O_2 → H_2O + *N,N*-dimethylaniline *N*-oxide + $NADP^+$ **EC:1.14.13.8**	None	Endoplasmic reticulum (4)	None

63. Dimethylaniline monooxygenase [N-oxide-forming] 3 (FMO3; P31513)	Essential hepatic enzyme that catalyzes the oxygenation of a wide variety of nitrogen- and sulfur-containing compounds including drugs as well as dietary compounds. Plays an important role in the metabolism of trimethylamine (TMA), via the production of trimethylamine N-oxide (TMAO) metabolite. TMA is generated by the action of gut microbiota using dietary precursors such as choline, choline containing compounds, betaine or L-carnitine. By regulating TMAO concentration, FMO3 directly impacts both platelet responsiveness and rate of thrombus formation./ H^+ + *N,N*-dimethylaniline + NADPH + O_2 → H_2O + *N,N*-dimethylaniline N-oxide + $NADP^+$ **EC:1.14.13.8**; NADPH + O_2 + trimethylamine → H_2O + $NADP^+$ + trimethylamine N-oxide **EC:1.14.13.148**; fenbendazole + O_2 + reduced [NADPH—hemoprotein reductase] → fenbendazole S-oxide + H+ + H_2O + oxidized [NADPH—hemoprotein reductase] **EC:1.14.14.73;** albendazole + O_2 + reduced [NADPH—hemoprotein reductase] → albendazole S-oxide + H^+ + H_2O + oxidized [NADPH—hemoprotein reductase] **EC:1.14.14.73**; (*S*)-nicotine + NADPH + O_2 → H_2O + $NADP^+$ + trans-*(S)*-nicotine $N^{1'}$-oxide	None	Endoplasmic reticulum (4), cytosol (4), plasma membrane (3), nucleus (3)	None
64. Dimethylaniline monooxygenase [N-oxide-forming] 4 (FMO4; P31512)	This protein is involved in the oxidative metabolism of a variety of xenobiotics such as drugs and pesticides./ H^+ + *N,N*-dimethylaniline + NADPH + O_2 → H_2O + *N,N*-dimethylaniline *N*-oxide + $NADP^+$ **EC:1.14.13.8**	None	Endoplasmic reticulum (3)	None
65. Flavin-containing monooxygenase 5 (FMO5;P49326)	Acts as Baeyer-Villiger monooxygenase on a broad range of substrates. Catalyzes the insertion of an oxygen atom into a carbon-carbon bond adjacent to a carbonyl, which converts ketones to esters. Active on	None	Endoplasmic reticulum (5), cytosol (4)	None

diverse carbonyl compounds, whereas soft nucleophiles are mostly non- or poorly reactive. In contrast with other forms of FMO it is non- or poorly active on 'classical' substrates such as drugs, pesticides, and dietary components containing soft nucleophilic heteroatoms (Probable). Able to oxidize drug molecules bearing a carbonyl group on an aliphatic chain, such as nabumetone and pentoxifylline. Also, in the absence of substrates, shows slow but yet significant NADPH oxidase activity. Acts as a positive modulator of cholesterol biosynthesis as well as glucose homeostasis, promoting metabolic aging via pleiotropic effects./ H^+ + N,N-dimethylaniline + NADPH + $O_2 \rightarrow H_2O$ + N,N-dimethylaniline N-oxide + $NADP^+$ **EC:1.14.13.8**; H^+ + NADPH + $O_2 \rightarrow H_2O_2$ + $NADP^+$ **EC:1.6.3.1**; H^+ + heptan-2-one + NADPH + $O_2 \rightarrow H_2O$ + $NADP^+$ + pentyl acetate; H^+ + NADPH + O_2 + octan-3-one $\rightarrow H_2O$ + $NADP^+$ + pentyl propanoate; H^+ + NADPH + O_2 + octan-3-one $\rightarrow$ ethyl hexanoate + H_2O + $NADP^+$; H^+ + hexan-3-one + NADPH + $O_2 \rightarrow$ ethyl butanoate + H_2O + $NADP^+$; H^+ + hexan-3-one + NADPH + O_2 $\rightarrow H_2O$ + $NADP^+$ + propyl propanoate; H^+ + heptan-4-one + NADPH + $O_2 \rightarrow H_2O$ + $NADP^+$ + propyl butanoate; geranial + H^+ + NADPH + $O_2 \rightarrow (1E)$-2,6-dimethylhepta-1,5-dien-1-yl formate + H_2O + $NADP^+$; H^+ + NADPH + O_2 + sulcatone $\rightarrow$ 4-methylpent-3-en-1-yl acetate + H_2O + $NADP^+$.

| 66. Kynurenine 3-monooxygenase (KMO;O15229) | Catalyzes the hydroxylation of L-kynurenine (L-Kyn) to form 3-hydroxy-L-kynurenine (L-3OHKyn). Required for synthesis of quinolinic acid, a neurotoxic NMDA receptor antagonist and potential endogenous | None | Mitochondrion (4), cytosol (4), plasma membrane (3), extracellular (3) | None |

	inhibitor of NMDA receptor signaling in axonal targeting, synaptogenesis and apoptosis during brain development. Quinolinic acid may also affect NMDA receptor signaling in pancreatic beta cells, osteoblasts, myocardial cells, and the gastrointestinal tract (Probable)./ H^+ + L-kynurenine + NADPH + $O_2 \rightarrow$ 3-hydroxy-L-kynurenine + H_2O + $NADP^+$ **EC:1.14.13.9**		
67. Nitric oxide synthase, brain (NOS1; P29475)	Produces nitric oxide (NO) which is a messenger molecule with diverse functions throughout the body. In the brain and peripheral nervous system, NO displays many properties of a neurotransmitter. Probably has nitrosylase activity and mediates cysteine S-nitrosylation of cytoplasmic target proteins such SRR./ H^+ + 2 L-arginine + 3 NADPH + 4 $O_2 \rightarrow$ 4 H_2O + 2 L-citrulline + 3 $NADP^+$ + 2 NO **EC:1.14.13.39**	DLG4 (5), PRKD1 (4), STUB1 (4). NOS1AP (3), SOX2 (2), ZDHHC23 (2)	Plasma membrane (5), nucleus (5) cytoskeleton (4), mitochondrion (4), endoplasmic reticulum (4), cytosol (4)	None
68. Nitric oxide synthase, inducible (NOS2; P35228)	Produces nitric oxide (NO) which is a messenger molecule with diverse functions throughout the body. In macrophages, NO mediates tumoricidal and bactericidal actions. Also has nitrosylase activity and mediates cysteine S-nitrosylation of cytoplasmic target proteins such PTGS2/COX2. As component of the iNOS-S100A8/9 transnitrosylase complex involved in the selective inflammatory stimulus-dependent S-nitrosylation of GAPDH on 'Cys-247' implicated in regulation of the GAIT complex activity and probably multiple targets including ANXA5, EZR, MSN and VIM. Involved in inflammation, enhances the synthesis of proinflammatory mediators such as IL6 and IL8./ H^+ + 2 L-arginine + 3 NADPH + 4 $O_2 \rightarrow$ 4 H_2O + 2 L-citrulline + 3 $NADP^+$ + 2 NO **EC:1.14.13.39**	SPSB2 (9), CALM1 (5), SPSB1 (5), CAV1 (2), CUL5 (2), FBXO45 (2), HSP90AA1 (2), RAC2 (2), SPSB4 (2), UCHL5 (2)	Peroxisome (5), plasma membrane (4), nucleus (4), cytosol (4), extracellular (3), cytoskeleton (3), mitochondrion (3)	None

69. Nitric oxide synthase, endotelial (NOS3; P29474)	Produces nitric oxide (NO) which is implicated in vascular smooth muscle relaxation through a cGMP-mediated signal transduction pathway. NO mediates vascular endothelial growth factor (VEGF)-induced angiogenesis in coronary vessels and promotes blood clotting through the activation of platelets./ $H^+ + 2$ L-arginine $+ 3$ NADPH $+ 4 O_2 \rightarrow 4 H_2O + 2$ L-citrulline $+ 3$ NADP$^+$ $+ 2$ NO **EC:1.14.13.39**	HSP90AA1 (7), CALM1 (6), CAV1 (4), NOSIP (4), NOSTRIN (4), ACTB (3), AKT1 (3), CDC37 (3), APOE (2), APP (2), GCDH (2), GOLGA2 (2), GUCY1B3 (2), PPP2R4 (2), ST13 (2).	Plasma membrane (5), cytoskeleton (5), golgi apparatus (5), nucleus (4), cytosol (4), extracellular (3), mitochondrion (3), endoplasmic reticulum (3)	None
70. Squalene monooxygenase (SQLE; Q14534)	Catalyzes the stereospecific oxidation of squalene to (S)-2,3-epoxysqualene, and is considered to be a rate-limiting enzyme in steroid biosynthesis./ O_2 + reduced [NADPH—hemoprotein reductase] + squalene $\rightarrow$ (S)-2,3-epoxysqualene $+ H^+ + H_2O$ + oxidized [NADPH—hemoprotein reductase] **EC:1.14.14.17**	FAF2 (2), MARCH6 (2), TREML2 (2)	Plasma membrane (4), endoplasmic reticulum (4), cytoskeleton (3), cytosol (3)	None
71. Ubiquinone biosynthesis monooxygenase COQ6, mitochondrial (COQ6; Q9Y2Z9)	FAD-dependent monooxygenase required for the C5-ring hydroxylation during ubiquinone biosynthesis. Catalyzes the hydroxylation of 3-hexaprenyl-4-hydroxybenzoic acid (HHB) to 3-hexaprenyl-4,5-dihydroxybenzoic acid (DHHB). The electrons required for the hydroxylation reaction may be funneled indirectly from NADPH via a ferredoxin/ferredoxin reductase system to COQ6 (By similarity). Is able to perform the deamination reaction at C4 of 3-hexaprenyl-4-amino-5-hydroxybenzoic acid (HHAB) to produce DHHB when expressed in yeast cells lacking COQ9, even if utilization of para-aminobenzoic acid (pABA) involving C4-deamination seems not to occur in bacteria, plants and mammals, where only C5 hydroxylation of HHB has been shown./ 4-hydroxy-3-*all-trans*-hexaprenylbenzoate $+ 2 H^+ + O_2 + 2$ reduced [2Fe-2S]-[ferredoxin] $\rightarrow$ 3,4-dihydroxy-5-*all-trans*-	COQ3 (3), AGR2 (2), COQ5 (2), COQ9 (2)	Mitochondrion (5), golgi apparatus (3)	None

	hexaprenylbenzoate + H_2O + 2 oxidized [2Fe-2S]-[ferredoxin]			
72. Metalloreductase STEAP3 (STEAP3; Q658P3)	Endosomal ferrireductase required for efficient transferrin-dependent iron uptake in erythroid cells. Participates in erythroid iron homeostasis by reducing Fe^{3+} to Fe^{2+}. Can also reduce of Cu^{2+} to Cu^{1+}, suggesting that it participates in copper homeostasis. Endosomal ferrireductase required for efficient transferrin-dependent iron uptake in erythroid cells. Participates in erythroid iron homeostasis by reducing Fe^{3+} to Fe^{2+}. Can also reduce of Cu^{2+} to Cu^{1+}, suggesting that it participates in copper homeostasis.	NRAS (3), BNIP3L (2), BSG (2), DLK1 (2), ENPP6 (2), FAM177A1 (2), HRAS (2), KRAS (2), NT5E (2), PKMYT1 (2), SLC39A4 (2), TEX29 (2), TMEFF1 (2)	Endosome (5), plasma membrane (4)	None
73. Methionine synthase reductase (MTRR; Q9UBK8)	Key enzyme in methionine and folate homeostasis responsible for the reactivation of methionine synthase (MTR/MS) activity by catalyzing the reductive methylation of MTR-bound cob(II)alamin. Cobalamin (vitamin B12) forms a complex with MTR to serve as an intermediary in methyl transfer reactions that cycles between MTR-bound methylcob(III)alamin and MTR bound-cob(I)alamin forms, and occasional oxidative escape of the cob(I)alamin intermediate during the catalytic cycle leads to the inactive cob(II)alamin species. The processing of cobalamin in the cytosol occurs in a multiprotein complex composed of at least MMACHC, MMADHC, MTRR and MTR which may contribute to shuttle safely and efficiently cobalamin towards MTR in order to produce methionine. Also necessary for the utilization of methyl groups from the folate cycle, thereby affecting transgenerational epigenetic inheritance. Also acts as a molecular chaperone for methionine synthase by stabilizing apoMTR and incorporating methylcob(III)alamin into apoMTR to form the	MTR (2)	Cytosol (4), nucleus (3)	None

holoenzyme. Also serves as an aquacob(III)alamin reductase by reducing aquacob(III)alamin to cob(II)alamin; this reduction leads to stimulation of the conversion of apoMTR and aquacob(III)alamin to MTR holoenzyme./ 2 [methionine synthase]-methylcob(III)alamin + H^+ + $NADP^+$ + 2 S-adenosyl-L-homocysteine $\rightarrow$ 2 [methionine synthase]-cob(II)alamin + NADPH + 2 S-adenosyl-L-methionine **EC:1.16.1.8;**
2 cob(II)alamin + H^+ + 2 H_2O + $NADP^+$ $\rightarrow$ 2 aquacob(III)alamin + NADPH

74. NADPH:adrenodoxin oxidoreductase, mitochondrial (FDXR; P22570)	Serves as the first electron transfer protein in all the mitochondrial P450 systems including cholesterol side chain cleavage in all steroidogenic tissues, steroid 11-beta hydroxylation in the adrenal cortex, 25-OH-vitamin D3-24 hydroxylation in the kidney, and sterol C-27 hydroxylation in the liver./ H^+ + $NADP^+$ + 2 reduced [adrenodoxin] $\rightarrow$ NADPH + 2 oxidized [adrenodoxin] **EC:1.18.1.6**	CBWD1 (2), TBC1D22A (2), THBS3 (2), TSSC1 (2), TTC39B (2)	Mitochondrion (5), nucleus (3)	None
75. Cytochrome b-245 heavy chain (CYBB; P04839)	Critical component of the membrane-bound oxidase of phagocytes that generates superoxide. It is the terminal component of a respiratory chain that transfers single electrons from cytoplasmic NADPH across the plasma membrane to molecular oxygen on the exterior. Also functions as a voltage-gated proton channel that mediates the H^+ currents of resting phagocytes. It participates in the regulation of cellular pH and is blocked by zinc.	IQGAP1 (2)	Plasma membrane (5), endoplasmic reticulum (5), nucleus (4), extracellular (3), mitochondrion (3), cytosol (3), golgi apparatus (3)	None
76. Dual oxidase 1 (DUOX1; Q9NRD9)	Generates hydrogen peroxide which is required for the activity of thyroid peroxidase/TPO and lactoperoxidase/LPO. Plays a role in thyroid hormones synthesis and lactoperoxidase-mediated antimicrobial defense at the surface of mucosa. May	None	Plasma membrane (5), endoplasmic reticulum (5), extracellular (3), mitochondrion (3), cytosol (3)	None

	have its own peroxidase activity through its N-terminal peroxidase-like domain./ $H^+ + NADH + O_2 \rightarrow H_2O_2 + NAD^+$ **EC:1.6.3.1;** $H^+ + NADPH + O_2 \rightarrow H_2O_2 + NADP^+$ **EC:1.6.3.1**			
77. Dual oxidase 2 (DUOX2; Q9NRD8)	Generates hydrogen peroxide which is required for the activity of thyroid peroxidase/TPO and lactoperoxidase/LPO. Plays a role in thyroid hormones synthesis and lactoperoxidase-mediated antimicrobial defense at the surface of mucosa. May have its own peroxidase activity through its N-terminal peroxidase-like domain./ $H^+ + NADH + O_2 \rightarrow H_2O_2 + NAD^+$ **EC:1.6.3.1;** $H^+ + NADPH + O_2 \rightarrow H_2O_2 + NADP^+$ **EC:1.6.3.1**	None	Plasma membrane (5), endoplasmic reticulum (5), cytosol (5), extracellular (4), mitochondrion (3)	None
78. 2-hydroxyacyl-CoA lyase 2 (ILVBL; A1L0T0)	Endoplasmic reticulum 2-OH acyl-CoA lyase involved in the cleavage (C1 removal) reaction in the fatty acid alpha-oxydation in a thiamine pyrophosphate (TPP)-dependent manner. Involved in the phytosphingosine degradation pathway./ 2-hydroxyoctadecanoyl-CoA $\rightarrow$ formyl-CoA + heptadecanal (2*R*)-hydroxyhexadecanoyl-CoA $\rightarrow$ formyl-CoA + pentadecanal	GBA (2)	Endoplasmic reticulum (5), plasma membrane (3)	None
79. Alkyldihydroxyacetonephosphate synthase, peroxisomal (AGPS; O00116)	Catalyzes the exchange of the acyl chain in acyl-dihydroxyacetonephosphate (acyl-DHAP) for a long chain fatty alcohol, yielding the first ether linked intermediate, i.e. alkyl-dihydroxyacetonephosphate (alkyl-DHAP), in the pathway of ether lipid biosynthesis./ A 1-acylglycerone 3-phosphate + a long chain fatty alcohol $\rightarrow$ 1-O-alkylglycerone 3-phosphate + a long-chain fatty acid + H^+ **EC:2.5.1.26;** 1-hexadecanoylglycerone 3-phosphate + hexadecan-1-ol $\rightarrow$ 1-O-hexadecylglycerone 3-phosphate + H^+ + hexadecanoate;	CAT (2), EMC2 (2), GNPAT (2), GORASP1 (2), PDHA1 (2), PEX5 (2)	Peroxisome (5), nucleus (5), cytosol (5), mitochondrion (4)	None

	1-hexadecanoylglycerone 3-phosphate + a long-chain fatty acid $\rightarrow$ a 1-acylglycerone 3-phosphate + hexadecanoate			
80. Phosphopantothenoylcystei ne decarboxylase (PPCDC; Q96CD2)	Necessary for the biosynthesis of coenzyme A. Catalyzes the decarboxylation of 4-phosphopantothenoylcysteine to form 4'-phosphopantotheine./ H^+ + N-[(R)-4-phosphopantothenoyl]-L-cysteine $\rightarrow$ CO_2 + D-pantetheine 4'-phosphate **EC:4.1.1.36**	ZNF232 (3), FOXR1 (2), TXN2 (2), WDYHV1 (2)	Cytosol (5), extracellular (3)	None
81. Cryptochrome-1 (CRY1; Q16526)	Transcriptional repressor which forms a core component of the circadian clock. The circadian clock, an internal time-keeping system, regulates various physiological processes through the generation of approximately 24 hour circadian rhythms in gene expression, which are translated into rhythms in metabolism and behavior. It is derived from the Latin roots 'circa' (about) and 'diem' (day) and acts as an important regulator of a wide array of physiological functions including metabolism, sleep, body temperature, blood pressure, endocrine, immune, cardiovascular, and renal function. Consists of two major components: the central clock, residing in the suprachiasmatic nucleus (SCN) of the brain, and the peripheral clocks that are present in nearly every tissue and organ system. Both the central and peripheral clocks can be reset by environmental cues, also known as Zeitgebers (German for 'timegivers'). The predominant Zeitgeber for the central clock is light, which is sensed by retina and signals directly to the SCN. The central clock entrains the peripheral clocks through neuronal and hormonal signals, body temperature and feeding-related cues, aligning all clocks with the external light/dark cycle. Circadian	FBXL3 (8), PER2 (7), PER1 (5), ARNTL (4), CSNK1E (4), CUL1 (3), FOXO1 (3), SKP1 (3), USP7 (3), AP2M1 (2), CRY2 (2), CSNK2A1 (2), CSNK2B (2), PLSCR1 (2), PPP2R1B (2), QPRT (2), TEKT4 (2), USP2 (2), XPO1 (2)	Nucleus (5), cytoskeleton (3), mitochondrion (3), cytosol (3)	None

rhythms allow an organism to achieve temporal homeostasis with its environment at the molecular level by regulating gene expression to create a peak of protein expression once every 24 hours to control when a particular physiological process is most active with respect to the solar day. Transcription and translation of core clock components (CLOCK, NPAS2, ARNTL/BMAL1, ARNTL2/BMAL2, PER1, PER2, PER3, CRY1 and CRY2) plays a critical role in rhythm generation, whereas delays imposed by post-translational modifications (PTMs) are important for determining the period (tau) of the rhythms (tau refers to the period of a rhythm and is the length, in time, of one complete cycle). A diurnal rhythm is synchronized with the day/night cycle, while the ultradian and infradian rhythms have a period shorter and longer than 24 hours, respectively. Disruptions in the circadian rhythms contribute to the pathology of cardiovascular diseases, cancer, metabolic syndromes and aging. A transcription/translation feedback loop (TTFL) forms the core of the molecular circadian clock mechanism. Transcription factors, CLOCK or NPAS2 and ARNTL/BMAL1 or ARNTL2/BMAL2, form the positive limb of the feedback loop, act in the form of a heterodimer and activate the transcription of core clock genes and clock-controlled genes (involved in key metabolic processes), harboring E-box elements (5'-CACGTG-3') within their promoters. The core clock genes: PER1/2/3 and CRY1/2 which are transcriptional repressors form the negative limb of the feedback loop and interact with the CLOCK|NPAS2-ARNTL/BMAL1|ARNTL2/BMAL2 heterodimer

inhibiting its activity and thereby negatively regulating their own expression. This heterodimer also activates nuclear receptors NR1D1/2 and RORA/B/G, which form a second feedback loop and which activate and repress ARNTL/BMAL1 transcription, respectively. CRY1 and CRY2 have redundant functions but also differential and selective contributions at least in defining the pace of the SCN circadian clock and its circadian transcriptional outputs. More potent transcriptional repressor in cerebellum and liver than CRY2, though more effective in lengthening the period of the SCN oscillator. On its side, CRY2 seems to play a critical role in tuning SCN circadian period by opposing the action of CRY1. With CRY2, is dispensable for circadian rhythm generation but necessary for the development of intercellular networks for rhythm synchrony. Capable of translocating circadian clock core proteins such as PER proteins to the nucleus. Interacts with CLOCK-ARNTL/BMAL1 independently of PER proteins and is found at CLOCK-ARNTL/BMAL1-bound sites, suggesting that CRY may act as a molecular gatekeeper to maintain CLOCK-ARNTL/BMAL1 in a poised and repressed state until the proper time for transcriptional activation. Represses the CLOCK-ARNTL/BMAL1 induced transcription of BHLHE40/DEC1. Represses the CLOCK-ARNTL/BMAL1 induced transcription of ATF4, MTA1, KLF10 and NAMPT. May repress circadian target genes expression in collaboration with HDAC1 and HDAC2 through histone deacetylation. Mediates the clock-control activation of ATR and modulates ATR-mediated DNA damage checkpoint.

In liver, mediates circadian regulation of cAMP signaling and gluconeogenesis by binding to membrane-coupled G proteins and blocking glucagon-mediated increases in intracellular cAMP concentrations and CREB1 phosphorylation. Inhibits hepatic gluconeogenesis by decreasing nuclear FOXO1 levels that downregulates gluconeogenic gene expression. Besides its role in the maintenance of the circadian clock, is also involved in the regulation of other processes. Represses glucocorticoid receptor NR3C1/GR-induced transcriptional activity by binding to glucocorticoid response elements (GREs). Plays a key role in glucose and lipid metabolism modulation, in part, through the transcriptional regulation of genes involved in these pathways, such as LEP or ACSL4. Represses PPARD and its target genes in the skeletal muscle and limits exercise capacity. Plays an essential role in the generation of circadian rhythms in the retina. Represses the transcriptional activity of NR1I2.

82. Cryptochrome-2 (CRY2; Q49AN0)	Transcriptional repressor which forms a core component of the circadian clock. The circadian clock, an internal time-keeping system, regulates various physiological processes through the generation of approximately 24 hour circadian rhythms in gene expression, which are translated into rhythms in metabolism and behavior. It is derived from the Latin roots 'circa' (about) and 'diem' (day) and acts as an important regulator of a wide array of physiological functions including metabolism, sleep, body temperature, blood pressure, endocrine, immune, cardiovascular, and renal function. Consists of two major components: the central clock, residing in the	FBXL3 (10), PER2 (4), PER1 (3), AP2M1 (2), ARNTL (2), CLOCK (2), CRY1 (2), CSNK1E (2), CSNK2B (2), CUL1 (2), DDB1 (2), MTUS2 (2), PDE9A (2), PPP5C (2), QPRT (2), SKP1 (2), TEKT4 (2), XPO1 (2), XRN2 (2)	Extracellular (5), nucleus (5), cytosol (5) None

suprachiasmatic nucleus (SCN) of the brain, and the peripheral clocks that are present in nearly every tissue and organ system. Both the central and peripheral clocks can be reset by environmental cues, also known as Zeitgebers (German for 'timegivers'). The predominant Zeitgeber for the central clock is light, which is sensed by retina and signals directly to the SCN. The central clock entrains the peripheral clocks through neuronal and hormonal signals, body temperature and feeding-related cues, aligning all clocks with the external light/dark cycle. Circadian rhythms allow an organism to achieve temporal homeostasis with its environment at the molecular level by regulating gene expression to create a peak of protein expression once every 24 hours to control when a particular physiological process is most active with respect to the solar day. Transcription and translation of core clock components (CLOCK, NPAS2, ARNTL/BMAL1, ARNTL2/BMAL2, PER1, PER2, PER3, CRY1 and CRY2) plays a critical role in rhythm generation, whereas delays imposed by post-translational modifications (PTMs) are important for determining the period (tau) of the rhythms (tau refers to the period of a rhythm and is the length, in time, of one complete cycle). A diurnal rhythm is synchronized with the day/night cycle, while the ultradian and infradian rhythms have a period shorter and longer than 24 hours, respectively. Disruptions in the circadian rhythms contribute to the pathology of cardiovascular diseases, cancer, metabolic syndromes and aging. A transcription/translation feedback loop (TTFL) forms the core of the molecular circadian clock mechanism. Transcription factors, CLOCK or

NPAS2 and ARNTL/BMAL1 or ARNTL2/BMAL2, form the positive limb of the feedback loop, act in the form of a heterodimer and activate the transcription of core clock genes and clock-controlled genes (involved in key metabolic processes), harboring E-box elements (5'-CACGTG-3') within their promoters. The core clock genes: PER1/2/3 and CRY1/2 which are transcriptional repressors form the negative limb of the feedback loop and interact with the CLOCK|NPAS2-ARNTL/BMAL1|ARNTL2/BMAL2 heterodimer inhibiting its activity and thereby negatively regulating their own expression. This heterodimer also activates nuclear receptors NR1D1/2 and RORA/B/G, which form a second feedback loop and which activate and repress ARNTL/BMAL1 transcription, respectively. CRY1 and CRY2 have redundant functions but also differential and selective contributions at least in defining the pace of the SCN circadian clock and its circadian transcriptional outputs. Less potent transcriptional repressor in cerebellum and liver than CRY1, though less effective in lengthening the period of the SCN oscillator. Seems to play a critical role in tuning SCN circadian period by opposing the action of CRY1. With CRY1, dispensable for circadian rhythm generation but necessary for the development of intercellular networks for rhythm synchrony. May mediate circadian regulation of cAMP signaling and gluconeogenesis by blocking glucagon-mediated increases in intracellular cAMP concentrations and in CREB1 phosphorylation. Besides its role in the maintenance of the circadian clock, is also involved in

the regulation of other processes. Plays a key role in glucose and lipid metabolism modulation, in part, through the transcriptional regulation of genes involved in these pathways, such as LEP or ACSL4. Represses glucocorticoid receptor NR3C1/GR-induced transcriptional activity by binding to glucocorticoid response elements (GREs). Represses the CLOCK-ARNTL/BMAL1 induced transcription of BHLHE40/DEC1. Represses the CLOCK-ARNTL/BMAL1 induced transcription of NAMPT. Represses PPARD and its target genes in the skeletal muscle and limits exercise capacity. Represses the transcriptional activity of NR1I2.

| 83. Apoptosis-inducing factor 1, mitochondrial (AIFM1; O95831) | Functions both as NADH oxidoreductase and as regulator of apoptosis. In response to apoptotic stimuli, it is released from the mitochondrion intermembrane space into the cytosol and to the nucleus, where it functions as a proapoptotic factor in a caspase-independent pathway. The soluble form (AIFsol) found in the nucleus induces 'parthanatos' i.e. caspase-independent fragmentation of chromosomal DNA (By similarity). Binds to DNA in a sequence-independent manner. Interacts with EIF3G, and thereby inhibits the EIF3 machinery and protein synthesis, and activates caspase-7 to amplify apoptosis. Plays a critical role in caspase-independent, pyknotic cell death in hydrogen peroxide-exposed cells. In contrast, participates in normal mitochondrial metabolism. Plays an important role in the regulation of respiratory chain biogenesis by interacting with CHCHD4 and controlling CHCHD4 mitochondrial import./ $A + H^+ + NADH \rightarrow AH_2 + NAD^+$ | AK2 (3), ARAF (3), BAG6 (3), CHCHD4 (3), EIF3G (3), HSPA1A (3), MLF2 (3), NDUFS5 (3), PGAM5 (3), RHOT2 (3), TOMM40 (3), TOR1AIP1 (3), APOOL (2), ATP5C1 (2), CFTR (2), CLPB (2), CLPP (2), CLPX (2), COX2 (2), COX4I1 (2), COX6B1 (2), CPOX (2), DLD (2), DNAJC11 (2), ENDOG (2), ERAL1 (2), GFER (2), GRPEL1 (2), HAX1 (2), HSD17B10 (2), HTRA2 (2), IARS2 (2), ILK (2), IMMT (2), LACTB (2), MLH1 (2), MTHFD1L (2), NDUFA13 (2), NDUFA8 (2), OMA1 (2), PARK2 (2), PHB (2), POLDIP2 (2), SAMM50 (2), SCO2 (2), SEC16A (2), SLC25A12 (2), SLC25A5 (2), TIMM13 (2), | Mitochondrion (5), nucleus (5), cytosol (4) | None |

		TIMM8A (2), TIMM8B (2), TSC22D4 (2), TTC19 (2), TUFM (2), TXN (2), XIAP (2)		
84. Ferroptosis suppressor protein 1 (AIFM2; Q9BRQ8)	A NAD(P)H-dependent oxidoreductase involved in cellular oxidative stress response. At the plasma membrane, catalyzes reduction of coenzyme Q/ubiquinone-10 to ubiquinol-10, a lipophilic radical-trapping antioxidant that prevents lipid oxidative damage and consequently ferroptosis. Cooperates with GPX4 to suppress phospholipid peroxidation and ferroptosis. This anti-ferroptotic function is independent of cellular glutathione levels. May play a role in mitochondrial stress signaling. Upon oxidative stress, associates with the lipid peroxidation end product 4-hydroxy-2-nonenal (HNE) forming a lipid adduct devoid of oxidoreductase activity, which then translocates from mitochondria into the nucleus triggering DNA damage and cell death. Capable of DNA binding in a non-sequence specific way./ H^+ + NADH + ubiquinone-10 $\rightarrow$ NAD^+ + ubiquinol-10	None.	Plasma membrane (5), mitochondrion (5), cytosol (5), extracellular (4), nucleus (4).	None
85. Iodotyrosine deiodinase 1 (IYD; Q6PHW0)	Catalyzes the oxidative NADPH-dependent deiodination of monoiodotyrosine (L-MIT) or diiodotyrosine (L-DIT). Acts during the hydrolysis of thyroglobulin to liberate iodide, which can then reenter the hormone-producing pathways. Acts more efficiently on monoiodotyrosine than on diiodotyrosine./ 2 iodide + L-tyrosine + 2 $NADP^+$ $\rightarrow$ 3,5-diiodo-L-tyrosine + H+ + 2 NADPH **EC:1.21.1.1**	DDRGK1 (2), TRIM69 (2)	Plasma membrane (5), nucleus (4)	None
86. [F-actin]-monooxygenase MICAL1 (MICAL1; Q8TDZ2)	Monooxygenase that promotes depolymerization of F-actin by mediating oxidation of specific methionine residues on actin to form methionine-sulfoxide, resulting in actin filament disassembly and preventing	EHD1 (2), NEDD9 (2), RAB1A (2), SLC25A41 (2), TAS2R7 (2), VIM (2)	Extracellular (5), cytoskeleton (5), cytosol (5), plasma	None

	repolymerization. In the absence of actin, it also functions as a NADPH oxidase producing H_2O_2. Acts as a cytoskeletal regulator that connects NEDD9 to intermediate filaments. Also acts as a negative regulator of apoptosis via its interaction with STK38 and STK38L; acts by antagonizing STK38 and STK38L activation by MST1/STK4. Involved in regulation of lamina-specific connectivity in the nervous system such as the development of lamina-restricted hippocampal connections. Through redox regulation of the actin cytoskeleton controls the intracellular distribution of secretory vesicles containing L1/neurofascin/NgCAM family proteins in neurons, thereby regulating their cell surface levels. May act as Rab effector protein and play a role in vesicle trafficking./ H^+ + L-methionyl-[F-actin] + NADPH + O_2 → H_2O + L-methionyl-(R)-S-oxide-[F-actin] + NADP$^+$ **EC:1.14.13.225**		membrane (4), nucleus (3)
87. [F-actin]-monooxygenase MICAL2 (MICAL2; O94851)	Nuclear monooxygenase that promotes depolymerization of F-actin by mediating oxidation of specific methionine residues on actin to form methionine-sulfoxide, resulting in actin filament disassembly and preventing repolymerization. In the absence of actin, it also functions as a NADPH oxidase producing H_2O_2. Acts as a key regulator of the SRF signaling pathway elicited by nerve growth factor and serum: mediates oxidation and subsequent depolymerization of nuclear actin, leading to increase MKL1/MRTF-A presence in the nucleus and promote SRF:MKL1/MRTF-A-dependent gene transcription. Does not activate SRF:MKL1/MRTF-A through RhoA./ H^+ + L-methionyl-[F-actin] + NADPH + O_2 → H_2O + L-methionyl-(R)-S-oxide-[F-actin] + NADP$^+$	None	Nucleus (5), cytoskeleton (4), cytosol (3) None

EC:1.14.13.225

88. [F-actin]-monooxygenase MICAL3 (MICAL3; Q7RTP6)	Monooxygenase that promotes depolymerization of F-actin by mediating oxidation of specific methionine residues on actin to form methionine-sulfoxide, resulting in actin filament disassembly and preventing repolymerization. In the absence of actin, it also functions as a NADPH oxidase producing H_2O_2. Seems to act as Rab effector protein and plays a role in vesicle trafficking. Involved in exocytic vesicles tethering and fusion: the monooxygenase activity is required for this process and implicates RAB8A associated with exocytotic vesicles. Required for cytokinesis. Contributes to stabilization and/or maturation of the intercellular bridge independently of its monooxygenase activity. Promotes recruitment of Rab8 and ERC1 to the intercellular bridge, and together these proteins are proposed to function in timely abscission./ H^+ + L-methionyl-[F-actin] + NADPH + $O_2 \rightarrow H_2O$ + L-methionyl-(R)-S-oxide-[F-actin] + $NADP^+$ **EC:1.14.13.225**	KIF23 (2), NINL (2)	Nucleus (5), cytosol (5), plasma membrane (4), extracellular (4), cytoskeleton (4)	None
89. FAD-dependent oxidoreductase domain-containing protein 1 (FOXRED1; Q96CU9)	Required for the assembly of the mitochondrial membrane respiratory chain NADH dehydrogenase (Complex I). Involved in mid-late stages of complex I assembly.	CLPP (2), HSPD1 (2), MTIF2 (2), NDUFS5 (2)	Mitochondrion (5)	None
90 Solute carrier family 52, riboflavin transporter, member 1 (SLC52A1; Q9NWF4)	Plasma membrane transporter mediating the uptake by cells of the water soluble vitamin B2/riboflavin that plays a key role in biochemical oxidation-reduction reactions of the carbohydrate, lipid, and amino acid metabolism. Humans are unable to synthesize vitamin B2/riboflavin and must obtain it via intestinal absorption.	None	Plasma membrane (5)	None
91. Solute carrier family 52, riboflavin transporter,	Plasma membrane transporter mediating the uptake by cells of the water soluble vitamin B2/riboflavin that	ADRB2 (2)	Plasma membrane (5)	None

member 2 (SLC52A2; Q9HAB3).	plays a key role in biochemical oxidation-reduction reactions of the carbohydrate, lipid, and amino acid metabolism. Humans are unable to synthesize vitamin B2/riboflavin and must obtain it via intestinal absorption. May also act as a receptor for 4-hydroxybutyrate.			
92. Solute carrier family 52, riboflavin transporter, member 3 (SLC52A3; Q9NQ40).	Plasma membrane transporter mediating the uptake by cells of the water soluble vitamin B2/riboflavin that plays a key role in biochemical oxidation-reduction reactions of the carbohydrate, lipid, and amino acid metabolism. Humans are unable to synthesize vitamin B2/riboflavin and must obtain it via intestinal absorption	None	Plasma membrane (5), nucleus (5)	None
93. Flavin reductase (NADPH) (BLVRB; P30043).	Broad specificity oxidoreductase that catalyzes the NADPH-dependent reduction of a variety of flavins, such as riboflavin, FAD or FMN, biliverdins, methemoglobin and PQQ (pyrroloquinoline quinone). Contributes to heme catabolism and metabolizes linear tetrapyrroles. Can also reduce the complexed Fe^{3+} iron to Fe^{2+} in the presence of FMN and NADPH. In the liver, converts biliverdin to bilirubin./ $NADP^+$ + reduced riboflavin $\rightarrow$ 2 H^+ + NADPH + riboflavin **EC:1.5.1.30;** bilirubin IXα + NAD^+ $\rightarrow$ biliverdin IXα + H^+ + NADH **EC:1.3.1.24;** bilirubin IXα + $NADP^+$ $\rightarrow$ biliverdin IXα + H^+ + NADPH **EC:1.3.1.24**	ARL6IP6 (2), FCGR2A (2), FSD1 (2), NENF (2)	Plasma membrane (5), nucleus (5), cytosol (5), extracellular (4)	5OOG; 5OOH
94. Riboflavin kinase (RFK; Q969G6)	Catalyzes the phosphorylation of riboflavin (vitamin B2) to form flavin-mononucleotide (FMN), hence rate-limiting enzyme in the synthesis of FAD. Essential for TNF-induced reactive oxygen species (ROS) production. Through its interaction with both TNFRSF1A and CYBA, physically and functionally couples TNFRSF1A to NADPH oxidase. TNF-activation of RFK may enhance the incorporation of	TNFRSF1A (2)	Cytosol (5), golgi apparatus (3)	None

	FAD in NADPH oxidase, a critical step for the assembly and activation of NADPH oxidase./ ATP + riboflavin $\rightarrow$ ADP + FMN + H$^+$ **EC:2.7.1.26**			
95. FAD synthase (FLAD1; Q8NFF5)	Catalyzes the adenylation of flavin mononucleotide (FMN) to form flavin adenine dinucleotide (FAD) coenzyme./ ATP + FMN + H$^+$ $\rightarrow$ diphosphate + FAD **EC:2.7.7.2**	CDKN1A (2), KLHL14 (2), PLEKHO2 (2), PRTFDC1 (2), REL (2), TIM23 (2)	Cytosol (5), plasma membrane (4), mitochondrion (4)	None

Table S2. FoldX structure-based calculation of the stability effects (ΔΔG) due to phosphorylation and disease-associated missense mutations on human flavoproteins. Flavoproteins were selected from the set analyzed by Vabulas and coworkers in terms of flavin-dependent proteasomal degradation [4]. For each protein, the gene and protein name, UniProt code and the structural model used (PDB code) is indicated. Phosphorylation sites were retrieved from PhosphoSitePlus® (https://www.phosphosite.org/). Those sites indicated with an asterisk (*) denoted sites identified in at least 3 high-throughput proteomic reports (HTPs). Disease-associated variants were retrieved from OMIM (https://www.omim.org/), UniProt (https://www.uniprot.org/) or ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) databases.

Protein (PDB code)	Type	Variant/site	ΔΔG (kcal·mol^{-1})	ASA (%)
CYBR3 (1UMK)	Phosphorylation	T31	0.37	100
	Phosphorylation	S38	0.71	30.2
	Phosphorylation	S66	1.16	32.6
	Phosphorylation	S82(*)	3.64	0
	Phosphorylation	T95	6.57	4.1
	Phosphorylation	S146	0.31	34.1
	Phosphorylation	T171	-0.10	67.9
	Phosphorylation	S174(*)	1.90	1
	Phosphorylation	T182	0.07	46.3
	Phosphorylation	T185	5.98	8
	Phosphorylation	Y43	1.31	15
	Phosphorylation	Y80	-0.84	28
	Phosphorylation	Y248	0.20	26
	Disease-associated	S128P	-1.71	16.9
	Disease-associated	R58Q	2.48	9.4
	Disease-associated	L149P	7.90	0.2
	Disease-associated	V106M	2.59	0.2
	Disease-associated	C204R	10.49	0
	Disease-associated	L73P	7.56	4.3
ACAD8 (1RXO)	Type	Variant/site	ΔΔG (kcal·mol^{-1})	ASA (%)
	Phosphorylation	S216	44.32	5
	Phosphorylation	S261(*)	-2.76	8
	Disease-associated	M128I	3.52	4.6
	Disease-associated	D134Y	13.90	3.4
	Disease-associated	G137R	60.41	14.9
	Disease-associated	M152T	29.46	0.1
	Disease-associated	V203I	5.90	0.1
	Disease-associated	R302Q	7.16	19.6

	Type	Variant/site	$\Delta\Delta G$ (kcal·mol^{-1})	ASA (%)
	Disease-associated	A320T	15.02	0
	Disease-associated	Q385R	35.20	1.5
ETFE (1EFV)	**Type**	**Variant/site**	**$\Delta\Delta G$ (kcal·mol^{-1})**	**ASA (%)**
	Phosphorylation	S65	6.79	0.3
	Phosphorylation	T172	4.60	20.7
	Phosphorylation	T182(*)	10.04	0.6
	Phosphorylation	T194(*)	-0.32	69.7
	Phosphorylation	Y192(*)	-0.19	95
	Disease-associated	R164Q	2.78	5.8
	Disease-associated	D128N	0.46	14.1
ACADM (1EGD)	**Type**	**Variant/site**	**$\Delta\Delta G$ (kcal·mol^{-1})**	**ASA (%)**
	Phosphorylation	T70	1.45	51.3
	Phosphorylation	S207	17.76	14.9
	Phosphorylation	T228	16.84	11.7
	Phosphorylation	T300	45.11	0
	Phosphorylation	S320	8.09	14.9
	Phosphorylation	T351(*)	15.16	5.3
	Phosphorylation	Y73	23.59	10.4
	Phosphorylation	Y183	11.14	2.2
	Phosphorylation	Y352	16.96	14.2
	Phosphorylation	Y353	8.15	17.8
	Phosphorylation	Y400(*)	18.08	28.6
	Disease-associated	K329E	4.81	31.9
	Disease-associated	R53C	6.94	53.5
	Disease-associated	G267R	9.80	12.7
	Disease-associated	I375R	26.74	0
	Disease-associated	C244R	26.81	1
	Disease-associated	M149I	5.01	0
ACADSB (2JIF)	**Type**	**Variant/site**	**$\Delta\Delta G$ (kcal·mol^{-1})**	**ASA (%)**
	Phosphorylation	S81	0.07	0
	Phosphorylation	T82	-0.01	55.7
	Phosphorylation	S183	0.02	31.6
	Phosphorylation	S185	6.05	24.2
	Phosphorylation	S289	33.85	1.1
	Phosphorylation	Y198(*)	3.55	57.2
	Phosphorylation	Y199(*)	8.68	7.7
	Phosphorylation	Y372	9.15	9.7
	Phosphorylation	Y373	57.70	9.3
	Phosphorylation	Y413(*)	35.85	18.7
	Disease-associated	L255F	17.20	2.4
	Disease-associated	I316V	5.41	0
IVD (1IVH)	**Type**	**Variant/site**	**$\Delta\Delta G$ (kcal·mol^{-1})**	**ASA (%)**
	Phosphorylation	S158	4.62	49.9
	Phosphorylation	T214(*)	27.58	3.7
	Phosphorylation	S222(*)	-2.69	37.9
	Phosphorylation	T334	22.12	1.7

	Type	Variant/site	$\Delta\Delta G$ (kcal·mol⁻¹)	ASA (%)
	Phosphorylation	Y161	2.93	32.3
	Phosphorylation	Y277	-3.98	71.9
	Phosphorylation	Y333	15.75	8.3
	Phosphorylation	Y344	51.97	0
	Disease-associated	L13P	16.42	0.3
	Disease-associated	G170V	40.29	1.6
	Disease-associated	R21C	10.68	7.9
	Disease-associated	A282V	-2.69	18.8
	Disease-associated	G123R	33.37	51.6
	Disease-associated	V314A	8.57	0.4
AIFM1 (5KVI)	**Type**	**Variant/site**	**$\Delta\Delta G$ (kcal·mol⁻¹)**	**ASA (%)**
	Phosphorylation	S176	5.84	5.1
	Phosphorylation	S182	3.47	5.4
	Phosphorylation	T188	0.82	29.8
	Phosphorylation	T190	0.22	43.8
	Phosphorylation	T263	0.53	63.3
	Phosphorylation	S266(*)	0.98	28.1
	Phosphorylation	S268(*)	-0.70	82.9
	Phosphorylation	S279(*)	0.91	63
	Phosphorylation	S292(*)	-0.02	58.9
	Phosphorylation	T328	6.40	13.6
	Phosphorylation	S371(*)	0.31	58.7
	Phosphorylation	S375(*)	-0.25	55.8
	Phosphorylation	S376	0.43	81.2
	Phosphorylation	S416	0.21	100
	Phosphorylation	S519(*)	1.26	23.3
	Phosphorylation	T521(*)	14.33	0
	Phosphorylation	S524(*)	4.29	15.2
	Phosphorylation	T526	6.27	17
	Phosphorylation	S530	3.13	5.6
	Phosphorylation	S532	2.94	1.5
	Phosphorylation	T542	0.79	45.5
	Phosphorylation	Y347	0.48	45.3
	Phosphorylation	Y443(*)	-1.40	60.9
	Disease-associated	G308E	5.04	4.3
	Disease-associated	G360R	3.44	53.1
	Disease-associated	T260A	1.38	12.2
	Disease-associated	E493V	-2.48	0.3
ETFA (2A1U)	**Type**	**Variant/site**	**$\Delta\Delta G$ (kcal·mol⁻¹)**	**ASA (%)**
	Phosphorylation	S32	0.59	82
	Phosphorylation	T37	13.59	1.1
	Phosphorylation	T42	3.32	31.3
	Phosphorylation	T45	1.22	45.4
	Phosphorylation	S52	4.67	2.4
	Phosphorylation	T93(*)	4.00	2.5
	Phosphorylation	S140(*)	1.02	46
	Phosphorylation	T143	2.27	17.2
	Phosphorylation	T171	3.24	16

	Type	Variant/site	ΔΔG (kcal·mol⁻¹)	ASA (%)
	Phosphorylation	S172	1.49	14.2
	Phosphorylation	S179	-0.48	98.7
	Phosphorylation	S190	0.10	87.3
	Phosphorylation	S192	0.15	100
	Phosphorylation	S207	1.04	22.7
	Phosphorylation	T213	1.94	31.7
	Phosphorylation	S227	0.27	45.5
	Phosphorylation	S248	0.36	14.8
	Phosphorylation	Y84(*)	8.56	0.2
	Phosphorylation	Y149	1.90	14
	Phosphorylation	Y235(*)	-0.08	27.5
	Disease-associated	V157G	4.08	8.2
	Disease-associated	T266M	-0.88	11.7
	Disease-associated	G116R	11.94	0
	Disease-associated	T171I	-2.31	16
ACADS (2VIG)	**Type**	**Variant/site**	**ΔΔG (kcal·mol⁻¹)**	**ASA (%)**
	Phosphorylation	S121	53.84	0.1
	Phosphorylation	T219	9.30	8.6
	Phosphorylation	T305	13.33	2
	Disease-associated	R46W	7.30	31.1
	Disease-associated	R107C	8.51	14
	Disease-associated	P55L	7.72	55
PPOX (3NKS)	**Type**	**Variant/site**	**ΔΔG (kcal·mol⁻¹)**	**ASA (%)**
	Phosphorylation	T242	-0.25	72.8
	Phosphorylation	T473	0.97	83.8
	Disease-associated	G232R	13.89	0
	Disease-associated	R59W	2.02	9.1
	Disease-associated	R168C	5.32	3.8
	Disease-associated	H20P	4.79	17.9
	Disease-associated	D349A	2.01	8.8
DLD (1ZMD)	**Type**	**Variant/site**	**ΔΔG (kcal·mol⁻¹)**	**ASA (%)**
	Phosphorylation	T135(*)	5.21	7.1
	Phosphorylation	S230	0.70	1.6
	Phosphorylation	T276	7.47	13.7
	Phosphorylation	S297(*)	-1.69	94
	Phosphorylation	T435	2.28	27.8
	Phosphorylation	S502	12.09	11.1
	Phosphorylation	Y153(*)	2.12	34.1
	Disease-associated	K72E	1.89	43.7
	Disease-associated	G229C	10.51	0
	Disease-associated	P488L	7.94	6.2
	Disease-associated	R495G	7.21	14.1
	Disease-associated	E375K	-0.82	15
	Disease-associated	D479V	5.00	39.6
ACADVL (2UXW)	**Type**	**Variant/site**	**ΔΔG (kcal·mol⁻¹)**	**ASA (%)**
	Phosphorylation	S328	0.30	50.4

	Type	Variant/site	$\Delta\Delta G$ (kcal·mol^{-1})	ASA (%)
	Phosphorylation	S485	0.82	74.7
	Phosphorylation	S489(*)	0.80	64.5
	Phosphorylation	S517	1.08	29.5
	Phosphorylation	S522(*)	-1.03	0
	Phosphorylation	S586	5.98	7.6
	Phosphorylation	Y631	5.18	9.6
	Disease-associated	K382Q	1.33	26.6
	Disease-associated	R613W	1.35	48.6
	Disease-associated	G401D	6.11	0
	Disease-associated	A416T	1.79	4.6
KDM1A (2DW4)	**Type**	**Variant/site**	**$\Delta\Delta G$ (kcal·mol^{-1})**	**ASA (%)**
	Phosphorylation	S172(*)	-0.21	79.5
	Phosphorylation	S299	2.7	12.6
	Phosphorylation	T305	3.12	3.0
	Phosphorylation	T335	1.73	37.9
	Phosphorylation	S346	3.23	11.3
	Phosphorylation	T588	3.32	5.4
	Phosphorylation	S611(*)	-0.01	83.5
	Phosphorylation	S683	-0.03	74.5
	Phosphorylation	S687	-0.12	61.5
	Disease-associated	E403K	-0.52	54.4
	Disease-associated	D580G	-0.87	59.9
	Disease-associated	Y785H	1.0	24.7
POR (3QE2)	**Type**	**Variant/site**	**$\Delta\Delta G$ (kcal·mol^{-1})**	**ASA (%)**
	Phosphorylation	S68	2.78	27.6
	Phosphorylation	T251	0.90	82
	Phosphorylation	T525	0.04	48.2
	Phosphorylation	T526	0.60	10
	Phosphorylation	Y245	1.59	5.7
	Phosphorylation	Y259	1.38	14.3
	Phosphorylation	Y373(*)	-1.34	37.5
	Phosphorylation	Y374(*)	21.15	0
	Phosphorylation	Y416	2.92	2.1
	Phosphorylation	Y564	14.35	0
	Phosphorylation	Y573	3.92	7.8
	Phosphorylation	Y575	6.96	1.6
	Phosphorylation	Y604	-0.67	43.2
	Disease-associated	V492E	2.17	3.4
	Disease-associated	A287P	5.40	0
	Disease-associated	C569Y	28.29	0.6
	Disease-associated	R457H	0.36	31
	Disease-associated	G539R	13.17	0.5
DHODH (6FMD)	**Type**	**Variant/site**	**$\Delta\Delta G$ (kcal·mol^{-1})**	**ASA (%)**
	Phosphorylation	T62	7.30	1.8
	Phosphorylation	S63	1.72	0
	Phosphorylation	T121	16.03	0
	Phosphorylation	S214	1.58	51

Type	Variant/site	ΔΔG (kcal·mol⁻¹)	ASA (%)
Phosphorylation	T356	3.40	8.7
Phosphorylation	T359	1.75	29.1
Phosphorylation	Y146	4.11	7.4
Phosphorylation	Y194	2.45	5.7
Phosphorylation	Y355	0.96	26
Disease-associated	G152R	10.37	1.1
Disease-associated	G202D	10.46	0
Disease-associated	R135C	2.00	12
Disease-associated	R346W	5.44	21.6

MTHFR (6FCX)	Type	Variant/site	ΔΔG (kcal·mol⁻¹)	ASA (%)
	Phosphorylation	T329	22.99	0
	Phosphorylation	T330	5.69	15.6
	Phosphorylation	T549	28.23	11.1
	Phosphorylation	Y90	27.12	0
	Disease-associated	A222V	3.31	0
	Disease-associated	R335C	2.04	55.5
	Disease-associated	W339G	7.58	11.1

PNPO (6H00)	Type	Variant/site	ΔΔG (kcal·mol⁻¹)	ASA (%)
	Phosphorylation	S164	-2.83	89.2
	Phosphorylation	S165	0.58	52.7
	Phosphorylation	Y157(*)	0.76	20.2
	Phosphorylation	Y212	-2.46	44.1
	Disease-associated	R229Q	6.27	20.7
	Disease-associated	R229W	17.57	20.7
	Disease-associated	R225H	1.94	91.5

Table S3. Details of the HDX experiment and digestion metrics.

Data Set	WT	S82D	P187S
HDX reaction details	50 mM K-HEPES, pD 7.4, 1 mM TCEP, 25 °C	50 mM K-HEPES, pD 7.4, 1 mM TCEP, 25 °C	50 mM K-HEPES, pD 7.4, 1 mM TCEP, 25 °C
HDX time course (sec)	10, 30, 120, 300, 1200, 3600, 10800	10, 30, 120, 300, 1200, 3600, 10800	10, 30, 120, 300, 1200, 3600, 10800
HDX control samples	fully deuterated WT NQO1	fully deuterated S82D NQO1	fully deuterated P187S NQO1
Average back-exchange	26.20%	26.60%	26.90%
# of Peptides	139	139	136
Sequence coverage	99%	99%	99%
Average peptide length / Redundancy	8.3/4.2	8.3/4.2	8.2/4.1
Replicates (biological or technical)	duplicate of 10, 300, 10800 sec and FD	duplicate of 10, 300, 10800 sec and FD	duplicate of 10, 300, 10800 sec and FD
Repeatability (average SD)	0.74%	0.99%	1.00%

Table S4. HDXMS data for all peptides characterized in this study. In some cases, some *proteoforms* (i.e. WT, S82D or P187S) displayed additional peptides detected in these analyses; a, h and d indicate apo-protein, holo-protein and holo-protein with dicoumarol.

Proteoform	From	To	Sequence	M$_{mono}$	Time(sec)	aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
1	7		MVGRRAL	801.4643	10	23.1	23.8	24.0	24.1	23.7	23.7	24.2	23.7	25.2
1	7		MVGRRAL	801.4643	30	28.8	28.7	28.9	28.8	28.4	29.2	29.0	28.8	30.4
1	7		MVGRRAL	801.4643	120	34.2	34.0	34.9	34.1	33.8	35.4	34.3	34.3	36.0
1	7		MVGRRAL	801.4643	300	36.2	36.4	37.1	36.3	35.8	37.5	36.1	36.3	37.7
1	7		MVGRRAL	801.4643	1200	36.8	36.1	37.9	36.5	36.2	39.0	36.8	36.7	37.2
1	7		MVGRRAL	801.4643	3600	37.5	37.3	38.9	37.5	37.5	37.8	37.3	36.8	38.4
1	7		MVGRRAL	801.4643	10800	41.5	41.7	43.3	41.9	41.0	42.0	41.3	40.5	44.4
6	10		ALIVL	527.3683	10	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.2
6	10		ALIVL	527.3683	30	0.0	0.2	0.0	0.2	0.2	0.2	0.0	0.5	0.4
6	10		ALIVL	527.3683	120	0.2	0.0	0.0	0.1	0.0	0.2	0.1	0.2	0.5
6	10		ALIVL	527.3683	300	0.0	0.0	0.0	0.0	0.1	0.4	0.2	0.0	0.3
6	10		ALIVL	527.3683	1200	0.0	0.0	0.0	0.0	0.0	0.6	0.4	0.0	0.5
6	10		ALIVL	527.3683	3600	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.4	0.4
6	10		ALIVL	527.3683	10800	0.0	0.0	0.0	0.5	0.1	0.2	0.4	0.3	0.2
7	10		LIVL	456.3312	10	0.2	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
7	10		LIVL	456.3312	30	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.4
7	10		LIVL	456.3312	120	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
7	10		LIVL	456.3312	300	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.1
7	10		LIVL	456.3312	1200	0.0	0.0	0.4	0.0	0.0	0.1	0.1	0.0	0.2
7	10		LIVL	456.3312	3600	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
7	10		LIVL	456.3312	10800	0.0	0.0	0.3	0.0	0.0	0.0	0.2	0.2	0.0
8	15		IVLAHSER	923.5189	10	29.6	25.0	23.1	35.2	27.7	24.1	33.5	25.8	23.7
8	15		IVLAHSER	923.5189	30	29.2	24.4	24.2	32.8	28.3	25.0	33.2	25.1	24.5
8	15		IVLAHSER	923.5189	120	32.9	25.6	24.6	32.9	28.5	25.3	33.3	26.7	25.9
8	15		IVLAHSER	923.5189	300	35.6	30.4	28.2	36.9	33.6	29.2	36.3	31.5	30.0
8	15		IVLAHSER	923.5189	1200	36.0	32.6	32.3	38.5	34.3	32.4	37.8	35.0	34.3
8	15		IVLAHSER	923.5189	3600	35.6	34.4	33.8	39.0	33.7	31.5	38.0	34.7	38.7
8	15		IVLAHSER	923.5189	10800	36.8	36.1	36.6	37.6	35.6	34.7	39.4	35.7	38.9
11	19		AHSERTSFN	1047.4734	10	71.3	33.3	31.4	87.1	40.2	32.6	85.7	40.2	32.0
11	19		AHSERTSFN	1047.4734	30	74.4	34.0	31.9	85.4	46.2	33.8	83.9	46.0	31.9
11	19		AHSERTSFN	1047.4734	120	82.6	37.9	33.1	83.3	57.0	33.5	85.1	60.1	33.5
11	19		AHSERTSFN	1047.4734	300	86.8	47.6	35.1	84.2	68.3	36.6	87.0	71.0	34.3
11	19		AHSERTSFN	1047.4734	1200	88.5	62.4	37.3	85.9	76.7	37.2	90.0	78.3	37.7
11	19		AHSERTSFN	1047.4734	3600	85.5	69.8	36.9	86.9	82.7	38.5	86.9	82.5	37.8
11	19		AHSERTSFN	1047.4734	10800	85.1	70.8	37.4	82.9	85.9	40.2	86.6	82.8	34.9
11	20		AHSERTSFNY	1210.5367	10	58.3	27.5	27.9	70.4	30.8	27.4	67.2	30.9	29.3
11	20		AHSERTSFNY	1210.5367	30	60.7	28.8	30.6	68.8	37.8	30.4	68.9	39.1	29.4
11	20		AHSERTSFNY	1210.5367	120	67.5	32.5	31.7	74.3	47.7	30.8	73.5	50.9	30.3
11	20		AHSERTSFNY	1210.5367	300	73.9	39.0	32.2	82.2	55.2	33.4	82.0	59.2	31.9
11	20		AHSERTSFNY	1210.5367	1200	81.7	52.3	35.8	85.6	63.8	34.6	88.1	65.5	32.5
11	20		AHSERTSFNY	1210.5367	3600	84.8	57.3	32.7	86.1	67.8	34.4	86.5	70.7	32.7
11	20		AHSERTSFNY	1210.5367	10800	85.8	59.5	32.5	84.7	71.4	33.0	86.5	74.1	33.4
16	20		TSFNY	630.2649	10	49.3	0.0	0.0	51.6	5.9	0.0	50.3	9.1	0.0
16	20		TSFNY	630.2649	30	48.0	0.6	1.3	51.0	15.6	0.0	51.1	18.9	0.0
16	20		TSFNY	630.2649	120	50.3	4.9	1.0	64.5	34.4	0.7	62.5	39.7	0.0
16	20		TSFNY	630.2649	300	58.7	18.7	1.4	77.8	43.9	1.5	73.8	49.1	0.0
16	20		TSFNY	630.2649	1200	71.5	40.4	2.6	84.2	50.4	1.4	86.5	50.7	0.9
16	20		TSFNY	630.2649	3600	81.7	48.7	1.9	85.9	50.4	3.4	84.6	51.5	0.0
16	20		TSFNY	630.2649	10800	83.5	51.1	0.8	82.9	51.3	2.0	84.4	55.6	0.0
20	23		YAMK	511.2464	10	0.2	0.4	0.6	0.4	0.0	0.3	0.5	0.4	1.1
20	23		YAMK	511.2464	30	0.8	0.2	0.5	0.4	0.1	0.5	0.3	0.5	0.8
20	23		YAMK	511.2464	120	0.5	0.4	0.5	0.8	0.2	0.5	1.0	0.3	0.8
20	23		YAMK	511.2464	300	0.3	0.4	0.7	2.9	0.0	0.7	2.6	0.5	0.7
20	23		YAMK	511.2464	1200	0.7	0.4	0.5	10.7	0.0	0.3	8.8	0.5	0.5
20	23		YAMK	511.2464	3600	1.6	0.5	0.4	17.5	0.5	0.5	19.4	0.6	0.0
20	23		YAMK	511.2464	10800	4.9	0.5	0.9	23.7	0.4	0.3	42.5	1.2	0.7
24	32		EAAAAALKK	871.5127	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24	32		EAAAAALKK	871.5127	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24	32		EAAAAALKK	871.5127	120	0.6	1.2	0.0	1.0	0.5	0.0	2.8	1.5	0.0
24	32		EAAAAALKK	871.5127	300	6.3	6.3	6.3	6.1	5.3	5.1	8.2	7.7	6.2
24	32		EAAAAALKK	871.5127	1200	11.8	12.1	11.0	11.1	10.5	10.3	12.2	12.5	10.8
24	32		EAAAAALKK	871.5127	3600	12.5	12.2	10.8	11.4	10.9	11.0	11.3	11.8	11.7
24	32		EAAAAALKK	871.5127	10800	11.9	11.5	10.7	11.0	11.5	11.3	12.2	11.0	12.0
24	33		EAAAAALKKK	999.6077	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24	33		EAAAAALKKK	999.6077	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24	33		EAAAAALKKK	999.6077	120	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.2	0.0
24	33		EAAAAALKKK	999.6077	300	4.6	4.7	4.3	5.3	4.3	3.6	6.7	5.7	4.3
24	33		EAAAAALKKK	999.6077	1200	11.2	11.6	11.1	11.9	10.9	10.7	12.7	12.0	11.0
24	33		EAAAAALKKK	999.6077	3600	17.1	17.3	17.3	17.7	16.9	16.1	16.9	17.6	17.3
24	33		EAAAAALKKK	999.6077	10800	18.5	19.2	18.6	18.6	18.5	18.6	18.8	18.4	18.2
24	39		EAAAAALKKKGWEVVE	1698.9305	10	7.0	7.6	7.1	7.9	7.9	6.9	8.3	7.8	7.5
24	39		EAAAAALKKKGWEVVE	1698.9305	30	12.5	12.6	12.2	12.8	12.3	12.0	13.1	12.5	12.2
24	39		EAAAAALKKKGWEVVE	1698.9305	120	19.2	19.9	20.8	19.8	18.8	18.4	20.8	20.0	20.8
24	39		EAAAAALKKKGWEVVE	1698.9305	300	24.0	24.2	24.0	24.6	23.6	23.6	25.3	24.7	24.3
24	39		EAAAAALKKKGWEVVE	1698.9305	1200	28.7	28.9	28.4	29.3	28.8	28.8	29.4	29.2	29.6
24	39		EAAAAALKKKGWEVVE	1698.9305	3600	32.0	31.9	32.2	32.0	31.7	31.8	31.9	31.8	32.1
24	39		EAAAAALKKKGWEVVE	1698.9305	10800	33.5	33.7	33.6	33.1	33.4	33.8	34.1	33.1	33.1
25	33		AAAAALKKK	870.5651	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
25	33		AAAAALKKK	870.5651	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
25	33		AAAAALKKK	870.5651	120	0.6	0.8	0.7	1.7	0.4	0.4	2.8	2.0	1.6
25	33		AAAAALKKK	870.5651	300	6.3	6.5	5.8	7.7	6.4	5.9	8.5	7.8	6.2
25	33		AAAAALKKK	870.5651	1200	14.6	15.1	13.9	16.3	14.4	13.8	16.7	15.8	14.8
25	33		AAAAALKKK	870.5651	3600	21.5	21.5	22.0	21.8	21.1	20.0	22.2	22.0	21.2
25	33		AAAAALKKK	870.5651	10800	22.9	23.2	23.1	23.2	23.5	23.4	23.5	22.9	22.2

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	corrected %D								
						aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
	25	39	AAAAALKKKGWEVVE	1569.8879	10	7.3	8.0	7.4	8.3	8.4	6.7	8.9	8.9	7.8
	25	39	AAAAALKKKGWEVVE	1569.8879	30	13.0	13.5	13.3	13.6	13.1	12.6	14.3	14.2	13.6
	25	39	AAAAALKKKGWEVVE	1569.8879	120	20.2	21.0	20.3	20.8	20.2	19.9	22.0	21.9	21.1
	25	39	AAAAALKKKGWEVVE	1569.8879	300	24.9	25.9	25.4	26.0	25.4	25.3	27.2	27.0	25.9
	25	39	AAAAALKKKGWEVVE	1569.8879	1200	30.4	30.9	30.7	31.4	31.0	31.3	31.0	31.8	31.1
	25	39	AAAAALKKKGWEVVE	1569.8879	3600	33.5	34.5	34.2	33.4	34.0	34.4	33.9	34.6	34.7
	25	39	AAAAALKKKGWEVVE	1569.8879	10800	35.2	36.0	35.7	35.2	35.9	35.9	35.7	35.3	35.4
	25	41	AAAAALKKKGWEVVESD	1771.9468	10	10.8	11.3	10.9	11.7	11.2	9.9	12.4	11.4	11.4
	25	41	AAAAALKKKGWEVVESD	1771.9468	30	16.8	16.7	16.2	16.9	16.4	15.9	17.7	16.9	16.7
	25	41	AAAAALKKKGWEVVESD	1771.9468	120	23.7	24.0	23.4	24.3	23.3	23.0	25.2	24.6	23.9
	25	41	AAAAALKKKGWEVVESD	1771.9468	300	28.2	28.5	27.7	28.7	28.3	27.5	29.6	29.2	28.1
	25	41	AAAAALKKKGWEVVESD	1771.9468	1200	32.8	32.8	32.5	33.2	32.8	32.9	33.4	33.5	33.0
	25	41	AAAAALKKKGWEVVESD	1771.9468	3600	35.3	35.7	35.4	35.7	35.6	35.4	35.9	35.8	35.6
	25	41	AAAAALKKKGWEVVESD	1771.9468	10800	37.0	37.1	36.9	37.1	37.2	37.1	37.3	36.5	36.7
	26	41	AAAALKKKGWEVVESD	1700.9097	10	11.7	12.4	12.0	12.7	12.3	11.1	13.1	12.4	12.3
	26	41	AAAALKKKGWEVVESD	1700.9097	30	17.5	17.9	17.8	17.9	17.6	17.2	18.4	18.2	17.7
	26	41	AAAALKKKGWEVVESD	1700.9097	120	24.5	25.3	24.9	25.6	25.1	24.5	26.2	26.2	25.3
	26	41	AAAALKKKGWEVVESD	1700.9097	300	30.0	30.5	30.2	30.8	30.5	29.9	31.4	31.5	30.5
	26	41	AAAALKKKGWEVVESD	1700.9097	1200	35.0	35.9	35.5	35.7	36.2	35.8	35.4	36.3	35.9
	26	41	AAAALKKKGWEVVESD	1700.9097	3600	38.0	39.0	38.9	38.1	39.2	38.9	38.4	39.1	38.8
	26	41	AAAALKKKGWEVVESD	1700.9097	10800	39.8	40.5	40.2	40.1	40.8	40.5	39.8	39.7	39.7
	27	39	AAALKKKGWEVVE	1427.8136	10	9.2	10.3	9.8	9.9	10.1	8.8	10.9	10.4	9.9
	27	39	AAALKKKGWEVVE	1427.8136	30	15.8	17.3	17.0	15.7	16.4	15.8	16.6	17.2	16.9
	27	39	AAALKKKGWEVVE	1427.8136	120	23.3	26.3	25.8	24.0	25.1	25.1	25.3	26.6	26.0
	27	39	AAALKKKGWEVVE	1427.8136	300	29.5	31.4	31.5	29.9	30.7	31.6	31.0	31.9	31.9
	27	39	AAALKKKGWEVVE	1427.8136	1200	34.6	37.2	38.1	34.7	37.3	38.4	35.2	37.3	38.3
	27	39	AAALKKKGWEVVE	1427.8136	3600	37.5	40.2	41.6	38.2	40.0	41.4	38.0	40.5	41.5
	27	39	AAALKKKGWEVVE	1427.8136	10800	39.8	42.3	43.3	40.1	41.5	42.8	40.1	41.5	42.9
	27	41	AAALKKKGWEVVESD	1629.8726	10	13.8	15.0	15.6	14.6	14.6	14.6	15.3	14.8	15.4
	27	41	AAALKKKGWEVVESD	1629.8726	30	20.0	21.1	22.7	20.0	20.6	21.0	20.8	21.7	20.5
	27	41	AAALKKKGWEVVESD	1629.8726	120	27.0	29.1	30.8	27.5	28.3	29.3	28.7	29.8	30.5
	27	41	AAALKKKGWEVVESD	1629.8726	300	33.0	34.4	36.4	33.9	33.9	35.4	34.3	35.2	35.6
	27	41	AAALKKKGWEVVESD	1629.8726	1200	38.1	40.0	42.2	38.1	39.9	41.6	38.4	40.5	41.5
	27	41	AAALKKKGWEVVESD	1629.8726	3600	41.0	43.0	45.1	41.1	42.8	44.2	41.1	43.1	43.9
	27	41	AAALKKKGWEVVESD	1629.8726	10800	42.9	44.4	45.9	42.4	44.0	45.4	42.7	43.3	44.6
	27	42	AAALKKKGWEVVESDL	1742.9567	10	14.2	15.2	16.7	15.0	14.7	14.7	15.9	15.1	16.3
	27	42	AAALKKKGWEVVESDL	1742.9567	30	20.2	21.1	24.0	19.9	20.3	21.2	21.0	21.1	22.9
	27	42	AAALKKKGWEVVESDL	1742.9567	120	26.7	28.2	31.2	26.9	27.0	28.7	28.1	28.5	30.5
	27	42	AAALKKKGWEVVESDL	1742.9567	300	31.7	32.9	35.8	31.6	32.3	33.9	32.8	33.6	34.8
	27	42	AAALKKKGWEVVESDL	1742.9567	1200	35.7	37.6	40.9	35.7	37.0	39.2	36.1	38.0	40.2
	27	42	AAALKKKGWEVVESDL	1742.9567	3600	38.0	40.2	43.7	38.1	39.4	41.5	38.7	40.1	41.8
	27	42	AAALKKKGWEVVESDL	1742.9567	10800	39.8	41.7	44.0	39.5	41.1	43.0	40.2	40.7	42.4
	30	39	LKKKGWEVVE	1214.7023	10	10.7	11.9	11.2	12.1	12.0	10.5	13.6	12.6	11.8
	30	39	LKKKGWEVVE	1214.7023	30	19.9	21.6	21.6	20.5	20.3	21.5	21.6	21.6	21.0
	30	39	LKKKGWEVVE	1214.7023	120	30.0	33.2	33.1	31.0	31.8	32.7	32.4	33.0	33.6
	30	39	LKKKGWEVVE	1214.7023	300	37.5	39.5	40.0	39.0	39.2	41.4	39.8	40.0	41.1
	30	39	LKKKGWEVVE	1214.7023	1200	44.7	46.4	48.8	45.8	45.8	49.9	45.8	47.2	48.4
	30	39	LKKKGWEVVE	1214.7023	3600	48.7	50.7	53.1	51.1	51.4	54.3	49.6	51.3	54.2
	30	39	LKKKGWEVVE	1214.7023	10800	51.9	53.2	55.1	53.0	53.5	55.1	52.2	53.4	56.4
	30	42	LKKKGWEVVESDL	1529.8453	10	17.4	19.2	18.9	18.4	18.6	16.5	19.6	19.2	18.0
	30	42	LKKKGWEVVESDL	1529.8453	30	24.5	27.0	27.2	24.3	26.0	24.6	25.6	26.9	25.1
	30	42	LKKKGWEVVESDL	1529.8453	120	32.0	36.1	35.9	32.2	34.6	34.3	33.5	36.1	34.3
	30	42	LKKKGWEVVESDL	1529.8453	300	37.7	42.3	41.9	38.1	41.7	41.4	38.9	42.4	40.3
	30	42	LKKKGWEVVESDL	1529.8453	1200	41.5	47.9	48.2	41.9	47.6	47.5	42.4	47.7	46.9
	30	42	LKKKGWEVVESDL	1529.8453	3600	44.7	51.0	51.4	45.2	50.9	50.7	44.9	50.2	49.7
	30	42	LKKKGWEVVESDL	1529.8453	10800	47.4	52.3	52.5	47.0	51.6	51.5	47.6	50.2	49.7
	34	39	GWEVVE	717.3333	10	14.8	15.3	14.6	15.9	15.2	12.3	17.5	15.9	14.0
	34	39	GWEVVE	717.3333	30	23.5	23.3	22.1	23.9	22.8	21.6	24.5	24.2	22.5
	34	39	GWEVVE	717.3333	120	38.0	39.0	38.1	37.8	35.5	33.5	40.3	40.4	39.0
	34	39	GWEVVE	717.3333	300	44.7	45.6	44.8	43.8	44.0	43.4	46.0	46.5	45.9
	34	39	GWEVVE	717.3333	1200	45.8	45.9	45.7	45.0	45.0	44.8	45.5	46.6	46.9
	34	39	GWEVVE	717.3333	3600	45.2	45.9	45.5	44.2	45.1	44.4	45.7	46.1	46.5
	34	39	GWEVVE	717.3333	10800	44.2	44.5	45.2	43.0	44.0	43.6	45.4	44.9	45.7
	36	41	EVVESD	676.2915	10	18.8	19.2	19.1	19.5	17.4	15.9	18.4	17.3	16.7
	36	41	EVVESD	676.2915	30	22.7	22.9	20.5	23.0	20.9	19.2	22.4	21.6	20.9
	36	41	EVVESD	676.2915	120	35.5	35.5	34.6	35.9	33.4	31.7	37.5	36.5	35.1
	36	41	EVVESD	676.2915	300	44.0	44.1	43.4	43.8	43.6	42.8	43.9	44.0	43.3
	36	41	EVVESD	676.2915	1200	44.7	43.6	44.4	44.1	44.5	44.6	44.4	44.0	43.8
	36	41	EVVESD	676.2915	3600	43.1	43.8	43.6	43.4	43.7	43.2	43.5	42.2	44.9
	36	41	EVVESD	676.2915	10800	43.6	43.5	43.7	43.4	44.0	43.8	43.5	42.4	42.4
	36	42	EVVESDL	789.3756	10	13.0	13.8	13.1	13.2	13.0	11.0	13.7	12.7	12.3
	36	42	EVVESDL	789.3756	30	17.2	18.4	17.5	17.1	17.1	16.4	18.0	17.6	16.8
	36	42	EVVESDL	789.3756	120	28.0	29.3	28.1	28.5	27.7	25.6	29.6	29.6	27.8
	36	42	EVVESDL	789.3756	300	34.2	35.0	34.0	34.1	35.1	33.9	33.7	34.5	33.5
	36	42	EVVESDL	789.3756	1200	35.1	35.5	35.0	34.5	36.1	35.4	33.5	34.6	35.0
	36	42	EVVESDL	789.3756	3600	34.0	35.0	34.7	33.9	35.5	34.2	33.6	33.6	33.2
	36	42	EVVESDL	789.3756	10800	34.3	34.4	34.1	33.8	34.8	34.6	33.0	32.3	32.3
	42	45	LYAM	496.2355	10	0.5	0.9	1.4	0.0	0.0	0.2	0.0	0.0	0.4
	42	45	LYAM	496.2355	30	1.1	0.9	1.5	1.1	0.0	0.6	0.4	0.2	0.8
	42	45	LYAM	496.2355	120	6.8	3.1	6.8	10.6	1.4	3.5	7.9	2.1	5.4
	42	45	LYAM	496.2355	300	24.4	13.9	15.8	36.6	11.1	8.3	27.5	12.5	15.3
	42	45	LYAM	496.2355	1200	49.9	40.9	41.2	82.2	42.8	31.5	60.4	40.8	42.0
	42	45	LYAM	496.2355	3600	65.4	58.3	59.2	95.2	77.3	58.7	82.2	58.1	59.6
	42	45	LYAM	496.2355	10800	85.9	77.9	77.2	95.5	94.1	80.3	94.7	76.7	76.6

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	corrected %D								
						aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
	46	54	NFNPIISRK	1087.6138	10	39.0	27.7	24.9	57.5	49.9	26.9	49.1	28.8	23.5
	46	54	NFNPIISRK	1087.6138	30	45.4	33.2	31.5	65.9	54.1	34.3	53.2	34.1	29.0
	46	54	NFNPIISRK	1087.6138	120	51.4	39.7	34.9	72.7	65.7	44.2	64.5	40.9	33.9
	46	54	NFNPIISRK	1087.6138	300	62.6	47.5	37.0	81.7	73.5	53.2	75.4	48.9	35.2
	46	54	NFNPIISRK	1087.6138	1200	74.1	55.5	44.6	89.2	80.2	63.6	84.3	58.4	43.5
	46	54	NFNPIISRK	1087.6138	3600	78.9	65.9	50.0	89.9	86.2	69.2	89.0	68.1	51.5
	46	54	NFNPIISRK	1087.6138	10800	86.7	76.0	58.0	89.0	89.6	77.4	90.0	75.0	57.6
	46	55	NFNPIISRKD	1202.6408	10	38.0	24.3	24.3	58.2	51.7	26.2	48.8	25.9	21.8
	46	55	NFNPIISRKD	1202.6408	30	45.1	31.0	31.8	64.9	56.0	33.0	52.4	31.8	29.4
	46	55	NFNPIISRKD	1202.6408	120	51.5	39.1	33.9	71.6	67.3	46.5	63.5	41.2	33.8
	46	55	NFNPIISRKD	1202.6408	300	62.3	48.5	36.2	78.8	74.1	56.6	74.0	49.4	34.2
	46	55	NFNPIISRKD	1202.6408	1200	73.9	58.8	44.3	86.2	79.1	66.5	81.2	61.5	43.0
	46	55	NFNPIISRKD	1202.6408	3600	77.0	67.7	53.4	87.2	84.1	71.0	85.7	68.6	53.9
	46	55	NFNPIISRKD	1202.6408	10800	84.4	76.2	61.9	86.8	86.6	78.2	87.4	75.1	61.0
	55	59	DITGK	532.2857	10	85.5	61.8	59.7	91.1	90.5	66.2	89.5	63.5	58.9
	55	59	DITGK	532.2857	30	88.4	66.4	58.8	88.9	89.4	73.7	88.8	69.7	58.2
	55	59	DITGK	532.2857	120	89.4	81.1	60.2	87.4	88.8	88.0	89.5	84.7	60.9
	55	59	DITGK	532.2857	300	91.5	90.5	64.6	89.7	91.9	91.2	91.1	91.4	64.1
	55	59	DITGK	532.2857	1200	92.0	91.3	77.5	89.8	91.7	92.6	92.0	91.5	78.0
	55	59	DITGK	532.2857	3600	89.0	90.8	87.7	90.6	90.5	89.1	90.6	89.1	87.8
	55	59	DITGK	532.2857	10800	90.3	90.8	90.7	88.2	90.9	91.1	90.3	87.9	89.5
	55	62	DITGKLKD	888.4916	10	79.7	53.0	46.2	87.7	87.4	54.8	87.3	58.3	46.5
	55	62	DITGKLKD	888.4916	30	86.0	66.1	54.7	86.4	86.8	69.7	87.4	70.7	54.7
	55	62	DITGKLKD	888.4916	120	86.6	81.4	67.6	85.6	86.3	85.8	87.6	85.1	68.8
	55	62	DITGKLKD	888.4916	300	88.8	87.6	74.6	87.1	88.7	88.3	88.6	89.0	75.0
	55	62	DITGKLKD	888.4916	1200	89.2	88.0	80.8	87.4	88.7	89.3	88.8	89.5	82.1
	55	62	DITGKLKD	888.4916	3600	87.0	87.7	85.8	87.4	87.8	85.4	88.5	86.7	85.6
	55	62	DITGKLKD	888.4916	10800	87.4	86.9	86.4	85.3	87.7	87.3	87.6	85.1	86.5
	55	65	DITGKLKDPAN	1170.6244	10	80.8	56.6	49.6	88.6	89.0	57.9	87.7	61.2	48.9
	55	65	DITGKLKDPAN	1170.6244	30	87.3	69.5	58.6	87.0	88.7	72.9	87.5	73.8	57.8
	55	65	DITGKLKDPAN	1170.6244	120	87.6	84.4	73.0	86.5	88.2	88.0	87.7	86.8	73.2
	55	65	DITGKLKDPAN	1170.6244	300	89.5	89.4	80.1	87.9	89.8	90.3	89.0	89.8	79.4
	55	65	DITGKLKDPAN	1170.6244	1200	89.9	89.8	84.8	87.6	89.9	91.3	89.1	89.9	84.9
	55	65	DITGKLKDPAN	1170.6244	3600	87.6	88.5	87.9	88.0	88.5	87.9	88.3	87.5	87.1
	55	65	DITGKLKDPAN	1170.6244	10800	88.1	88.4	88.8	86.4	88.6	89.0	87.6	86.4	86.6
	55	68	DITGKLKDPANFQY	1608.8148	10	81.4	57.6	49.2	90.1	91.0	57.0	89.4	61.6	48.4
	55	68	DITGKLKDPANFQY	1608.8148	30	88.0	71.1	60.7	88.2	90.4	73.6	88.6	75.0	59.7
	55	68	DITGKLKDPANFQY	1608.8148	120	88.5	86.9	74.0	87.4	90.0	89.9	89.1	88.9	74.3
	55	68	DITGKLKDPANFQY	1608.8148	300	90.7	92.1	80.4	89.5	92.3	92.7	91.1	92.6	79.9
	55	68	DITGKLKDPANFQY	1608.8148	1200	91.1	92.5	86.2	89.4	92.3	93.6	91.7	92.7	86.8
	55	68	DITGKLKDPANFQY	1608.8148	3600	89.5	91.7	90.7	89.9	91.2	91.0	90.9	90.9	91.3
	55	68	DITGKLKDPANFQY	1608.8148	10800	90.0	91.5	91.7	88.7	91.2	91.8	90.5	89.8	91.0
	60	65	LKDPAN	656.3493	10	80.2	50.6	39.9	93.9	91.4	51.8	91.7	56.8	42.3
	60	65	LKDPAN	656.3493	30	91.2	69.1	52.8	91.6	90.5	71.3	91.8	76.5	56.4
	60	65	LKDPAN	656.3493	120	90.7	87.2	80.5	91.1	90.0	91.4	91.5	91.3	84.5
	60	65	LKDPAN	656.3493	300	93.6	93.2	92.1	93.0	93.5	94.2	93.8	94.7	94.1
	60	65	LKDPAN	656.3493	1200	94.6	92.5	93.2	93.7	92.4	94.3	95.0	94.4	93.7
	60	65	LKDPAN	656.3493	3600	90.2	91.1	91.6	92.4	91.2	89.7	94.1	91.2	91.1
	60	65	LKDPAN	656.3493	10800	91.7	89.4	88.8	91.0	89.8	90.5	93.9	88.9	90.0
	60	68	LKDPANFQY	1094.5397	10	81.8	58.8	42.9	90.6	95.7	53.2	89.1	62.6	41.5
	60	68	LKDPANFQY	1094.5397	30	89.3	77.8	61.4	87.8	95.3	75.6	87.3	79.8	57.8
	60	68	LKDPANFQY	1094.5397	120	88.9	93.6	80.6	87.4	94.9	94.2	87.6	93.5	78.9
	60	68	LKDPANFQY	1094.5397	300	91.9	98.1	87.3	90.5	96.9	98.2	90.7	96.2	84.2
	60	68	LKDPANFQY	1094.5397	1200	92.7	98.9	92.7	89.9	97.5	98.7	91.4	97.1	90.5
	60	68	LKDPANFQY	1094.5397	3600	90.6	97.8	97.0	91.0	95.8	95.8	90.2	95.0	94.5
	60	68	LKDPANFQY	1094.5397	10800	91.4	96.9	97.6	89.9	95.2	95.7	90.2	92.3	93.8
	60	71	LKDPANFQYPAE	1391.6721	10	75.7	49.1	36.3	90.2	90.9	44.0	87.5	51.0	36.5
	60	71	LKDPANFQYPAE	1391.6721	30	87.0	65.1	49.5	86.7	90.0	63.1	87.2	67.9	51.1
	60	71	LKDPANFQYPAE	1391.6721	120	88.2	86.9	66.8	86.2	90.4	85.7	87.1	87.8	68.4
	60	71	LKDPANFQYPAE	1391.6721	300	91.2	94.5	75.4	88.3	93.4	93.9	90.5	93.5	75.5
	60	71	LKDPANFQYPAE	1391.6721	1200	91.5	94.5	83.2	87.8	92.4	93.8	91.0	93.6	83.3
	60	71	LKDPANFQYPAE	1391.6721	3600	91.0	93.8	89.8	88.7	91.3	91.8	89.5	92.1	90.2
	60	71	LKDPANFQYPAE	1391.6721	10800	91.1	93.6	93.1	88.3	91.1	91.5	89.8	90.8	92.7
	60	74	LKDPANFQYPAESVL	1690.8566	10	69.7	38.3	30.9	88.7	89.8	39.2	84.6	40.8	28.9
	60	74	LKDPANFQYPAESVL	1690.8566	30	83.4	54.7	40.6	86.0	89.5	56.7	85.5	57.2	39.1
	60	74	LKDPANFQYPAESVL	1690.8566	120	87.4	79.3	52.5	85.5	89.5	81.7	86.4	80.7	51.9
	60	74	LKDPANFQYPAESVL	1690.8566	300	90.6	91.5	59.2	88.5	91.8	91.6	89.1	90.1	57.6
	60	74	LKDPANFQYPAESVL	1690.8566	1200	90.9	94.3	69.8	88.3	92.0	93.4	89.9	91.9	69.4
	60	74	LKDPANFQYPAESVL	1690.8566	3600	89.5	93.8	82.3	88.9	90.8	91.5	88.9	90.7	81.8
	60	74	LKDPANFQYPAESVL	1690.8566	10800	90.4	93.6	91.9	89.0	91.0	92.0	89.4	90.0	90.4
	62	68	DPANFQY	853.3606	10	80.8	54.8	40.5	90.3	91.0	48.5	89.1	58.0	40.5
	62	68	DPANFQY	853.3606	30	88.5	70.9	56.6	87.8	90.8	69.9	87.5	73.4	56.4
	62	68	DPANFQY	853.3606	120	88.7	87.4	72.8	87.0	90.6	90.9	87.6	88.2	71.8
	62	68	DPANFQY	853.3606	300	91.5	93.2	79.9	89.9	92.9	94.8	90.8	92.7	78.4
	62	68	DPANFQY	853.3606	1200	92.3	93.6	85.8	89.6	93.2	95.1	91.2	92.7	85.5
	62	68	DPANFQY	853.3606	3600	90.3	92.3	91.6	90.6	92.0	92.9	90.4	90.6	91.7
	62	68	DPANFQY	853.3606	10800	91.0	92.1	93.3	89.5	91.6	92.6	89.7	90.1	92.9
	66	73	FQYPAESV	939.4338	10	70.5	33.6	18.7	90.7	90.9	32.9	87.1	36.7	19.3
	66	73	FQYPAESV	939.4338	30	84.3	49.3	26.7	89.2	90.6	52.2	87.8	53.2	27.6
	66	73	FQYPAESV	939.4338	120	87.4	76.2	36.6	88.4	90.8	80.9	87.9	80.2	38.1
	66	73	FQYPAESV	939.4338	300	89.4	88.8	44.2	90.5	92.5	92.5	89.8	89.6	44.2
	66	73	FQYPAESV	939.4338	1200	90.1	90.8	59.0	90.3	92.7	93.2	90.2	90.8	61.1
	66	73	FQYPAESV	939.4338	3600	88.5	90.6	77.7	90.8	91.7	91.5	89.7	90.1	79.5
	66	73	FQYPAESV	939.4338	10800	89.3	90.5	88.7	90.0	91.6	91.1	89.4	88.7	87.3

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	corrected %D								
						aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
	66	74	FQYPAESVL	1052.5179	10	61.6	29.8	20.1	85.4	86.8	30.6	81.5	33.8	20.0
	66	74	FQYPAESVL	1052.5179	30	76.5	44.2	27.4	85.0	88.2	46.7	84.8	48.0	27.6
	66	74	FQYPAESVL	1052.5179	120	84.5	70.2	35.8	83.5	88.1	73.0	86.4	74.5	38.1
	66	74	FQYPAESVL	1052.5179	300	87.4	85.1	42.4	85.2	89.8	86.8	88.1	85.8	42.7
	66	74	FQYPAESVL	1052.5179	1200	88.8	89.7	54.4	85.6	89.0	90.3	89.0	88.3	56.5
	66	74	FQYPAESVL	1052.5179	3600	86.0	88.8	71.4	86.9	89.4	88.8	88.1	87.8	72.0
	66	74	FQYPAESVL	1052.5179	10800	86.5	88.9	84.6	85.4	88.8	87.9	87.2	84.8	78.2
	69	73	PAESV	501.2435	10	65.8	15.4	3.4	93.5	90.9	18.9	88.0	20.3	3.4
	69	73	PAESV	501.2435	30	86.6	32.3	5.0	91.8	91.1	37.2	91.3	39.0	5.1
	69	73	PAESV	501.2435	120	90.6	71.2	10.5	91.2	91.8	78.1	91.4	77.5	11.6
	69	73	PAESV	501.2435	300	92.6	88.8	19.5	92.5	94.0	93.8	93.5	90.5	21.2
	69	73	PAESV	501.2435	1200	93.2	91.4	41.5	92.7	94.2	95.0	93.7	92.4	45.4
	69	73	PAESV	501.2435	3600	91.8	92.1	70.2	93.2	92.9	92.4	92.9	90.9	74.1
	69	73	PAESV	501.2435	10800	92.3	92.2	90.2	91.3	92.3	93.1	92.7	90.7	90.2
	69	74	PAESVL	614.3275	10	53.9	7.9	0.8	90.9	88.5	12.8	82.5	12.4	0.0
	69	74	PAESVL	614.3275	30	78.4	21.9	1.0	89.6	89.3	27.2	88.4	28.7	0.0
	69	74	PAESVL	614.3275	120	88.6	60.0	5.1	89.4	90.2	65.0	90.1	68.6	3.2
	69	74	PAESVL	614.3275	300	91.3	84.4	6.9	90.6	92.2	87.4	91.8	87.2	9.8
	69	74	PAESVL	614.3275	1200	91.7	90.5	29.0	90.6	92.2	93.2	91.6	90.7	31.8
	69	74	PAESVL	614.3275	3600	90.5	90.5	57.0	91.1	91.5	91.1	91.4	90.1	60.9
	69	74	PAESVL	614.3275	10800	91.2	91.1	83.7	88.8	91.7	91.4	90.4	88.6	83.6
	69	76	PAESVLAY	848.428	10	31.2	7.4	4.7	61.5	59.0	9.4	53.0	11.2	2.4
	69	76	PAESVLAY	848.428	30	49.2	16.3	8.7	60.9	61.3	18.7	58.2	19.2	8.0
	69	76	PAESVLAY	848.428	120	59.3	39.5	10.3	59.7	60.8	45.5	59.2	45.3	11.1
	69	76	PAESVLAY	848.428	300	61.3	56.8	13.1	61.0	62.0	60.7	61.0	59.4	11.8
	69	76	PAESVLAY	848.428	1200	62.4	62.0	23.4	61.0	62.1	65.2	61.6	62.1	25.3
	69	76	PAESVLAY	848.428	3600	60.6	62.3	40.6	60.8	62.0	63.5	60.9	61.5	42.8
	69	76	PAESVLAY	848.428	10800	61.1	62.4	58.0	61.5	62.4	63.1	61.8	61.7	59.4
	74	87	LAYKEGHLSPDIVA	1511.7984	10	32.5	12.8	6.2	56.3	59.0	13.9	51.0	16.0	6.7
	74	87	LAYKEGHLSPDIVA	1511.7984	30	45.7	23.8	9.8	57.8	61.4	26.1	53.3	28.1	11.1
	74	87	LAYKEGHLSPDIVA	1511.7984	120	52.1	47.7	16.9	66.5	69.0	50.5	57.9	50.2	18.1
	74	87	LAYKEGHLSPDIVA	1511.7984	300	60.6	60.0	24.8	78.0	80.0	62.0	67.0	61.3	25.7
	74	87	LAYKEGHLSPDIVA	1511.7984	1200	70.3	66.3	40.0	83.4	87.6	65.4	77.9	67.2	42.5
	74	87	LAYKEGHLSPDIVA	1511.7984	3600	76.1	71.4	51.6	87.7	90.9	67.1	81.5	71.8	52.8
	74	87	LAYKEGHLSPDIVA	1511.7984	10800	82.4	79.0	63.2	90.1	94.4	77.6	88.4	79.7	62.6
	75	90	AYKEGHLSPDIVAEQK	1783.9105	10	20.5	10.3	7.1	34.7	35.2	11.0	31.3	11.8	7.1
	75	90	AYKEGHLSPDIVAEQK	1783.9105	30	28.5	18.2	9.9	36.8	38.5	19.5	32.5	20.0	9.7
	75	90	AYKEGHLSPDIVAEQK	1783.9105	120	32.5	32.3	17.6	44.9	45.0	35.3	36.3	33.4	16.4
	75	90	AYKEGHLSPDIVAEQK	1783.9105	300	38.4	38.6	22.9	54.0	53.5	40.1	44.2	38.6	22.4
	75	90	AYKEGHLSPDIVAEQK	1783.9105	1200	46.9	43.1	32.9	60.0	60.9	41.8	53.3	43.1	31.6
	75	90	AYKEGHLSPDIVAEQK	1783.9105	3600	53.3	46.7	37.3	68.5	65.1	44.3	59.0	47.8	37.1
	75	90	AYKEGHLSPDIVAEQK	1783.9105	10800	59.8	54.1	41.7	76.1	71.3	52.4	68.5	54.2	41.5
	75	91	AYKEGHLSPDIVAEQKK	1912.0054	10	19.2	9.2	7.9	30.8	30.6	11.0	29.3	10.5	7.8
	75	91	AYKEGHLSPDIVAEQKK	1912.0054	30	26.6	16.6	11.7	32.9	34.0	18.8	30.8	18.8	10.5
	75	91	AYKEGHLSPDIVAEQKK	1912.0054	120	30.2	29.5	17.6	40.4	40.0	32.9	34.0	30.4	17.8
	75	91	AYKEGHLSPDIVAEQKK	1912.0054	300	35.6	34.2	22.1	48.0	47.0	36.3	41.3	34.5	20.1
	75	91	AYKEGHLSPDIVAEQKK	1912.0054	1200	43.4	38.1	31.3	53.6	53.5	38.4	49.4	38.6	29.7
	75	91	AYKEGHLSPDIVAEQKK	1912.0054	3600	49.4	41.5	34.7	61.1	57.7	39.7	54.6	43.0	34.2
	75	91	AYKEGHLSPDIVAEQKK	1912.0054	10800	55.4	48.8	38.2	68.6	62.5	46.9	62.9	49.1	37.7
	77	90	KEGHLSPDIVAEQK	1549.81	10	12.7	7.2	9.3	20.4	17.7	8.0	20.2	7.8	9.2
	77	90	KEGHLSPDIVAEQK	1549.81	30	17.9	12.0	13.4	24.2	22.0	14.6	21.9	13.1	13.1
	77	90	KEGHLSPDIVAEQK	1549.81	120	21.4	19.7	18.6	34.6	30.7	23.8	26.2	19.3	18.5
	77	90	KEGHLSPDIVAEQK	1549.81	300	27.5	23.0	20.0	44.2	40.0	25.4	34.6	23.3	19.3
	77	90	KEGHLSPDIVAEQK	1549.81	1200	37.4	27.2	24.1	52.2	48.3	27.4	44.4	28.0	21.2
	77	90	KEGHLSPDIVAEQK	1549.81	3600	45.3	32.3	23.4	62.4	54.3	29.9	51.4	34.2	23.4
	77	90	KEGHLSPDIVAEQK	1549.81	10800	52.5	42.0	27.9	71.2	62.5	39.7	61.7	42.5	28.2
	77	91	KEGHLSPDIVAEQKK	1677.905	10	12.5	6.7	9.1	18.6	15.9	7.7	18.7	7.2	8.5
	77	91	KEGHLSPDIVAEQKK	1677.905	30	16.8	10.9	13.2	22.0	19.2	14.2	20.0	11.4	12.5
	77	91	KEGHLSPDIVAEQKK	1677.905	120	19.7	17.1	17.8	31.5	26.9	23.3	24.0	16.5	17.6
	77	91	KEGHLSPDIVAEQKK	1677.905	300	24.7	20.5	19.1	40.0	35.5	23.2	30.9	20.7	18.2
	77	91	KEGHLSPDIVAEQKK	1677.905	1200	33.6	23.7	22.3	47.2	42.7	25.3	40.4	24.6	19.3
	77	91	KEGHLSPDIVAEQKK	1677.905	3600	40.8	28.0	21.6	56.3	47.9	27.1	46.8	30.0	21.9
	77	91	KEGHLSPDIVAEQKK	1677.905	10800	47.4	36.8	25.3	64.8	55.9	35.5	57.2	38.1	26.1
	77	92	KEGHLSPDIVAEQKKL	1790.989	10	7.8	5.2	8.2	14.5	15.6	7.8	13.7	5.9	5.2
	77	92	KEGHLSPDIVAEQKKL	1790.989	30	11.4	9.1	11.9	17.6	18.2	12.3	15.2	9.4	8.7
	77	92	KEGHLSPDIVAEQKKL	1790.989	120	14.1	14.8	15.1	25.8	24.9	20.0	18.6	15.5	13.7
	77	92	KEGHLSPDIVAEQKKL	1790.989	300	18.6	18.3	16.9	33.7	32.6	20.9	24.8	18.8	16.0
	77	92	KEGHLSPDIVAEQKKL	1790.989	1200	26.4	22.2	20.0	40.0	39.9	21.2	32.9	23.1	16.6
	77	92	KEGHLSPDIVAEQKKL	1790.989	3600	33.8	26.0	20.5	49.1	44.9	25.3	39.1	27.7	17.3
	77	92	KEGHLSPDIVAEQKKL	1790.989	10800	40.1	33.4	22.4	57.2	51.3	32.8	48.8	34.0	21.3
	88	92	EQKKL	644.3857	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	88	92	EQKKL	644.3857	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	88	92	EQKKL	644.3857	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	88	92	EQKKL	644.3857	300	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	88	92	EQKKL	644.3857	1200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	88	92	EQKKL	644.3857	3600	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	88	92	EQKKL	644.3857	10800	0.0	0.0	0.0	1.9	0.0	0.0	0.0	0.0	0.0
	91	95	KLEAA	530.3064	10	10.3	10.8	12.9	10.9	8.1	9.1	5.0	5.1	13.4
	91	95	KLEAA	530.3064	30	21.8	21.6	23.8	21.8	17.7	19.5	12.7	14.1	23.9
	91	95	KLEAA	530.3064	120	31.4	32.0	33.0	32.3	31.4	32.0	26.9	29.0	32.5
	91	95	KLEAA	530.3064	300	33.5	33.4	34.4	34.1	33.8	34.5	33.0	33.0	33.8
	91	95	KLEAA	530.3064	1200	35.0	34.1	35.9	37.1	34.5	35.3	35.2	34.5	36.1
	91	95	KLEAA	530.3064	3600	41.6	39.1	41.0	48.8	39.2	38.9	42.4	40.7	43.6
	91	95	KLEAA	530.3064	10800	55.9	52.3	53.7	60.8	51.6	50.3	56.7	52.7	55.5

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	corrected %D								
						aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
	91	96	KLEAAD	645.3333	10	6.1	6.5	7.8	6.7	4.3	4.6	0.7	1.3	7.6
	91	96	KLEAAD	645.3333	30	15.4	14.8	16.3	15.0	11.7	12.7	7.8	9.4	16.9
	91	96	KLEAAD	645.3333	120	22.4	22.6	23.4	22.9	22.0	22.7	19.1	20.7	23.9
	91	96	KLEAAD	645.3333	300	24.0	23.5	24.0	23.8	23.9	25.3	23.3	23.6	23.4
	91	96	KLEAAD	645.3333	1200	24.6	24.1	24.2	25.6	24.3	24.0	24.3	24.5	23.5
	91	96	KLEAAD	645.3333	3600	30.9	28.9	28.4	36.8	28.8	27.1	31.2	30.1	30.9
	91	96	KLEAAD	645.3333	10800	42.7	40.8	40.8	46.1	40.5	37.8	42.7	40.7	42.8
	92	95	LEAA	402.2114	10	9.9	10.9	14.7	10.9	7.8	9.7	4.9	4.9	15.6
	92	95	LEAA	402.2114	30	23.8	23.1	29.4	24.5	21.0	23.6	14.9	15.8	29.9
	92	95	LEAA	402.2114	120	38.3	38.8	41.4	40.0	38.3	41.8	33.0	35.0	42.4
	92	95	LEAA	402.2114	300	42.5	41.9	44.0	44.0	42.7	44.1	42.0	42.0	43.9
	92	95	LEAA	402.2114	1200	46.1	46.0	47.4	52.0	49.5	47.0	48.8	47.0	50.1
	92	95	LEAA	402.2114	3600	57.1	56.1	56.7	70.6	55.2	54.0	59.1	57.1	63.5
	92	95	LEAA	402.2114	10800	79.3	75.4	78.0	88.2	76.3	73.1	80.3	76.0	81.0
	92	96	LEAAD	517.2384	10	11.5	12.0	13.9	11.6	9.1	10.2	6.7	6.6	14.1
	92	96	LEAAD	517.2384	30	21.6	21.7	23.8	21.5	17.8	19.5	13.2	14.8	23.7
	92	96	LEAAD	517.2384	120	30.5	31.0	32.4	31.6	30.1	31.8	26.1	27.9	32.0
	92	96	LEAAD	517.2384	300	32.2	32.4	33.9	33.1	32.5	33.9	32.0	31.9	33.3
	92	96	LEAAD	517.2384	1200	34.4	33.9	35.7	37.1	34.1	35.2	34.9	34.2	36.4
	92	96	LEAAD	517.2384	3600	42.3	40.6	42.5	48.9	40.4	40.8	42.5	41.9	44.7
	92	96	LEAAD	517.2384	10800	55.9	53.7	54.7	60.5	53.3	51.8	55.9	53.5	56.2
	96	102	DLVIFQF	880.4694	10	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.0	0.0
	96	102	DLVIFQF	880.4694	30	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	96	102	DLVIFQF	880.4694	120	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0
	96	102	DLVIFQF	880.4694	300	0.0	0.0	0.2	0.0	0.0	0.0	0.3	0.0	0.1
	96	102	DLVIFQF	880.4694	1200	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.1	0.1
	96	102	DLVIFQF	880.4694	3600	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
	96	102	DLVIFQF	880.4694	10800	0.1	0.0	0.2	0.1	0.0	0.0	0.4	0.0	0.0
	97	101	LVIFQ	618.3741	10	1.0	1.2	0.6	0.0	0.0	0.0	0.0	0.0	0.0
	97	101	LVIFQ	618.3741	30	0.4	0.4	0.1	0.0	0.2	0.0	0.0	0.0	0.0
	97	101	LVIFQ	618.3741	120	0.1	0.8	0.4	0.7	0.3	0.0	0.0	0.0	0.0
	97	101	LVIFQ	618.3741	300	0.4	0.4	0.0	0.0	0.3	0.0	0.0	0.0	0.0
	97	101	LVIFQ	618.3741	1200	1.2	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0
	97	101	LVIFQ	618.3741	3600	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	97	101	LVIFQ	618.3741	10800	0.6	0.4	0.0	0.6	0.0	0.8	0.0	0.0	0.0
	98	102	VIFQF	652.3584	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	98	102	VIFQF	652.3584	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	98	102	VIFQF	652.3584	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	98	102	VIFQF	652.3584	300	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	98	102	VIFQF	652.3584	1200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	98	102	VIFQF	652.3584	3600	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	98	102	VIFQF	652.3584	10800	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	103	107	PLQWF	689.3537	10	27.0	0.0	0.0	21.2	4.9	0.0	22.1	0.0	0.0
	103	107	PLQWF	689.3537	30	21.1	0.0	0.0	38.0	17.4	0.0	27.5	7.3	0.0
	103	107	PLQWF	689.3537	120	41.1	0.0	0.0	71.4	27.6	0.0	50.7	22.5	0.0
	103	107	PLQWF	689.3537	300	65.8	12.9	0.0	76.6	28.9	0.0	67.8	27.3	0.0
	103	107	PLQWF	689.3537	1200	83.9	28.6	0.0	76.9	28.0	0.0	73.0	29.1	0.0
	103	107	PLQWF	689.3537	3600	79.6	27.9	0.0	77.1	28.8	0.0	76.5	42.6	10.3
	103	107	PLQWF	689.3537	10800	80.3	27.9	0.0	71.7	44.9	1.1	76.1	59.1	12.6
	103	109	PLQWFGV	845.4436	10	20.2	0.0	0.0	33.5	16.2	0.0	28.7	1.6	0.0
	103	109	PLQWFGV	845.4436	30	25.8	0.0	0.0	44.4	27.3	0.0	36.7	6.2	0.0
	103	109	PLQWFGV	845.4436	120	44.9	4.3	0.0	63.2	35.4	1.2	53.8	17.6	0.0
	103	109	PLQWFGV	845.4436	300	58.6	13.3	0.0	70.7	36.7	4.6	69.9	29.0	1.8
	103	109	PLQWFGV	845.4436	1200	69.2	30.2	0.0	79.6	36.8	14.5	76.9	45.5	4.9
	103	109	PLQWFGV	845.4436	3600	69.9	35.2	3.5	84.3	40.4	17.8	81.8	60.1	17.7
	103	109	PLQWFGV	845.4436	10800	74.3	37.4	9.9	84.6	49.3	18.3	81.3	71.2	21.7
	107	113	FGVPAIL	715.4269	10	0.2	0.0	0.0	25.8	11.5	0.0	13.1	0.0	0.0
	107	113	FGVPAIL	715.4269	30	9.9	0.0	0.0	34.8	19.0	0.0	18.8	0.0	0.0
	107	113	FGVPAIL	715.4269	120	23.2	0.0	0.0	46.1	21.8	0.0	29.4	0.0	0.0
	107	113	FGVPAIL	715.4269	300	33.1	1.0	0.0	58.5	28.9	1.9	39.6	8.9	0.0
	107	113	FGVPAIL	715.4269	1200	39.8	15.7	0.0	63.0	52.8	17.1	51.6	21.9	0.0
	107	113	FGVPAIL	715.4269	3600	45.6	21.1	0.0	66.9	63.3	21.5	62.3	28.0	4.1
	107	113	FGVPAIL	715.4269	10800	57.8	25.0	10.8	67.6	66.3	28.2	67.3	38.1	14.1
	108	113	GVPAIL	568.3584	10	1.3	0.0	0.0	26.5	11.5	0.0	12.4	0.0	0.0
	108	113	GVPAIL	568.3584	30	9.3	0.0	0.0	34.1	17.3	0.0	16.6	0.0	0.0
	108	113	GVPAIL	568.3584	120	19.0	0.0	0.0	46.1	20.9	0.7	28.6	0.8	0.0
	108	113	GVPAIL	568.3584	300	32.5	2.4	0.0	60.4	29.9	5.1	40.2	9.7	0.0
	108	113	GVPAIL	568.3584	1200	40.9	14.5	0.2	63.9	52.1	17.5	52.6	19.4	0.5
	108	113	GVPAIL	568.3584	3600	43.2	19.6	2.1	64.2	63.7	23.4	61.5	23.8	7.3
	108	113	GVPAIL	568.3584	10800	55.5	26.4	12.5	64.8	65.7	31.9	65.8	35.1	17.2
	114	119	KGWFER	821.4184	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	119	KGWFER	821.4184	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	119	KGWFER	821.4184	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	119	KGWFER	821.4184	300	0.1	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0
	114	119	KGWFER	821.4184	1200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	119	KGWFER	821.4184	3600	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	119	KGWFER	821.4184	10800	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	121	KGWFERVF	1067.5552	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	121	KGWFERVF	1067.5552	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	121	KGWFERVF	1067.5552	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	121	KGWFERVF	1067.5552	300	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	121	KGWFERVF	1067.5552	1200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	121	KGWFERVF	1067.5552	3600	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	114	121	KGWFERVF	1067.5552	10800	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	corrected %D								
						aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
	116	121	WFERVF	882.4388	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	116	121	WFERVF	882.4388	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	116	121	WFERVF	882.4388	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	116	121	WFERVF	882.4388	300	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	116	121	WFERVF	882.4388	1200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	116	121	WFERVF	882.4388	3600	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	116	121	WFERVF	882.4388	10800	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	120	124	VFIGE	563.2955	10	21.3	20.3	0.0	52.0	31.7	0.0	42.9	23.1	0.0
	120	124	VFIGE	563.2955	30	32.5	22.7	0.0	51.7	39.0	0.0	45.6	22.4	0.0
	120	124	VFIGE	563.2955	120	45.3	25.5	0.0	64.8	57.8	0.0	48.5	33.1	0.0
	120	124	VFIGE	563.2955	300	49.0	34.4	0.0	81.3	75.6	1.9	57.0	43.5	0.0
	120	124	VFIGE	563.2955	1200	56.9	47.0	0.0	84.5	83.2	15.5	75.0	59.1	1.4
	120	124	VFIGE	563.2955	3600	72.9	61.3	2.6	86.1	81.5	36.7	83.1	74.1	4.9
	120	124	VFIGE	563.2955	10800	81.4	77.1	12.7	86.1	84.1	51.9	84.3	80.6	17.2
	120	125	VFIGEF	710.3639	10	30.4	31.5	0.0	63.9	49.5	0.0	49.7	35.3	0.0
	120	125	VFIGEF	710.3639	30	46.6	37.8	0.0	63.0	59.0	0.0	58.2	39.1	0.0
	120	125	VFIGEF	710.3639	120	57.7	41.1	0.0	73.2	75.0	0.0	63.7	49.9	0.0
	120	125	VFIGEF	710.3639	300	63.4	51.9	0.0	85.2	88.1	0.0	70.0	62.1	0.0
	120	125	VFIGEF	710.3639	1200	70.1	66.0	0.0	88.6	93.2	7.7	79.7	75.9	0.0
	120	125	VFIGEF	710.3639	3600	79.2	75.8	0.7	90.2	92.6	27.8	86.4	85.0	2.1
	120	125	VFIGEF	710.3639	10800	85.9	84.9	8.1	88.0	92.0	42.6	86.8	86.9	10.6
	122	125	IGEF	464.2271	10	62.0	64.4	0.0	97.0	80.3	0.0	76.8	68.9	0.0
	122	125	IGEF	464.2271	30	82.3	77.0	0.0	91.4	85.5	0.2	85.9	75.9	0.0
	122	125	IGEF	464.2271	120	88.5	80.2	0.0	91.0	91.7	0.7	90.3	82.0	0.0
	122	125	IGEF	464.2271	300	97.8	88.5	0.0	97.0	98.9	4.2	98.3	93.6	0.0
	122	125	IGEF	464.2271	1200	100.1	97.7	1.6	96.7	98.7	18.4	98.9	97.8	4.3
	122	125	IGEF	464.2271	3600	97.1	99.4	4.8	98.0	98.3	43.1	96.6	97.3	10.7
	122	125	IGEF	464.2271	10800	97.9	99.6	17.6	98.2	98.6	64.4	98.2	98.1	25.3
	122	127	IGEFAY	698.3275	10	39.1	60.4	0.0	92.4	84.6	0.0	71.4	75.2	0.0
	122	127	IGEFAY	698.3275	30	68.9	76.5	0.0	88.9	90.3	0.0	83.5	85.7	0.0
	122	127	IGEFAY	698.3275	120	85.6	88.9	0.5	88.6	94.1	0.0	88.2	89.1	0.9
	122	127	IGEFAY	698.3275	300	91.8	94.8	0.1	93.0	95.8	0.0	93.5	95.5	0.0
	122	127	IGEFAY	698.3275	1200	93.2	98.4	0.2	92.8	98.0	4.1	95.4	97.2	0.0
	122	127	IGEFAY	698.3275	3600	91.2	97.9	0.0	93.7	96.6	21.1	93.2	98.0	4.3
	122	127	IGEFAY	698.3275	10800	92.2	98.4	9.8	94.9	96.7	42.0	95.2	98.0	17.8
	125	129	FAYTY	663.2904	10	67.5	89.8	30.1	89.3	91.1	31.2	84.9	89.1	30.0
	125	129	FAYTY	663.2904	30	81.6	89.5	29.6	91.3	91.6	31.3	89.5	89.5	30.1
	125	129	FAYTY	663.2904	120	90.7	89.7	30.8	90.1	91.5	30.9	90.5	89.6	32.4
	125	129	FAYTY	663.2904	300	92.7	93.0	28.6	89.1	91.0	28.5	89.8	89.9	30.5
	125	129	FAYTY	663.2904	1200	91.4	92.7	30.1	89.6	92.0	29.7	89.1	90.2	30.6
	125	129	FAYTY	663.2904	3600	90.6	91.4	39.1	91.0	90.2	44.3	89.2	89.6	42.4
	125	129	FAYTY	663.2904	10800	89.9	90.4	56.8	88.5	91.6	69.7	89.3	88.9	60.9
	126	129	AYTY	516.222	10	85.0	94.9	42.4	95.8	91.2	42.7	90.1	92.1	43.4
	126	129	AYTY	516.222	30	89.2	92.6	41.8	93.0	89.9	42.9	90.7	92.9	43.4
	126	129	AYTY	516.222	120	92.0	91.3	42.9	93.2	90.4	43.0	91.4	90.9	44.0
	126	129	AYTY	516.222	300	94.9	95.5	43.5	95.5	92.8	43.5	95.0	94.5	45.3
	126	129	AYTY	516.222	1200	95.7	95.0	46.6	95.0	91.9	49.6	96.4	94.1	50.6
	126	129	AYTY	516.222	3600	94.1	95.0	60.7	95.9	92.6	65.5	94.7	94.9	64.5
	126	129	AYTY	516.222	10800	94.0	95.7	80.6	97.0	94.1	85.5	97.0	94.7	83.2
	126	131	AYTYAA	658.2962	10	77.3	85.0	44.8	94.6	95.5	47.7	92.9	93.7	46.5
	126	131	AYTYAA	658.2962	30	87.3	90.7	43.9	94.0	95.8	46.9	93.2	94.0	45.5
	126	131	AYTYAA	658.2962	120	93.6	91.3	52.1	94.6	96.1	56.4	93.1	93.7	55.3
	126	131	AYTYAA	658.2962	300	95.1	92.7	62.3	95.9	96.8	66.8	95.2	95.3	64.7
	126	131	AYTYAA	658.2962	1200	95.7	93.6	66.9	95.2	96.8	71.5	96.0	95.2	70.5
	126	131	AYTYAA	658.2962	3600	94.9	92.4	74.6	95.6	95.6	82.1	94.8	94.9	79.6
	126	131	AYTYAA	658.2962	10800	94.4	91.8	87.2	95.3	96.0	93.1	95.1	93.5	89.4
	127	130	YTYA	516.222	10	85.0	94.9	42.4	95.8	91.2	42.7	90.0	92.1	43.4
	127	130	YTYA	516.222	30	89.2	92.6	41.8	93.0	89.9	42.9	90.7	92.9	43.4
	127	130	YTYA	516.222	120	92.0	91.3	42.9	93.2	90.4	43.0	91.4	90.9	44.0
	127	130	YTYA	516.222	300	94.9	95.5	43.5	95.5	92.8	43.5	95.0	94.5	45.3
	127	130	YTYA	516.222	1200	95.7	95.0	46.6	95.0	91.9	49.6	96.4	94.1	50.6
	127	130	YTYA	516.222	3600	94.1	95.0	60.7	95.9	92.6	65.5	94.7	94.9	64.5
	127	130	YTYA	516.222	10800	94.0	95.7	80.6	97.0	94.1	85.5	97.0	94.7	83.2
	128	132	TYAAM	555.2363	10	50.7	54.3	32.7	75.9	65.8	34.6	92.5	87.3	34.3
	128	132	TYAAM	555.2363	30	68.1	70.7	44.8	81.8	77.9	46.0	92.7	90.2	45.2
	128	132	TYAAM	555.2363	120	89.4	89.2	72.5	92.6	93.6	73.8	94.1	93.3	71.7
	128	132	TYAAM	555.2363	300	94.4	95.1	88.2	95.7	97.2	90.6	94.7	95.6	88.3
	128	132	TYAAM	555.2363	1200	94.7	95.3	92.4	95.2	96.9	95.7	94.6	95.1	94.3
	128	132	TYAAM	555.2363	3600	93.8	95.7	92.7	94.7	96.2	94.7	94.4	94.3	94.3
	128	132	TYAAM	555.2363	10800	94.0	93.9	93.2	94.3	96.0	95.1	93.7	92.5	91.7
	129	132	YAAM	454.1886	10	46.8	50.4	26.4	77.7	63.2	27.9	95.4	89.2	27.5
	129	132	YAAM	454.1886	30	66.9	69.1	38.5	81.7	75.7	40.5	95.3	92.1	39.7
	129	132	YAAM	454.1886	120	91.0	92.2	71.9	94.1	92.7	73.0	96.1	94.2	71.9
	129	132	YAAM	454.1886	300	99.2	100.2	92.1	99.5	100.2	94.5	100.6	100.9	93.1
	129	132	YAAM	454.1886	1200	99.4	100.4	97.9	100.1	99.5	100.0	101.5	99.4	97.3
	129	132	YAAM	454.1886	3600	98.7	100.1	96.9	100.3	99.7	99.8	99.8	99.0	100.1
	129	132	YAAM	454.1886	10800	99.7	99.3	97.5	100.1	100.9	99.2	98.8	97.6	96.0
	132	141	MYDKGPFRSK	1227.607	10	15.3	12.5	12.3	37.8	13.6	12.6	59.9	47.5	12.4
	132	141	MYDKGPFRSK	1227.607	30	18.7	13.3	13.4	37.3	15.5	13.4	70.3	57.4	13.5
	132	141	MYDKGPFRSK	1227.607	120	23.4	18.6	17.3	46.7	23.1	18.4	78.5	69.9	19.7
	132	141	MYDKGPFRSK	1227.607	300	30.5	29.0	24.3	54.3	38.8	25.2	85.4	81.1	25.8
	132	141	MYDKGPFRSK	1227.607	1200	43.0	50.8	31.3	74.6	64.9	31.2	89.4	88.1	31.0
	132	141	MYDKGPFRSK	1227.607	3600	54.4	68.5	30.1	81.8	76.6	31.5	89.4	87.4	31.6
	132	141	MYDKGPFRSK	1227.607	10800	69.5	76.1	31.4	85.8	82.2	31.7	89.9	86.9	33.2

90

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	corrected %D								
						aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
	133	141	YDKGPFRSK	1096.5665	10	13.8	11.1	10.4	34.5	12.5	11.3	57.4	43.7	10.1
	133	141	YDKGPFRSK	1096.5665	30	15.5	11.6	11.2	34.8	13.2	10.7	69.0	53.3	10.2
	133	141	YDKGPFRSK	1096.5665	120	19.9	14.5	12.4	42.9	19.4	13.4	77.2	66.5	13.9
	133	141	YDKGPFRSK	1096.5665	300	25.5	24.7	18.4	49.6	34.9	19.0	83.3	77.3	18.0
	133	141	YDKGPFRSK	1096.5665	1200	37.5	47.9	24.7	70.3	61.4	24.5	88.3	85.1	24.6
	133	141	YDKGPFRSK	1096.5665	3600	49.1	66.4	23.9	79.3	73.1	25.2	88.3	85.0	25.8
	133	141	YDKGPFRSK	1096.5665	10800	65.3	73.2	25.2	82.7	77.4	25.3	88.7	84.4	25.2
	134	141	DKGPFRSK	933.5032	10	16.7	13.9	13.7	36.1	14.3	14.5	51.7	40.9	12.5
	134	141	DKGPFRSK	933.5032	30	19.2	13.3	13.7	35.7	14.9	13.8	62.7	51.4	13.9
	134	141	DKGPFRSK	933.5032	120	22.2	11.9	13.2	42.8	18.2	14.7	74.9	65.7	14.2
	134	141	DKGPFRSK	933.5032	300	25.4	21.5	14.7	49.0	30.6	15.7	82.3	76.1	14.9
	134	141	DKGPFRSK	933.5032	1200	33.0	40.8	16.0	67.8	54.1	15.8	89.3	86.1	16.9
	134	141	DKGPFRSK	933.5032	3600	47.0	60.4	16.2	77.1	65.7	16.8	89.9	86.5	15.4
	134	141	DKGPFRSK	933.5032	10800	62.6	68.8	14.7	81.1	73.8	16.1	89.0	85.9	13.4
	142	145	KAVL	429.2951	10	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.7	0.0
	142	145	KAVL	429.2951	30	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0
	142	145	KAVL	429.2951	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	142	145	KAVL	429.2951	300	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
	142	145	KAVL	429.2951	1200	0.0	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0
	142	145	KAVL	429.2951	3600	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	142	145	KAVL	429.2951	10800	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0
	142	148	KAVLSIT	730.4589	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	142	148	KAVLSIT	730.4589	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	142	148	KAVLSIT	730.4589	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	142	148	KAVLSIT	730.4589	300	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	142	148	KAVLSIT	730.4589	1200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	142	148	KAVLSIT	730.4589	3600	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	142	148	KAVLSIT	730.4589	10800	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	142	155	KAVLSITTGGSGSM	1307.6755	10	36.3	24.8	16.2	37.6	26.7	17.6	44.9	35.7	16.5
	142	155	KAVLSITTGGSGSM	1307.6755	30	37.9	27.9	20.9	39.5	31.9	21.8	46.7	37.0	21.2
	142	155	KAVLSITTGGSGSM	1307.6755	120	42.1	31.8	24.2	42.8	37.6	26.0	50.3	40.5	25.5
	142	155	KAVLSITTGGSGSM	1307.6755	300	45.6	37.9	26.9	45.9	42.5	28.7	51.1	44.0	27.4
	142	155	KAVLSITTGGSGSM	1307.6755	1200	48.5	43.6	32.0	50.2	45.0	34.1	51.8	45.4	33.0
	142	155	KAVLSITTGGSGSM	1307.6755	3600	50.4	43.7	34.6	51.2	44.4	36.2	51.7	46.2	34.7
	142	155	KAVLSITTGGSGSM	1307.6755	10800	50.7	43.0	35.8	49.6	44.2	38.1	50.5	47.1	35.6
	146	156	SITTGGSGSMY	1059.4543	10	47.6	32.9	21.0	53.7	34.4	21.2	69.5	56.4	21.8
	146	156	SITTGGSGSMY	1059.4543	30	50.5	37.9	27.2	53.8	41.3	26.9	70.3	58.7	27.7
	146	156	SITTGGSGSMY	1059.4543	120	54.8	43.1	31.8	58.5	49.2	32.2	75.4	63.9	33.5
	146	156	SITTGGSGSMY	1059.4543	300	60.9	51.3	35.6	63.0	55.6	35.9	78.4	69.7	36.6
	146	156	SITTGGSGSMY	1059.4543	1200	66.1	58.7	42.9	69.4	58.9	42.3	79.9	71.5	43.7
	146	156	SITTGGSGSMY	1059.4543	3600	68.1	58.7	45.7	72.2	58.4	44.8	78.7	72.7	46.6
	146	156	SITTGGSGSMY	1059.4543	10800	70.3	59.3	48.5	72.2	59.1	47.7	78.9	75.1	48.9
	149	155	TGGSGSM	595.2272	10	73.1	59.1	37.0	78.8	60.3	37.9	90.7	84.6	37.8
	149	155	TGGSGSM	595.2272	30	78.3	68.2	48.4	80.1	68.1	48.8	89.2	85.5	48.7
	149	155	TGGSGSM	595.2272	120	85.3	75.1	56.8	84.3	77.2	57.9	89.3	87.6	59.0
	149	155	TGGSGSM	595.2272	300	92.6	87.8	64.0	91.0	88.9	65.2	90.8	91.9	65.1
	149	155	TGGSGSM	595.2272	1200	95.1	93.9	76.4	93.3	94.3	77.3	90.3	91.8	77.5
	149	155	TGGSGSM	595.2272	3600	92.1	93.3	83.2	93.4	93.1	81.3	91.2	89.6	83.6
	149	155	TGGSGSM	595.2272	10800	92.3	92.8	89.2	91.7	92.9	89.2	90.9	87.2	87.9
	149	156	TGGSGSMY	758.2905	10	58.7	47.6	29.7	68.1	48.8	30.9	89.8	83.2	29.1
	149	156	TGGSGSMY	758.2905	30	61.9	54.3	39.7	69.0	55.3	39.7	87.7	83.8	38.8
	149	156	TGGSGSMY	758.2905	120	68.5	60.1	45.8	72.5	61.0	47.1	87.9	86.3	46.9
	149	156	TGGSGSMY	758.2905	300	74.9	69.8	51.7	76.4	70.7	52.9	90.0	91.2	52.7
	149	156	TGGSGSMY	758.2905	1200	77.3	74.4	62.6	80.5	75.2	63.8	92.1	91.8	63.7
	149	156	TGGSGSMY	758.2905	3600	75.9	73.8	67.4	80.6	74.7	66.7	90.5	89.8	68.3
	149	156	TGGSGSMY	758.2905	10800	77.2	74.4	71.4	81.0	75.5	71.8	90.0	87.3	71.6
	156	164	YSLQGIHGD	988.4614	10	23.1	15.0	12.2	41.4	16.5	13.2	86.5	80.0	13.7
	156	164	YSLQGIHGD	988.4614	30	27.3	19.5	18.3	40.6	19.9	17.6	84.9	81.3	19.0
	156	164	YSLQGIHGD	988.4614	120	28.8	21.8	19.8	42.6	22.3	20.3	86.0	85.6	24.1
	156	164	YSLQGIHGD	988.4614	300	32.0	25.6	23.6	48.2	26.6	24.8	88.2	91.3	27.7
	156	164	YSLQGIHGD	988.4614	1200	39.6	34.5	32.4	50.6	35.3	33.1	88.8	92.3	37.2
	156	164	YSLQGIHGD	988.4614	3600	44.4	37.4	36.4	55.6	39.1	38.5	87.7	90.5	39.4
	156	164	YSLQGIHGD	988.4614	10800	46.2	39.3	36.4	57.6	43.6	36.4	87.3	87.3	38.2
	157	164	SLQGIHGD	825.3981	10	14.2	2.6	2.8	36.7	2.9	2.5	89.5	77.4	4.6
	157	164	SLQGIHGD	825.3981	30	15.1	3.3	4.0	33.4	3.4	2.2	88.1	76.6	4.8
	157	164	SLQGIHGD	825.3981	120	16.6	5.6	5.7	35.2	5.5	5.2	88.5	82.6	9.5
	157	164	SLQGIHGD	825.3981	300	20.2	10.0	9.4	38.9	10.7	10.1	91.3	91.3	14.9
	157	164	SLQGIHGD	825.3981	1200	29.3	19.9	18.1	45.2	20.8	18.2	90.9	91.9	25.6
	157	164	SLQGIHGD	825.3981	3600	34.9	24.2	22.8	51.0	26.2	25.4	90.6	89.3	27.6
	157	164	SLQGIHGD	825.3981	10800	37.2	27.4	22.8	53.7	32.8	22.4	89.9	86.9	27.3
	157	165	SLQGIHGDM	956.4386	10	12.0	0.6	0.7	34.2	1.1	0.9	88.3	77.9	1.3
	157	165	SLQGIHGDM	956.4386	30	13.5	1.9	1.8	30.2	1.9	0.6	85.7	78.7	2.2
	157	165	SLQGIHGDM	956.4386	120	14.2	3.2	3.9	32.2	3.8	3.6	87.1	84.1	8.6
	157	165	SLQGIHGDM	956.4386	300	17.5	7.8	6.9	35.7	8.3	7.5	90.2	91.9	12.9
	157	165	SLQGIHGDM	956.4386	1200	27.0	17.3	16.4	41.9	18.1	16.2	91.3	92.4	22.8
	157	165	SLQGIHGDM	956.4386	3600	32.2	22.1	19.9	47.7	23.7	22.5	89.4	90.7	26.2
	157	165	SLQGIHGDM	956.4386	10800	36.9	29.7	21.4	53.2	34.3	21.2	90.0	89.3	30.0
	166	173	NVILWPIQ	981.5647	10	0.1	0.1	0.0	0.0	0.2	0.1	0.4	1.3	0.0
	166	173	NVILWPIQ	981.5647	30	0.3	0.1	0.1	0.2	0.3	0.0	0.8	6.1	0.0
	166	173	NVILWPIQ	981.5647	120	0.3	0.1	0.0	0.3	0.2	0.0	3.2	20.4	0.0
	166	173	NVILWPIQ	981.5647	300	0.4	0.1	0.0	0.1	0.3	0.0	11.2	31.7	0.0
	166	173	NVILWPIQ	981.5647	1200	0.3	0.2	0.0	6.1	0.3	0.0	36.8	47.1	0.0
	166	173	NVILWPIQ	981.5647	3600	0.0	0.1	0.0	14.5	0.2	0.0	66.7	66.2	0.0
	166	173	NVILWPIQ	981.5647	10800	0.0	0.0	0.0	20.2	0.2	0.1	85.5	84.8	0.0

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	corrected %D								
						aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
	167	173	VILWPIQ	867.5218	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	167	173	VILWPIQ	867.5218	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.7	0.0
	167	173	VILWPIQ	867.5218	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.2	0.0
	167	173	VILWPIQ	867.5218	300	0.0	0.0	0.0	0.0	0.0	0.0	6.0	35.4	0.0
	167	173	VILWPIQ	867.5218	1200	0.0	0.0	0.0	3.5	0.0	0.0	31.9	48.6	0.0
	167	173	VILWPIQ	867.5218	3600	0.0	0.0	0.0	11.5	0.0	0.0	59.5	64.8	0.0
	167	173	VILWPIQ	867.5218	10800	0.0	0.0	0.0	17.0	0.0	0.0	82.2	83.6	1.3
	168	173	ILWPIQ	768.4534	10	0.1	0.1	0.0	0.0	0.1	0.0	0.4	1.2	0.0
	168	173	ILWPIQ	768.4534	30	0.4	0.2	0.0	0.0	0.2	0.0	0.9	3.3	0.0
	168	173	ILWPIQ	768.4534	120	0.5	0.1	0.0	0.3	0.1	0.0	2.6	13.7	0.0
	168	173	ILWPIQ	768.4534	300	0.7	0.1	0.0	0.0	0.1	0.0	7.9	24.7	0.0
	168	173	ILWPIQ	768.4534	1200	1.2	0.3	0.0	3.9	0.2	0.0	27.5	38.6	0.0
	168	173	ILWPIQ	768.4534	3600	1.4	0.2	0.0	10.5	0.1	0.0	54.5	58.3	0.0
	168	173	ILWPIQ	768.4534	10800	1.4	0.1	0.0	14.3	0.0	0.0	77.7	80.2	0.0
	168	175	ILWPIQSG	912.5069	10	0.1	0.0	0.0	0.0	0.0	0.2	0.0	0.8	0.1
	168	175	ILWPIQSG	912.5069	30	0.3	0.2	0.0	0.2	0.0	0.1	0.4	2.2	0.0
	168	175	ILWPIQSG	912.5069	120	0.2	0.3	0.0	0.1	0.2	0.0	1.8	7.7	0.1
	168	175	ILWPIQSG	912.5069	300	0.2	0.0	0.0	0.1	0.1	0.0	6.2	16.3	0.0
	168	175	ILWPIQSG	912.5069	1200	0.2	0.4	0.0	4.2	0.1	0.0	22.6	32.4	0.2
	168	175	ILWPIQSG	912.5069	3600	0.0	0.0	0.0	12.0	0.0	0.0	49.0	54.5	0.0
	168	175	ILWPIQSG	912.5069	10800	0.0	0.0	0.0	17.4	0.0	0.0	80.1	79.2	0.0
	174	178	SGILH	525.2911	10	0.6	0.4	0.5	1.0	0.8	0.5	1.3	0.9	0.7
	174	178	SGILH	525.2911	30	0.6	0.1	0.5	0.6	1.0	0.4	1.6	0.8	0.6
	174	178	SGILH	525.2911	120	0.4	0.4	0.4	0.7	0.6	0.4	1.6	1.2	0.5
	174	178	SGILH	525.2911	300	0.8	0.6	0.5	1.0	0.6	0.7	5.1	2.6	0.7
	174	178	SGILH	525.2911	1200	0.1	0.6	0.8	2.4	0.8	0.8	9.9	11.0	0.7
	174	178	SGILH	525.2911	3600	1.3	0.7	0.5	6.6	0.3	1.8	27.4	24.6	2.5
	174	178	SGILH	525.2911	10800	1.4	1.0	0.6	13.7	1.6	1.0	51.9	37.5	2.5
	174	181	SGILHFCG	832.3902	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0
	174	181	SGILHFCG	832.3902	30	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
	174	181	SGILHFCG	832.3902	120	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.2	0.0
	174	181	SGILHFCG	832.3902	300	0.0	0.0	0.0	0.0	0.0	0.0	2.1	1.3	0.0
	174	181	SGILHFCG	832.3902	1200	0.0	0.0	0.0	0.7	0.0	0.0	8.9	7.7	0.0
	174	181	SGILHFCG	832.3902	3600	0.0	0.0	0.0	5.9	0.0	0.0	27.5	20.0	0.0
	174	181	SGILHFCG	832.3902	10800	0.0	0.0	0.0	11.8	0.0	0.0	55.4	35.6	0.0
	178	181	HFCG	462.1685	10	0.6	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0
	178	181	HFCG	462.1685	30	0.6	0.0	0.6	0.0	0.0	0.0	0.6	0.1	0.0
	178	181	HFCG	462.1685	120	0.7	0.8	0.0	0.0	0.0	0.0	0.7	0.2	0.0
	178	181	HFCG	462.1685	300	0.1	0.5	0.0	0.0	0.0	0.0	2.5	0.0	0.0
	178	181	HFCG	462.1685	1200	0.5	0.5	0.2	4.3	0.0	0.0	19.7	4.0	0.0
	178	181	HFCG	462.1685	3600	0.1	0.6	0.0	16.1	0.0	0.0	49.2	20.8	0.0
	178	181	HFCG	462.1685	10800	0.8	0.6	0.2	20.3	0.0	0.0	79.3	49.2	0.0
	178	182	HFCGF	609.2369	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	178	182	HFCGF	609.2369	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	178	182	HFCGF	609.2369	120	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	178	182	HFCGF	609.2369	300	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	178	182	HFCGF	609.2369	1200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	178	182	HFCGF	609.2369	3600	0.0	0.0	0.0	0.0	0.0	0.0	6.0	1.8	0.0
	178	182	HFCGF	609.2369	10800	0.0	0.0	0.0	0.0	0.0	0.0	21.6	19.0	0.0
	182	185	FQVL	505.29	10	0.0	0.0	0.0	0.1	0.3	0.0	0.3	0.3	0.1
	182	185	FQVL	505.29	30	0.0	0.0	0.0	0.0	0.1	0.0	0.4	0.3	0.2
	182	185	FQVL	505.29	120	0.1	0.0	0.0	0.2	0.0	0.0	0.3	0.5	0.2
	182	185	FQVL	505.29	300	0.0	0.0	0.0	0.2	0.3	0.1	0.5	0.8	0.2
	182	185	FQVL	505.29	1200	0.1	0.2	0.0	1.9	0.3	0.0	6.3	10.9	0.4
	182	185	FQVL	505.29	3600	0.5	0.0	0.0	7.4	0.2	0.0	25.3	30.4	0.7
	182	185	FQVL	505.29	10800	1.5	0.3	0.2	16.5	1.3	0.5	43.0	43.2	1.3
WT/S82D	182	189	FQVLEPQL	972.528	10	0.4	0.4	0.0	0.0	0.7	0.0			
WT/S82D	182	189	FQVLEPQL	972.528	30	1.3	1.5	0.2	0.8	1.7	0.0			
WT/S82D	182	189	FQVLEPQL	972.528	120	5.4	7.8	5.1	7.4	7.7	4.9			
WT/S82D	182	189	FQVLEPQL	972.528	300	11.1	15.3	13.1	12.4	15.3	13.0			
WT/S82D	182	189	FQVLEPQL	972.528	1200	18.0	18.5	17.9	19.0	19.1	18.5			
WT/S82D	182	189	FQVLEPQL	972.528	3600	19.0	18.5	18.1	22.1	19.1	18.5			
WT/S82D	182	189	FQVLEPQL	972.528	10800	19.4	18.6	18.1	25.5	19.5	18.6			
WT/S82D	182	190	FQVLEPQLT	1073.5757	10	1.4	1.5	0.6	1.6	2.2	1.2			
WT/S82D	182	190	FQVLEPQLT	1073.5757	30	2.6	3.2	1.6	3.0	3.4	1.5			
WT/S82D	182	190	FQVLEPQLT	1073.5757	120	6.3	7.9	5.4	8.5	7.8	5.4			
WT/S82D	182	190	FQVLEPQLT	1073.5757	300	11.0	13.5	11.1	12.4	14.1	11.8			
WT/S82D	182	190	FQVLEPQLT	1073.5757	1200	17.1	16.4	15.2	21.2	17.2	16.2			
WT/S82D	182	190	FQVLEPQLT	1073.5757	3600	18.5	16.3	15.5	25.5	17.2	16.5			
WT/S82D	182	190	FQVLEPQLT	1073.5757	10800	20.0	16.8	15.8	29.3	18.0	16.4			
WT/S82D	182	191	FQVLEPQLTY	1236.639	10	3.3	0.7	0.5	6.6	1.1	0.3			
WT/S82D	182	191	FQVLEPQLTY	1236.639	30	7.6	3.0	1.8	10.0	2.9	1.7			
WT/S82D	182	191	FQVLEPQLTY	1236.639	120	15.5	13.1	10.3	18.0	12.6	9.8			
WT/S82D	182	191	FQVLEPQLTY	1236.639	300	20.4	23.5	20.8	22.7	23.1	20.6			
WT/S82D	182	191	FQVLEPQLTY	1236.639	1200	26.4	27.1	27.4	29.2	27.4	27.7			
WT/S82D	182	191	FQVLEPQLTY	1236.639	3600	27.4	26.7	27.2	32.0	27.3	27.7			
WT/S82D	182	191	FQVLEPQLTY	1236.639	10800	28.2	26.8	27.3	35.7	27.6	27.6			
WT/S82D	183	190	QVLEPQLT	926.5073	10	1.1	1.3	0.6	1.4	1.5	1.0			
WT/S82D	183	190	QVLEPQLT	926.5073	30	2.3	2.8	1.6	2.9	2.8	1.3			
WT/S82D	183	190	QVLEPQLT	926.5073	120	6.4	9.0	7.7	9.1	8.6	7.5			
WT/S82D	183	190	QVLEPQLT	926.5073	300	12.5	16.1	14.2	14.4	16.0	13.8			
WT/S82D	183	190	QVLEPQLT	926.5073	1200	20.2	19.4	19.3	23.2	19.8	19.2			
WT/S82D	183	190	QVLEPQLT	926.5073	3600	21.5	19.3	18.7	25.5	19.8	19.5			
WT/S82D	183	190	QVLEPQLT	926.5073	10800	22.3	19.2	18.7	27.3	19.7	19.0			

						corrected %D								
Proteoform	From	To	Sequence	M_{mono}	Time(sec)	aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
WT/S82D	183	191	QVLEPQLTY	1089.5706	10	5.0	1.9	1.9	8.9	1.7	0.6			
WT/S82D	183	191	QVLEPQLTY	1089.5706	30	10.4	4.7	2.8	12.7	4.0	3.0			
WT/S82D	183	191	QVLEPQLTY	1089.5706	120	19.3	16.7	12.3	21.9	15.5	11.7			
WT/S82D	183	191	QVLEPQLTY	1089.5706	300	25.0	27.7	23.7	27.0	26.8	24.2			
WT/S82D	183	191	QVLEPQLTY	1089.5706	1200	31.9	31.5	31.3	34.3	31.7	31.9			
WT/S82D	183	191	QVLEPQLTY	1089.5706	3600	32.8	31.3	31.3	36.3	31.3	31.6			
WT/S82D	183	191	QVLEPQLTY	1089.5706	10800	33.7	31.8	31.5	38.3	31.7	31.5			
P187S	186	190	ESQLT	576.2755	10							11.0	1.4	0.7
P187S	186	190	ESQLT	576.2755	30							21.9	10.5	3.2
P187S	186	190	ESQLT	576.2755	120							33.8	25.6	18.3
P187S	186	190	ESQLT	576.2755	300							47.1	30.6	29.1
P187S	186	190	ESQLT	576.2755	1200							61.3	31.7	31.2
P187S	186	190	ESQLT	576.2755	3600							62.8	37.2	31.0
P187S	186	190	ESQLT	576.2755	10800							66.4	52.4	31.8
P187S	186	191	ESQLTY	739.7706	10							31.3	8.3	0.0
P187S	186	191	ESQLTY	739.7706	30							41.0	21.7	0.6
P187S	186	191	ESQLTY	739.7706	120							48.7	42.9	21.0
P187S	186	191	ESQLTY	739.7706	300							59.0	48.8	40.3
P187S	186	191	ESQLTY	739.7706	1200							70.7	48.2	49.4
P187S	186	191	ESQLTY	739.7706	3600							71.2	51.2	49.6
P187S	186	191	ESQLTY	739.7706	10800							72.2	62.2	48.9
	190	201	TYSIGHTPADAR	1287.6208	10	73.2	49.6	50.2	85.4	50.1	52.2	90.7	56.4	50.3
	190	201	TYSIGHTPADAR	1287.6208	30	80.1	54.8	55.5	83.8	57.5	59.3	90.2	67.6	56.3
	190	201	TYSIGHTPADAR	1287.6208	120	82.3	61.7	62.9	86.0	63.9	64.8	90.6	83.0	65.0
	190	201	TYSIGHTPADAR	1287.6208	300	87.3	72.6	69.7	90.1	75.2	70.9	92.6	92.0	71.3
	190	201	TYSIGHTPADAR	1287.6208	1200	92.1	83.4	79.5	92.8	84.6	80.0	93.5	94.1	79.4
	190	201	TYSIGHTPADAR	1287.6208	3600	90.0	84.6	82.8	91.5	85.2	82.3	92.3	91.9	82.8
	190	201	TYSIGHTPADAR	1287.6208	10800	90.6	87.0	84.2	90.8	89.3	84.7	92.0	90.4	83.9
	191	201	YSIGHTPADAR	1186.5731	10	77.8	54.1	54.8	88.4	56.2	56.2	85.3	56.2	54.0
	191	201	YSIGHTPADAR	1186.5731	30	86.0	60.0	60.3	88.3	61.8	62.6	84.7	66.0	61.9
	191	201	YSIGHTPADAR	1186.5731	120	87.5	67.4	68.3	88.3	70.7	71.9	85.3	80.1	70.9
	191	201	YSIGHTPADAR	1186.5731	300	90.5	78.7	76.4	91.2	82.4	78.5	88.1	89.5	77.0
	191	201	YSIGHTPADAR	1186.5731	1200	91.5	90.1	86.3	91.5	92.5	88.2	89.3	92.1	86.4
	191	201	YSIGHTPADAR	1186.5731	3600	89.1	90.5	90.0	91.6	92.2	89.0	87.2	89.7	87.7
	191	201	YSIGHTPADAR	1186.5731	10800	89.3	90.4	91.1	90.0	92.2	92.1	87.6	87.7	92.3
	191	203	YSIGHTPADARIQ	1427.7157	10	67.5	43.4	44.3	80.6	44.7	44.7	82.4	45.6	44.3
	191	203	YSIGHTPADARIQ	1427.7157	30	80.9	49.0	49.7	85.3	50.5	50.3	84.7	55.3	50.7
	191	203	YSIGHTPADARIQ	1427.7157	120	87.2	59.7	61.2	87.7	62.5	62.5	86.8	74.1	63.1
	191	203	YSIGHTPADARIQ	1427.7157	300	90.5	76.6	75.4	90.7	79.4	76.4	90.6	89.6	78.7
	191	203	YSIGHTPADARIQ	1427.7157	1200	91.4	94.1	91.3	90.7	94.9	91.3	91.9	96.2	92.1
	191	203	YSIGHTPADARIQ	1427.7157	3600	89.5	94.5	94.6	90.8	94.7	92.9	89.5	94.3	96.7
	191	203	YSIGHTPADARIQ	1427.7157	10800	90.2	94.2	95.7	91.0	93.8	95.0	90.7	93.1	95.2
	192	201	SIGHTPADAR	1023.5097	10	77.3	60.6	61.1	87.0	62.0	60.0	86.4	62.0	61.6
	192	201	SIGHTPADAR	1023.5097	30	84.8	67.6	66.7	86.7	67.0	69.0	84.8	68.3	67.4
	192	201	SIGHTPADAR	1023.5097	120	87.7	71.9	75.6	86.2	73.0	75.9	84.7	78.4	78.7
	192	201	SIGHTPADAR	1023.5097	300	88.8	85.1	85.8	87.2	85.4	86.0	87.8	87.6	86.9
	192	201	SIGHTPADAR	1023.5097	1200	87.9	88.2	92.9	90.2	89.0	89.5	89.8	89.6	89.0
	192	201	SIGHTPADAR	1023.5097	3600	87.8	89.9	89.6	88.7	86.3	84.9	86.8	86.1	87.2
	192	201	SIGHTPADAR	1023.5097	10800	88.8	90.0	89.5	88.1	88.2	89.0	88.1	87.4	91.7
	192	203	SIGHTPADARIQ	1264.6524	10	66.7	46.6	47.7	79.4	45.5	46.0	82.3	45.5	48.1
	192	203	SIGHTPADARIQ	1264.6524	30	81.9	54.4	53.6	85.4	51.5	54.4	84.8	54.7	56.6
	192	203	SIGHTPADARIQ	1264.6524	120	86.9	62.2	67.1	86.8	63.1	68.1	86.9	70.8	71.5
	192	203	SIGHTPADARIQ	1264.6524	300	91.0	78.0	80.9	89.3	81.2	83.5	90.1	85.8	84.2
	192	203	SIGHTPADARIQ	1264.6524	1200	92.5	91.0	96.3	88.3	91.9	93.4	92.9	91.5	95.0
	192	203	SIGHTPADARIQ	1264.6524	3600	87.9	89.8	92.9	89.9	89.0	90.9	90.7	89.3	95.0
	192	203	SIGHTPADARIQ	1264.6524	10800	90.0	90.5	93.0	89.5	91.7	94.3	90.8	90.1	94.8
	202	205	IQIL	485.3213	10	9.0	0.0	0.0	25.3	0.9	0.1	44.8	0.4	0.0
	202	205	IQIL	485.3213	30	24.0	0.0	0.0	42.2	0.5	0.0	64.5	1.3	0.0
	202	205	IQIL	485.3213	120	47.0	0.8	3.8	75.4	4.2	5.7	87.6	11.7	7.3
	202	205	IQIL	485.3213	300	72.5	11.0	15.4	92.4	15.5	17.7	95.4	28.7	18.5
	202	205	IQIL	485.3213	1200	94.2	26.5	31.9	95.2	29.8	31.9	97.3	58.8	36.7
	202	205	IQIL	485.3213	3600	94.0	47.0	53.8	96.2	52.8	54.7	95.4	85.5	63.6
	202	205	IQIL	485.3213	10800	95.9	78.9	87.1	97.4	82.7	85.4	97.4	96.7	93.7
	204	211	ILEGWKKR	1028.6131	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	204	211	ILEGWKKR	1028.6131	30	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0
	204	211	ILEGWKKR	1028.6131	120	2.2	0.0	0.0	5.4	0.0	0.0	7.6	0.0	0.0
	204	211	ILEGWKKR	1028.6131	300	7.0	0.0	0.0	10.3	0.0	0.0	19.2	0.0	0.0
	204	211	ILEGWKKR	1028.6131	1200	13.3	0.5	4.3	23.6	0.2	2.6	35.0	5.3	4.0
	204	211	ILEGWKKR	1028.6131	3600	24.7	6.4	9.4	37.5	6.5	8.3	39.2	8.2	9.5
	204	211	ILEGWKKR	1028.6131	10800	37.3	8.0	10.8	40.0	7.9	9.9	42.7	16.1	12.0
	204	212	ILEGWKKRL	1141.6972	10	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.1	0.0
	204	212	ILEGWKKRL	1141.6972	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	204	212	ILEGWKKRL	1141.6972	120	0.0	0.0	0.0	0.0	0.0	0.0	1.9	0.0	0.0
	204	212	ILEGWKKRL	1141.6972	300	0.2	0.0	0.0	3.3	0.0	0.0	9.9	0.0	0.0
	204	212	ILEGWKKRL	1141.6972	1200	5.2	0.4	0.0	14.6	1.0	1.2	22.3	3.2	1.5
	204	212	ILEGWKKRL	1141.6972	3600	14.0	3.0	5.5	24.3	3.5	4.5	25.0	5.5	3.1
	204	212	ILEGWKKRL	1141.6972	10800	23.6	4.8	6.6	27.1	5.4	5.5	28.3	11.2	7.0
	206	211	EGWKKR	802.445	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	206	211	EGWKKR	802.445	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	206	211	EGWKKR	802.445	120	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0
	206	211	EGWKKR	802.445	300	0.0	0.0	0.0	1.8	0.0	0.0	9.3	0.0	0.4
	206	211	EGWKKR	802.445	1200	4.9	0.0	0.0	13.4	0.0	0.2	21.3	0.0	1.0
	206	211	EGWKKR	802.445	3600	14.2	0.0	1.5	22.0	0.0	1.0	24.5	1.5	0.3
	206	211	EGWKKR	802.445	10800	22.6	0.1	2.2	25.6	0.8	1.7	28.8	7.4	3.1

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
											corrected %D			
	206	212	EGWKKRL	915.529	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	206	212	EGWKKRL	915.529	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	206	212	EGWKKRL	915.529	120	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0
	206	212	EGWKKRL	915.529	300	0.0	0.0	0.0	0.0	0.0	0.0	3.9	0.0	0.0
	206	212	EGWKKRL	915.529	1200	0.0	0.0	0.0	6.6	0.0	0.0	14.5	0.0	0.0
	206	212	EGWKKRL	915.529	3600	8.1	0.0	0.1	14.3	0.0	0.0	15.7	0.0	0.0
	206	212	EGWKKRL	915.529	10800	15.1	0.0	0.0	15.2	0.0	0.0	15.6	0.1	0.0
	212	216	LENIW	673.3435	10	7.2	7.6	5.7	10.9	9.8	6.7	18.2	16.7	9.3
	212	216	LENIW	673.3435	30	29.2	30.5	28.6	30.9	31.4	29.5	39.0	41.7	35.5
	212	216	LENIW	673.3435	120	53.0	57.4	55.9	54.3	57.9	57.7	53.6	57.1	57.2
	212	216	LENIW	673.3435	300	56.6	59.9	60.9	58.6	60.3	63.3	58.4	61.2	61.1
	212	216	LENIW	673.3435	1200	73.5	78.5	79.6	75.8	79.7	81.4	77.5	79.9	82.0
	212	216	LENIW	673.3435	3600	87.4	91.4	92.4	90.5	92.5	93.8	89.4	91.7	93.0
	212	216	LENIW	673.3435	10800	90.5	93.3	94.1	92.2	94.4	95.2	91.2	91.8	93.9
	212	218	LENIWDE	917.4131	10	8.8	9.2	7.7	11.3	11.4	8.8	17.6	16.8	10.4
	212	218	LENIWDE	917.4131	30	22.0	22.3	20.7	23.5	24.1	21.5	28.5	29.5	24.4
	212	218	LENIWDE	917.4131	120	35.6	38.5	37.8	37.4	40.2	40.0	38.6	41.1	39.8
	212	218	LENIWDE	917.4131	300	43.9	47.1	46.8	46.8	49.1	50.6	49.6	51.5	49.6
	212	218	LENIWDE	917.4131	1200	63.4	67.5	68.1	66.9	70.7	71.8	68.9	72.2	72.5
	212	218	LENIWDE	917.4131	3600	77.5	82.5	84.0	82.3	86.4	86.5	80.1	83.4	83.8
	212	218	LENIWDE	917.4131	10800	86.2	89.8	91.1	90.3	93.5	94.4	87.9	88.9	91.8
	212	219	LENIWDET	1018.4607	10	25.1	25.6	25.0	27.8	27.6	26.3	32.3	31.8	27.3
	212	219	LENIWDET	1018.4607	30	34.7	35.6	35.2	36.3	36.6	36.0	40.1	41.0	38.1
	212	219	LENIWDET	1018.4607	120	45.3	49.1	49.2	46.9	49.6	50.8	47.9	50.4	50.7
	212	219	LENIWDET	1018.4607	300	53.0	56.4	57.6	55.0	57.4	59.5	57.5	59.2	58.9
	212	219	LENIWDET	1018.4607	1200	68.9	73.5	74.8	70.8	74.6	76.9	73.2	75.9	77.2
	212	219	LENIWDET	1018.4607	3600	81.4	86.0	87.7	84.0	87.7	88.7	82.9	85.6	88.1
	212	219	LENIWDET	1018.4607	10800	89.1	93.1	94.0	91.7	94.4	95.6	90.3	90.9	93.7
	212	221	LENIWDETPL	1228.5976	10	21.5	22.0	21.1	23.5	23.5	21.8	26.4	26.1	21.5
	212	221	LENIWDETPL	1228.5976	30	29.2	30.3	29.5	30.7	31.3	30.0	33.1	33.7	29.9
	212	221	LENIWDETPL	1228.5976	120	38.8	41.2	40.9	40.3	41.8	42.2	40.2	42.2	39.8
	212	221	LENIWDETPL	1228.5976	300	44.8	47.6	48.0	47.4	48.2	49.2	50.3	52.4	46.4
	212	221	LENIWDETPL	1228.5976	1200	58.8	62.0	63.1	63.6	63.0	64.2	69.1	72.5	62.3
	212	221	LENIWDETPL	1228.5976	3600	70.6	72.9	74.1	76.2	74.1	75.1	81.4	83.1	71.0
	212	221	LENIWDETPL	1228.5976	10800	77.3	78.9	79.3	82.9	80.3	81.0	87.6	87.6	75.8
	212	222	LENIWDETPLY	1391.6609	10	20.5	19.7	18.9	24.1	21.2	19.5	34.0	31.2	20.8
	212	222	LENIWDETPLY	1391.6609	30	28.4	26.9	26.3	31.6	27.8	26.2	43.1	40.8	29.3
	212	222	LENIWDETPLY	1391.6609	120	36.5	36.2	36.1	40.5	36.8	36.3	50.2	50.6	42.4
	212	222	LENIWDETPLY	1391.6609	300	42.4	42.5	43.2	46.0	43.1	43.6	59.2	60.9	52.8
	212	222	LENIWDETPLY	1391.6609	1200	56.4	57.2	58.2	62.8	58.0	58.5	75.7	78.6	69.2
	212	222	LENIWDETPLY	1391.6609	3600	69.8	71.1	72.4	76.8	72.0	72.5	86.6	87.4	76.5
	212	222	LENIWDETPLY	1391.6609	10800	79.1	79.5	80.5	83.6	80.6	80.9	91.1	90.5	80.0
	213	216	ENIW	560.2595	10	9.4	8.5	5.7	14.8	10.1	6.8	26.5	23.6	10.6
	213	216	ENIW	560.2595	30	29.8	26.3	24.0	32.6	27.6	24.3	42.1	39.8	30.3
	213	216	ENIW	560.2595	120	45.5	44.6	44.8	46.5	45.7	45.5	46.7	46.5	45.5
	213	216	ENIW	560.2595	300	48.0	49.4	50.3	50.5	50.3	51.5	50.7	50.8	51.6
	213	216	ENIW	560.2595	1200	70.3	72.0	73.5	72.9	72.8	74.8	74.4	75.5	78.5
	213	216	ENIW	560.2595	3600	87.3	88.5	89.1	90.3	90.0	92.0	90.6	90.4	91.7
	213	216	ENIW	560.2595	10800	91.2	91.5	91.7	93.5	92.3	92.9	92.8	91.6	94.4
	213	218	ENIWDE	804.329	10	3.7	2.7	1.3	7.2	4.6	2.0	14.7	12.9	4.7
	213	218	ENIWDE	804.329	30	16.4	13.5	12.4	19.0	15.3	13.7	24.6	22.9	17.5
	213	218	ENIWDE	804.329	120	27.2	27.3	27.3	29.2	28.6	28.4	29.7	30.0	28.3
	213	218	ENIWDE	804.329	300	34.1	35.5	35.1	36.7	37.0	38.3	39.6	40.6	38.0
	213	218	ENIWDE	804.329	1200	55.7	59.9	60.9	60.0	62.6	63.6	62.2	65.0	65.9
	213	218	ENIWDE	804.329	3600	72.6	77.8	79.5	78.0	81.2	82.3	74.9	79.0	80.8
	213	218	ENIWDE	804.329	10800	83.5	87.5	88.9	89.4	91.8	93.2	85.5	86.5	89.1
	213	221	ENIWDETPL	1115.5135	10	22.3	21.9	20.4	24.7	23.1	21.6	28.2	27.4	22.5
	213	221	ENIWDETPL	1115.5135	30	29.0	28.1	26.6	31.4	29.4	28.2	34.1	33.7	29.1
	213	221	ENIWDETPL	1115.5135	120	36.9	37.2	36.0	39.2	38.4	38.2	40.1	40.5	36.9
	213	221	ENIWDETPL	1115.5135	300	42.2	42.8	42.7	45.4	44.0	44.5	49.5	50.3	44.2
	213	221	ENIWDETPL	1115.5135	1200	55.5	57.3	57.7	61.2	58.8	60.2	69.6	71.7	60.1
	213	221	ENIWDETPL	1115.5135	3600	67.0	68.4	69.6	73.3	71.0	71.8	82.0	83.6	70.3
	213	221	ENIWDETPL	1115.5135	10800	73.8	74.8	75.7	81.5	77.0	78.0	89.0	88.7	75.3
	213	222	ENIWDETPLY	1278.5768	10	18.9	17.5	16.5	22.6	18.9	17.1	33.4	29.6	18.7
	213	222	ENIWDETPLY	1278.5768	30	25.8	23.2	22.6	29.8	24.3	22.6	42.5	38.5	25.8
	213	222	ENIWDETPLY	1278.5768	120	33.2	31.4	31.0	37.2	32.3	31.9	48.6	47.2	37.9
	213	222	ENIWDETPLY	1278.5768	300	38.3	36.9	37.4	42.2	38.0	38.0	56.6	57.3	49.5
	213	222	ENIWDETPLY	1278.5768	1200	52.8	52.7	53.4	60.2	54.5	54.4	73.7	75.7	64.8
	213	222	ENIWDETPLY	1278.5768	3600	67.2	67.7	68.6	74.7	69.6	69.7	85.0	84.8	73.1
	213	222	ENIWDETPLY	1278.5768	10800	77.0	77.0	77.5	82.2	79.5	79.0	90.0	90.4	78.5
	215	221	IWDETPL	872.428	10	34.9	35.3	33.7	35.4	35.5	35.1	34.2	34.8	34.0
	215	221	IWDETPL	872.428	30	34.4	35.6	34.5	34.9	35.7	35.4	33.5	34.5	33.8
	215	221	IWDETPL	872.428	120	35.2	34.5	34.3	34.9	35.6	34.8	34.4	36.0	33.1
	215	221	IWDETPL	872.428	300	35.6	36.3	35.6	36.0	36.6	37.5	42.7	46.8	37.5
	215	221	IWDETPL	872.428	1200	40.1	44.2	45.5	46.1	44.6	46.4	64.1	70.3	47.4
	215	221	IWDETPL	872.428	3600	46.8	52.6	55.9	57.5	52.3	55.6	77.8	83.0	58.3
	215	221	IWDETPL	872.428	10800	57.8	62.2	65.1	68.9	62.0	65.1	88.9	91.5	65.9
	219	222	TPLY	492.2584	10	1.2	0.3	0.5	8.0	0.2	0.7	33.1	23.2	1.4
	219	222	TPLY	492.2584	30	3.8	0.3	0.8	13.5	0.7	0.7	43.2	35.0	4.6
	219	222	TPLY	492.2584	120	4.5	0.9	0.5	15.7	0.7	1.1	46.4	43.7	20.9
	219	222	TPLY	492.2584	300	5.1	1.2	2.4	20.0	1.1	2.1	54.4	56.9	38.6
	219	222	TPLY	492.2584	1200	16.1	9.1	11.5	36.1	11.7	12.1	78.4	80.5	46.8
	219	222	TPLY	492.2584	3600	32.2	29.8	31.2	51.3	31.9	31.9	91.5	90.5	46.8
	219	222	TPLY	492.2584	10800	45.8	44.4	44.7	59.7	45.7	45.7	92.5	90.6	46.8

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
	222	228	YFAPSSL	783.3803	10	28.4	20.6	18.1	48.0	22.8	18.7	87.6	79.8	16.6
	222	228	YFAPSSL	783.3803	30	32.3	24.8	22.4	49.6	27.4	21.5	86.5	81.4	24.6
	222	228	YFAPSSL	783.3803	120	39.8	34.9	31.2	56.0	37.8	31.7	87.0	86.3	37.3
	222	228	YFAPSSL	783.3803	300	47.9	45.3	42.2	59.9	46.8	43.0	88.0	90.3	50.6
	222	228	YFAPSSL	783.3803	1200	59.2	59.5	58.3	67.9	60.0	59.3	89.4	90.8	66.5
	222	228	YFAPSSL	783.3803	3600	66.4	65.8	65.8	72.9	67.7	65.4	88.2	89.6	71.1
	222	228	YFAPSSL	783.3803	10800	71.0	68.3	65.9	75.0	75.4	66.3	87.8	87.0	78.3
	223	228	FAPSSL	620.317	10	38.3	31.0	27.3	53.4	34.5	28.7	88.0	79.9	24.2
	223	228	FAPSSL	620.317	30	42.8	35.8	33.0	59.2	39.2	32.7	86.6	81.6	30.0
	223	228	FAPSSL	620.317	120	53.9	46.9	43.3	68.3	52.6	44.1	87.0	86.1	45.8
	223	228	FAPSSL	620.317	300	61.8	57.4	54.9	69.2	61.6	57.1	88.7	91.4	56.4
	223	228	FAPSSL	620.317	1200	66.1	63.2	64.2	71.2	66.2	64.0	91.0	91.8	67.3
	223	228	FAPSSL	620.317	3600	66.5	64.5	64.4	72.5	68.4	65.6	88.4	89.6	68.9
	223	228	FAPSSL	620.317	10800	69.7	66.8	64.2	74.3	73.4	64.1	87.5	86.2	79.3
	229	232	FDLN	507.2329	10	67.0	68.2	49.0	74.8	72.7	50.3	87.7	88.7	53.1
	229	232	FDLN	507.2329	30	84.6	84.7	61.4	85.5	88.1	65.6	90.8	90.3	67.0
	229	232	FDLN	507.2329	120	90.7	89.6	85.4	91.4	92.8	87.7	91.5	91.2	87.2
	229	232	FDLN	507.2329	300	93.2	93.0	91.2	94.6	96.9	94.2	93.3	94.3	91.0
	229	232	FDLN	507.2329	1200	93.8	93.0	92.5	94.9	96.9	95.3	93.0	95.1	93.5
	229	232	FDLN	507.2329	3600	92.9	92.7	92.7	94.5	94.1	94.2	92.0	92.4	93.4
	229	232	FDLN	507.2329	10800	93.2	92.6	92.3	93.1	95.2	94.2	92.8	91.3	90.9
	232	236	NFQAG	535.2391	10	75.7	77.2	52.9	81.9	78.8	54.8	91.8	90.0	56.5
	232	236	NFQAG	535.2391	30	89.9	90.6	73.9	90.8	90.8	77.3	90.5	91.6	75.8
	232	236	NFQAG	535.2391	120	91.9	91.3	89.9	93.1	93.5	90.2	90.6	92.2	89.0
	232	236	NFQAG	535.2391	300	92.5	93.3	90.5	92.6	93.7	92.4	90.8	92.4	89.5
	232	236	NFQAG	535.2391	1200	92.7	93.8	92.0	92.3	93.6	93.6	92.9	92.1	91.0
	232	236	NFQAG	535.2391	3600	92.0	92.0	91.6	93.0	92.3	92.5	90.9	91.0	91.6
	232	236	NFQAG	535.2391	10800	91.7	92.1	91.6	91.7	93.0	92.4	90.9	89.6	90.1
	232	237	NFQAGF	682.3075	10	53.2	53.8	32.1	66.1	56.2	34.1	88.8	83.9	36.6
	232	237	NFQAGF	682.3075	30	66.4	65.3	51.1	72.7	67.5	54.1	86.7	85.4	54.7
	232	237	NFQAGF	682.3075	120	78.5	77.6	64.1	82.1	83.6	66.3	87.2	87.2	66.3
	232	237	NFQAGF	682.3075	300	88.1	89.1	71.6	89.0	91.2	74.1	88.7	90.4	74.8
	232	237	NFQAGF	682.3075	1200	90.9	91.3	85.9	90.5	91.4	88.9	90.9	89.7	87.8
	232	237	NFQAGF	682.3075	3600	89.4	90.3	89.6	90.5	90.8	90.1	89.6	89.1	89.5
	232	237	NFQAGF	682.3075	10800	89.4	89.9	89.4	89.8	91.0	90.6	89.0	87.2	88.9
	232	238	NFQAGFL	795.3915	10	49.0	47.4	31.8	62.2	51.2	33.3	89.0	86.0	33.3
	232	238	NFQAGFL	795.3915	30	59.2	58.7	45.0	66.7	61.4	46.5	86.8	87.2	46.3
	232	238	NFQAGFL	795.3915	120	68.1	68.1	56.3	73.5	75.0	58.3	87.0	90.0	58.5
	232	238	NFQAGFL	795.3915	300	77.8	79.8	63.2	79.8	86.5	65.4	89.9	93.5	65.4
	232	238	NFQAGFL	795.3915	1200	86.6	90.2	75.4	87.2	93.6	77.5	91.2	93.3	81.3
	232	238	NFQAGFL	795.3915	3600	89.2	92.8	83.2	90.8	93.6	84.5	90.0	92.4	88.2
	232	238	NFQAGFL	795.3915	10800	89.9	92.6	90.7	90.5	93.3	91.3	89.8	90.2	90.1
	233	237	FQAGF	568.2645	10	47.4	46.7	24.0	62.7	49.5	24.7	89.7	82.8	29.6
	233	237	FQAGF	568.2645	30	63.6	61.4	45.9	71.8	64.4	48.1	87.3	85.7	50.1
	233	237	FQAGF	568.2645	120	77.6	77.9	62.8	82.1	82.7	65.1	88.1	87.9	66.7
	233	237	FQAGF	568.2645	300	89.1	90.1	73.3	90.6	90.9	75.3	89.4	91.3	77.3
	233	237	FQAGF	568.2645	1200	91.7	91.3	87.9	92.2	91.5	89.7	92.3	90.8	90.0
	233	237	FQAGF	568.2645	3600	90.0	91.0	91.2	91.3	91.5	91.0	90.2	90.1	90.7
	233	237	FQAGF	568.2645	10800	90.9	91.3	91.8	91.9	91.5	92.5	90.8	89.2	91.4
	233	238	FQAGFL	681.3486	10	33.6	33.4	12.4	50.1	37.8	14.1	86.8	83.3	15.3
	233	238	FQAGFL	681.3486	30	46.2	46.5	31.3	54.6	51.0	32.5	83.7	84.4	32.5
	233	238	FQAGFL	681.3486	120	56.6	58.4	43.7	63.1	66.0	45.0	84.1	88.0	45.6
	233	238	FQAGFL	681.3486	300	69.8	71.7	51.3	71.8	82.5	54.3	87.9	92.6	55.5
	233	238	FQAGFL	681.3486	1200	83.2	87.8	66.4	83.8	93.9	69.1	89.0	93.0	73.7
	233	238	FQAGFL	681.3486	3600	86.6	91.9	77.6	88.7	93.6	79.6	87.7	91.4	85.3
	233	238	FQAGFL	681.3486	10800	88.2	91.6	88.4	88.6	92.9	89.9	87.1	88.3	90.2
	233	240	FQAGFLMK	940.4841	10	28.8	21.5	9.9	48.0	27.3	8.5	87.0	79.9	14.3
	233	240	FQAGFLMK	940.4841	30	38.4	30.9	22.8	50.0	36.8	20.8	84.5	82.2	26.1
	233	240	FQAGFLMK	940.4841	120	47.8	46.1	31.8	58.8	59.2	32.9	86.0	87.4	38.8
	233	240	FQAGFLMK	940.4841	300	64.2	66.2	40.2	68.6	80.7	41.5	87.3	91.1	48.0
	233	240	FQAGFLMK	940.4841	1200	84.7	89.1	61.3	86.0	95.1	63.3	89.4	90.9	69.8
	233	240	FQAGFLMK	940.4841	3600	90.2	93.6	78.2	91.4	95.3	80.8	88.2	90.2	86.3
	233	240	FQAGFLMK	940.4841	10800	90.5	93.6	91.1	92.0	95.1	92.7	87.6	88.5	90.2
	239	244	MKKEVQ	761.4106	10	48.1	43.9	43.5	61.1	47.5	44.6	88.8	81.8	45.3
	239	244	MKKEVQ	761.4106	30	57.4	51.7	48.2	62.7	54.8	50.3	86.3	82.0	49.3
	239	244	MKKEVQ	761.4106	120	72.8	70.5	68.6	77.1	75.7	69.6	87.3	86.2	69.8
	239	244	MKKEVQ	761.4106	300	85.9	86.2	83.8	86.5	88.4	85.0	88.8	89.6	85.3
	239	244	MKKEVQ	761.4106	1200	91.1	89.3	90.2	88.3	89.4	91.7	90.6	89.5	90.4
	239	244	MKKEVQ	761.4106	3600	87.0	89.1	89.4	89.6	89.1	88.9	89.1	88.1	90.3
	239	244	MKKEVQ	761.4106	10800	88.7	89.1	89.2	88.5	87.9	89.2	88.9	86.5	89.7
	239	248	MKKEVQDEEK	1262.6176	10	45.7	39.6	38.3	59.8	46.3	39.9	87.4	80.2	39.9
	239	248	MKKEVQDEEK	1262.6176	30	60.4	56.8	53.8	68.4	62.9	58.1	86.6	82.0	55.0
	239	248	MKKEVQDEEK	1262.6176	120	79.0	77.3	73.7	81.3	81.6	80.3	86.8	85.3	77.7
	239	248	MKKEVQDEEK	1262.6176	300	88.6	88.3	85.9	89.6	91.4	88.1	87.8	89.7	87.9
	239	248	MKKEVQDEEK	1262.6176	1200	91.9	90.1	89.2	91.2	92.5	92.4	90.1	89.6	87.6
	239	248	MKKEVQDEEK	1262.6176	3600	88.3	88.8	88.2	90.1	91.0	88.6	88.5	87.7	88.7
	239	248	MKKEVQDEEK	1262.6176	10800	89.1	88.4	89.3	91.1	90.2	90.6	88.4	85.6	89.2
	242	248	EVQDEEK	875.3872	10	39.5	32.9	29.4	53.4	41.9	31.5	84.9	79.0	32.0
	242	248	EVQDEEK	875.3872	30	62.5	58.1	53.6	66.6	64.4	54.5	85.6	81.2	53.9
	242	248	EVQDEEK	875.3872	120	84.1	80.9	79.4	83.5	85.1	82.1	86.0	85.3	80.5
	242	248	EVQDEEK	875.3872	300	89.6	88.5	87.4	89.0	90.5	87.9	86.7	88.4	86.5
	242	248	EVQDEEK	875.3872	1200	90.2	88.9	89.1	90.4	90.9	90.5	87.5	87.8	86.6
	242	248	EVQDEEK	875.3872	3600	86.5	87.8	87.6	87.2	88.8	84.7	86.6	84.5	84.9
	242	248	EVQDEEK	875.3872	10800	87.4	87.2	87.6	87.0	88.3	87.9	85.4	82.2	85.1

Proteoform	From	To	Sequence	M_{mono}	Time(sec)	corrected %D								
						aWT	hWT	dWT	aS82D	hS82D	dS82D	aP187S	hP187S	dP187S
	242	251	EVQDEEKNKK	1245.6201	10	59.0	55.9	56.5	66.5	59.5	55.0	84.6	80.0	54.8
	242	251	EVQDEEKNKK	1245.6201	30	72.5	69.8	69.4	73.0	71.7	67.4	84.4	81.7	68.5
	242	251	EVQDEEKNKK	1245.6201	120	85.9	82.2	83.2	83.1	82.8	80.6	84.1	83.9	81.9
	242	251	EVQDEEKNKK	1245.6201	300	89.2	88.5	86.9	86.5	88.2	85.7	85.7	86.4	85.4
	242	251	EVQDEEKNKK	1245.6201	1200	89.4	90.1	89.8	87.5	88.4	86.9	87.4	86.1	85.1
	242	251	EVQDEEKNKK	1245.6201	3600	87.8	89.2	87.6	86.0	85.6	83.4	86.0	83.9	85.3
	242	251	EVQDEEKNKK	1245.6201	10800	87.2	87.5	87.5	85.6	85.7	85.1	85.0	81.4	84.7
	243	248	VQDEEK	746.3446	10	38.1	33.2	29.3	54.6	41.5	35.3	79.2	76.9	29.8
	243	248	VQDEEK	746.3446	30	62.5	58.9	55.7	68.7	64.1	62.0	78.3	79.2	54.9
	243	248	VQDEEK	746.3446	120	79.7	78.4	77.9	83.4	80.5	84.5	79.4	82.9	79.3
	243	248	VQDEEK	746.3446	300	86.2	85.7	84.8	89.5	87.2	89.1	83.6	90.9	85.8
	243	248	VQDEEK	746.3446	1200	87.5	85.4	85.7	90.1	87.5	90.2	87.2	86.0	85.1
	243	248	VQDEEK	746.3446	3600	82.5	87.0	84.4	87.5	88.0	83.5	84.4	81.6	83.0
	243	248	VQDEEK	746.3446	10800	83.8	86.7	84.3	87.8	86.4	87.2	84.2	80.4	84.3
	249	254	NKKFGL	705.4174	10	28.9	19.4	18.6	48.8	20.3	18.7	89.3	80.8	20.9
	249	254	NKKFGL	705.4174	30	34.7	27.6	24.6	48.0	26.8	25.1	87.0	81.3	27.2
	249	254	NKKFGL	705.4174	120	42.5	38.5	38.7	54.4	39.3	38.7	87.3	83.8	41.0
	249	254	NKKFGL	705.4174	300	46.1	42.0	40.6	57.6	42.6	41.9	88.8	89.5	46.6
	249	254	NKKFGL	705.4174	1200	51.7	47.3	44.1	60.6	47.7	44.9	90.9	89.7	64.8
	249	254	NKKFGL	705.4174	3600	60.8	59.1	54.4	68.3	59.1	54.6	89.0	88.3	81.0
	249	254	NKKFGL	705.4174	10800	67.6	66.7	64.6	72.6	68.7	64.8	88.1	85.7	88.7
	255	271	SVGHHLGKSIPTDNQIK	1829.9748	10	28.7	21.5	23.0	48.3	23.1	22.2	91.6	94.0	60.4
	255	271	SVGHHLGKSIPTDNQIK	1829.9748	30	32.3	25.5	27.6	47.7	26.7	26.6	88.8	98.0	61.9
	255	271	SVGHHLGKSIPTDNQIK	1829.9748	120	39.7	32.2	37.6	55.2	36.8	39.3	89.4	99.2	65.6
	255	271	SVGHHLGKSIPTDNQIK	1829.9748	300	51.0	42.8	49.7	61.2	49.0	52.3	92.4	100.3	71.4
	255	271	SVGHHLGKSIPTDNQIK	1829.9748	1200	65.0	61.0	66.1	72.8	66.0	67.9	94.5	101.3	85.8
	255	271	SVGHHLGKSIPTDNQIK	1829.9748	3600	68.7	69.0	71.0	76.2	71.4	72.2	92.4	99.5	95.2
	255	271	SVGHHLGKSIPTDNQIK	1829.9748	10800	73.1	73.8	74.1	79.8	76.4	74.2	93.2	96.4	100.6
	256	271	VGHHLGKSIPTDNQIK	1742.9428	10	29.6	20.7	21.9	48.0	21.4	22.2	89.3	91.1	63.3
	256	271	VGHHLGKSIPTDNQIK	1742.9428	30	33.3	23.5	25.9	47.4	25.3	25.8	85.8	95.8	65.2
	256	271	VGHHLGKSIPTDNQIK	1742.9428	120	40.6	31.8	39.2	56.1	35.7	40.8	86.6	95.5	69.1
	256	271	VGHHLGKSIPTDNQIK	1742.9428	300	53.1	44.0	51.8	63.1	49.8	54.1	89.5	96.8	73.6
	256	271	VGHHLGKSIPTDNQIK	1742.9428	1200	67.8	63.4	70.1	73.8	68.4	71.0	91.9	97.5	84.8
	256	271	VGHHLGKSIPTDNQIK	1742.9428	3600	71.5	71.6	74.6	77.5	73.4	75.0	90.1	96.0	93.4
	256	271	VGHHLGKSIPTDNQIK	1742.9428	10800	75.4	76.8	78.2	81.3	77.6	77.2	90.8	92.9	98.5
	261	271	GKSIPTDNQIK	1199.651	10	29.7	21.5	21.9	47.9	22.8	22.8	88.2	84.8	66.6
	261	271	GKSIPTDNQIK	1199.651	30	34.5	24.1	27.2	48.4	26.0	27.8	86.8	84.6	68.6
	261	271	GKSIPTDNQIK	1199.651	120	44.0	32.4	40.1	59.0	37.3	41.4	87.1	85.9	70.6
	261	271	GKSIPTDNQIK	1199.651	300	58.1	46.8	55.2	67.2	53.8	55.8	88.3	89.2	74.1
	261	271	GKSIPTDNQIK	1199.651	1200	75.5	67.7	73.3	79.7	72.9	73.9	89.4	89.2	79.4
	261	271	GKSIPTDNQIK	1199.651	3600	77.9	75.4	76.7	80.5	76.9	75.0	88.2	87.3	86.6
	261	271	GKSIPTDNQIK	1199.651	10800	79.9	77.9	77.7	81.7	78.6	77.0	87.5	85.5	88.8
	262	271	KSIPTDNQIK	1142.6295	10	31.8	22.6	23.7	50.0	23.3	24.1	87.3	84.9	72.4
	262	271	KSIPTDNQIK	1142.6295	30	36.8	24.7	28.8	51.1	26.8	26.4	85.8	84.2	74.9
	262	271	KSIPTDNQIK	1142.6295	120	46.1	32.4	40.9	60.1	38.5	41.6	86.4	85.7	75.4
	262	271	KSIPTDNQIK	1142.6295	300	57.6	46.2	56.0	67.6	53.8	56.1	87.1	88.4	76.5
	262	271	KSIPTDNQIK	1142.6295	1200	73.5	64.7	72.3	77.6	70.2	72.3	88.7	87.3	78.5
	262	271	KSIPTDNQIK	1142.6295	3600	76.2	73.2	75.0	78.8	74.2	73.1	88.3	87.5	82.2
	262	271	KSIPTDNQIK	1142.6295	10800	78.1	75.6	75.8	80.7	76.1	75.5	87.3	85.1	87.0
	263	271	SIPTDNQIK	1014.5346	10	35.3	28.2	28.4	51.5	28.8	28.3	87.4	87.3	84.5
	263	271	SIPTDNQIK	1014.5346	30	40.4	31.1	33.0	53.2	32.3	33.6	86.1	87.3	87.0
	263	271	SIPTDNQIK	1014.5346	120	51.9	39.6	48.6	64.6	46.1	49.7	86.3	86.8	87.3
	263	271	SIPTDNQIK	1014.5346	300	67.3	55.1	65.4	76.2	64.5	67.1	88.0	89.6	87.9
	263	271	SIPTDNQIK	1014.5346	1200	84.1	76.9	83.9	87.1	83.5	85.1	89.9	88.5	87.8
	263	271	SIPTDNQIK	1014.5346	3600	85.9	85.2	87.2	87.2	86.5	85.9	88.5	87.9	88.9
	263	271	SIPTDNQIK	1014.5346	10800	87.2	86.8	87.3	87.2	87.1	86.7	87.6	85.8	90.6
	264	271	IPTDNQIK	927.5025	10	35.8	26.2	26.3	52.4	26.6	26.7	87.4	87.6	86.8
	264	271	IPTDNQIK	927.5025	30	41.1	27.6	32.9	54.4	29.9	33.0	85.2	87.4	88.4
	264	271	IPTDNQIK	927.5025	120	53.5	37.7	49.3	66.2	45.7	49.4	85.9	86.3	88.9
	264	271	IPTDNQIK	927.5025	300	69.2	55.2	67.8	77.8	65.4	68.3	87.4	88.6	89.8
	264	271	IPTDNQIK	927.5025	1200	85.8	78.4	85.1	88.0	84.3	86.3	89.9	87.8	88.8
	264	271	IPTDNQIK	927.5025	3600	87.1	87.4	89.5	87.6	87.5	87.2	87.0	86.1	90.2
	264	271	IPTDNQIK	927.5025	10800	88.0	88.5	88.9	87.5	88.1	88.3	87.8	86.9	88.2

Table S5. Parameters for the fittings to hyperbolic functions for the reduction of FAD by NADH. Experiments were carried out at 6°C. Data for WT are from [7].

	Fast FAD reduction			Slow FAD reduction		
	(A→B)			(B→C)		
	k_{HT1}	K_{d1}^{NADH}	k_{HT1}/K_{d1}^{NADH}	k_{HT2}	K_{d2}^{NADH}	k_{HT2}/K_{d2}^{NADH}
	(s^{-1})	(μM)	($\mu M^{-1} \cdot s^{-1}$)	(s^{-1})	(μM)	($\mu M^{-1} \cdot s^{-1}$)
WT	281±14	15±2	19±3	14±2	8.2±3.6	1.9±1.1
S82D	368±33	23±5	16±5	9.8±0.9	43±8	0.23±0.06
P187S	480±79	59±18	8.1±3.8	1.8±0.5	75±34	0.024±0.06

Table S6. Observed rate constants for the anaerobic reoxidation of NQO1$_{hq}$ variants (i.e. NQO1$_{holo}$ reduced with NADH) by equimolecular concentrations of DCPIP. Data are mean ± SD from n > 3. Experiments carried out at 6ºC. Data for WT are from [7].

Variant	NQO1$_{hq}$ reoxidation by DCPIP	
	k_{obs1} (s^{-1})	k_{obs2} (s^{-1})
WT	> 500	160 ± 14
S82D	277 ± 18	72 ± 9
P187S	226 ± 14	32 ± 4

Table S7. KIEs and Arrhenius parameters for the HT in the reduction of NQO1 variants by NADH. All values correspond to data obtained with equimolecular concentrations of the reactants in the stopped-flow equipment. (n > 3, mean ± SD). Analysis were performed as previously described [7].

Parameters for the fast FAD reduction

	By NADH			By NADD			KIE	ΔE_{aD1-H1} (kcal·mol^{-1})	A_{H1}/A_{D1}
	k^a_{obs1} (s^{-1})	E_{aH1} (kcal·mol^{-1})	A_{H1} (s^{-1})	k^a_{obs1} (s^{-1})	E_{aD1} (kcal·mol^{-1})	A_{D1} (s^{-1})			
WT [b]	78±1	6.1±0.2	(5.3±1.2)·10^6	44±2	6.3 ± 0.4	(4.1±1.1)·10^6	1.8 ± 0.1	0.2±0.4	1.3±0.6
S82D	68±3	4.9±0.1	(4.9±1.0)·10^5	41±1	5.4±0.2	(6.4±3.3)·10^5	1.7±0.1	0.5±0.2	0.8±0.4
P187S	90±1	4.4±0.1	(2.7±0.8)·10^5	46±1	4.2±0.4	(9.6±9.0)·10^4	2.0±0.1	-0.2±0.4	2.8±1.6

Parameters for the slow FAD reduction

	By NADH			By NADD			KIE	ΔE_{aD2-H2} (kcal·mol^{-1})	A_{H2}/A_{D2}
	k^a_{obs2} (s^{-1})	E_{aH2} (kcal·mol^{-1})	A_{H2} (s^{-1})	k^a_{obs2} (s^{-1})	E_{aD2} (kcal·mol^{-1})	A_{D2} (s^{-1})			
WT [b]	8.9±0.9	10.9±0.5	(3.4±0.9)·10^9	5.3±0.2	9.8±0.5	(2.6±0.6)·10^8	1.7±0.2	-1.1±0.7	13±6
S82D	2.0±0.8	13.0±0.1	(3.0±0.8)·10^{10}	1.4±0.3	17.7±0.3	(1.0±0.5)·10^{14}	1.4±0.6	4.7±0.3	(3±2)·10^{-4}
P187S	0.13±0.01	4.8±0.2	(7.6±3.1)·10^2	0.13±0.01	4.7±0.2	(6.3±2.0)·10^2	1.0±0.1	-0.1±0.3	1.2±0.6

[a]Values at 6 °C. [b]Data from [7].

Table S8. Thermodynamic stability of the CTD inferred from proteolysis kinetics.

	NQO1$_{apo}$		NQO1$_{holo}$		NQO1$_{dic}$	
	k_{prot}* (M^{-1}·min^{-1})	$\Delta\Delta G_{prot}$** (kcal·mol^{-1})	k_{prot} (M^{-1}·min^{-1})	$\Delta\Delta G_{prot}$ (kcal·mol^{-1})	k_{prot} (M^{-1}·min^{-1})	$\Delta\Delta G_{prot}$ (kcal·mol^{-1})
WT	4.07±0.24 ·10^6	0**	3.28±0.10 ·10^4	2.86±0.04	1.84±0.19 ·10^4	3.20±0.07
S82D	2.69±0.20 ·10^6	0.25±0.06	5.30±0.27 ·10^5	1.21±0.05	5.33±0.37 ·10^4	2.57±0.03
P187S	1.05±0.12 ·10^7	-0.56±0.08	1.20±0.24 ·10^7	-0.64±0.12	4.53±0.49 ·10^5	1.30±0.07

* Second-order rate constants for proteolysis with trypsin.

** Change in unfolding free energy using WT-NQO1$_{apo}$ as reference.

Table S9. Apparent binding thermodynamic parameters for the interaction between NQO1 and dicoumarol.

	N*	K_d** (nM)	ΔG** (kcal·mol^{-1})	ΔH** (kcal·mol^{-1})	-TΔS** (kcal·mol^{-1})	ΔC_p*** (kcal·mol^{-1}·K^{-1})
WT	1.05±0.04	18±11	-10.5±0.3	-8.7±0.3	-1.8±0.4	-0.37±0.04
S82D	1.14±0.06	91±36	-9.6±0.2	-12.0±0.8	2.4±0.9	-0.59±0.07
P187S	0.90±0.09	560±270	-8.4±0.3	-31.7±4.8	23.3±4.7	-2.02±0.23

* Mean±s.d. from at least five experiments at different temperatures.

** Mean±s.d. from at least three experiments at 25ºC.

*** Slope from a linear fitting of the binding ΔH vs. Temperature.

References

[1] J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau, L. Serrano, The FoldX web server: an online force field, Nucleic Acids Res, 33 (2005) W382-388.

[2] R. Fraczkiewicz, W. Braun, Exact and Efficient Analytical Calculation of the Accessible Surface Areas and Their Gradients for Macromolecules, J.Comp.Chem., 19 (1998) 319-333.

[3] H. Wickham, ggplot2: Elegant Graphics for Data Analysis. , Springer-Verlag, 2016.

[4] A. Martinez-Limon, M. Alriquet, W.H. Lang, G. Calloni, I. Wittig, R.M. Vabulas, Recognition of enzymes lacking bound cofactor by protein quality control, Proc Natl Acad Sci U S A, 113 (2016) 12156-12161.

[5] N. Tokuriki, F. Stricher, J. Schymkowitz, L. Serrano, D.S. Tawfik, The stability effects of protein mutations appear to be universally distributed, J Mol Biol, 369 (2007) 1318-1332.

[6] P. Vankova, E. Salido, D.J. Timson, P. Man, A.L. Pey, A dynamic core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer, Biomolecules, 9 (2019) 728.

[7] E. Anoz-Carbonell, D.J. Timson, A.L. Pey, M. Medina, The Catalytic Cycle of the Antioxidant and Cancer-Associated Human NQO1 Enzyme: Hydride Transfer, Conformational Dynamics and Functional Cooperativity, Antioxidants (Basel), 9 (2020).

Different phenotypic outcome due to site-specific phosphorylation in the cancer-associated NQO1 enzyme studied by phosphomimetic mutations.

Check for updates
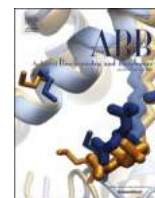
Different phenotypic outcome due to site-specific phosphorylation in the cancer-associated NQO1 enzyme studied by phosphomimetic mutations

Juan Luis Pacheco-Garcia [a], Ernesto Anoz-Carbonell [b], Dmitry S. Loginov [c], Pavla Vankova [d], Eduardo Salido [e], Petr Man [c], Milagros Medina [b], Rogelio Palomino-Morales [f], Angel L. Pey [g],[*]

[a] *Departamento de Química Física, Universidad de Granada, Av. Fuentenueva s/n, 18071, Granada, Spain*
[b] *Departamento de Bioquímica y Biología Molecular y Cellular, Facultad de Ciencias, Instituto de Biocomputación y Física de Sistemas Complejos (BIFI) (GBsC-CSIC Joint Unit), Universidad de Zaragoza, 50009, Zaragoza, Spain*
[c] *Institute of Microbiology - BioCeV, Academy of Sciences of the Czech Republic, Prumyslova 595, Vestec, 252 50, Czech Republic*
[d] *Institute of Biotechnology - BioCeV, Academy of Sciences of the Czech Republic, Prumyslova 595, Vestec, 252 50, Czech Republic*
[e] *Center for Rare Diseases (CIBERER), Hospital Universitario de Canarias, Universidad de la laguna, 38320, Tenerife, Spain*
[f] *Department of Biochemistry and Molecular Biology I, Faculty of Sciences and Biomedical Research Center (CIBM), University of Granada, Granada, Spain*
[g] *Departamento de Química Física, Unidad de Excelencia en Química Aplicada a Biomedicina y Medioambiente e Instituto de Biotecnología, Universidad de Granada, Av. Fuentenueva s/n, 18071, Granada, Spain*

ABSTRACT

Protein phosphorylation is a common phenomenon in human flavoproteins although the functional consequences of this site-specific modification are largely unknown. Here, we evaluated the effects of site-specific phosphorylation (using phosphomimetic mutations at sites S40, S82 and T128) on multiple functional aspects as well as in the structural stability of the antioxidant and disease-associated human flavoprotein NQO1 using biophysical and biochemical methods. *In vitro* biophysical studies revealed effects of phosphorylation at different sites such as decreased binding affinity for FAD and structural stability of its binding site (S82), conformational stability (S40 and S82) and reduced catalytic efficiency and functional cooperativity (T128). Local stability measurements by H/D exchange in different ligation states provided structural insight into these effects. Transfection of eukaryotic cells showed that phosphorylation at sites S40 and S82 may reduce steady-levels of NQO1 protein by enhanced proteasome-induced degradation. We show that site-specific phosphorylation of human NQO1 may cause pleiotropic and counterintuitive effects on this multifunctional protein with potential implications for its relationships with human disease. Our approach allows to establish relationships between site-specific phosphorylation, functional and structural stability effects *in vitro* and inside cells paving the way for more detailed analyses of phosphorylation at the flavoproteome scale.

1. Introduction

Human NAD(P)H:quinone oxidoreductase 1 (NQO1) is a multifunctional protein whose dysregulation has been associated with developing cancer, Alzheimer's and Parkinson's disease [1–3]. It catalyzes the two-electron reduction of a wide variety of quinones, including maintenance of vitamins in their reduced state, superoxide scavenging, detoxification of xenobiotics and activation of cancer pro-drugs [1,3–5]. Additionally, NQO1 interacts with other proteins (PPI, e.g. p53 and HIF-1) and with RNA molecules [1,6–9].

NQO1 forms dimers with two active sites, each of them containing a tightly bound FAD molecule required for catalysis and also providing

intracellular stability [10–13]. The monomer contains two domains: an N-terminal domain (NTD; residues 1–225) containing the FAD binding site and most of the monomer:monomer interface (MMI), whereas the C-terminal domain (CTD, residues 226–274) completes the active sites [11–15].

NQO1 displays a remarkable functional chemistry. During its catalytic cycle, two half-reactions occur: i) in the reductive half-reaction the FAD molecules are reduced upon NAD(P)H binding and hydride transfer (HT), which constitutes the rate-limiting step of the catalytic cycle; ii) in the oxidative half-reaction, upon release of NAD(P)$^+$, the substrate binds and is reduced by FADH$_2$ reconstituting the holo-enzyme [15,16]. Enzyme kinetic analyses have supported the existence of non-equivalent active sites that yield negative cooperativity to this catalytic cycle

---

* Corresponding author.
  E-mail address: angelpey@ugr.es (A.L. Pey).

[16–18].

In addition, ligands bound to NQO1 modulate its multiple functions. In the FAD-free state (NQO1$_{apo}$), the enzyme is inactive, intracellularly unstable and displays altered PPI [10,19]. FAD binding renders a conformational state (NQO1$_{holo}$), ready to initiate catalysis, intracellularly stable and with a different interactomic pattern than that of NQO1$_{apo}$ [10,19]. Upon NAD(P)H binding, the FAD is reduced to FADH$_2$ constituting a catalytically relevant state that also shuttles the proteins to interact with microtubules and enhances PPI [6,16,20]. Binding of dicoumarol (Dic), a competitive inhibitor of both NAD(P)H and the substrate [16], leads to an inactive state likely reflecting the ternary complex (NQO1holo:NAD(P)H) or transition state analogue, also abrogating PPI [6]. It is interesting to note that we have recently characterized the effects of FAD and Dic binding to WT NQO1, revealing that ligand binding sequentially stabilizes the FAD (FBS) and Dic (DBS) binding sites, and these stabilizing effects propagate far beyond these sites affecting the entire structure of the enzyme [21].

Phosphorylation is a common phenomenon in human flavoproteins, although the functional consequences of these events are largely unknown (https://www.phosphosite.org/). By July 30, 2022, there were 12 phosphorylation events identified in human NQO1 based on high-throughput approaches (Fig. 1A). Most of them are localized in the NTD (Fig. 1A). Three of these sites (S40, S82 and T128) have been characterized in some detail at the molecular level by phosphomimetic mutations or by direct phosphorylation [2,18,22,23](Fig. 1B). The phosphomimetic mutation S82D affects a partially solvent exposed residue far from the FBS and the DBS (at minimal distances of 14 Å and 18.5 Å, respectively)(Fig. 1B). S82D causes local destabilization of the structure, and this effect propagates to the FBS and reduces by ~10-fold the affinity for the cofactor and thus its intracellular stability likely due to increased population of the NQO1$_{apo}$ state [22]. In addition, S82D decreases the catalytic efficiency for the slow FAD reduction pathway, thus exacerbating the functional negative cooperativity displayed by WT NQO1 [18]. The kinase responsible for phosphorylation of S82 is unknown. S40 is fully buried in the crystal structure and also locates far from the FBS and DBS (minimal distances of 14.5 Å and 20 Å, respectively). T128 is a solvent exposed residue that locates closer to the DBS (minimal distance of 6.8 Å) and at 12 Å from the FBS (Fig. 1B). Importantly, it is next to Y129, that undergoes important conformational rearrangements upon Dic binding [24]. S40 and T128 sites are both phosphorylated by the Akt kinase, although phosphorylation of the T128 site seems to be important to trigger phosphorylation at both sites [2]. Phosphorylation at S40 and T128 may facilitate ubiquitination and degradation of NQO1, with a more important role of phosphorylation at T128 [2]. This decrease in intracellular stability due to phosphorylation at S40/T128 may lower the antioxidant and cytoprotective capacity of NQO1 in cultured cells [2]. This study has also associated phosphorylation of T128 with increased propensity to develop Parkinson's disease [2].

Overall, these studies suggest that phosphorylation of NQO1 at different sites (S40, S82 and T128) may impair the stability of the enzyme inside cells. However, the structural and functional consequences of these phosphorylation events individually on the multifunctional NQO1 are not well known. In this work, we have carried out a comparative and multidisciplinary structure-function analysis of these phosphorylation events using *phosphomimetic* mutations (S40D, S82D ad T128D) and *unphosphorylatable* variants (S40A, S82A and T128A).

2. Materials and methods

2.1. Protein expression and purification

Mutations were introduced by site-directed mutagenesis on the wild-type (WT) NQO1 cDNA cloned into the pET-15b vector (pET-15b-NQO1) by GenScript (Leiden, The Netherlands). Codons were optimized for expression in *E. coli* and mutagenesis was confirmed by sequencing.

Typically, 40 mL of Luria-Bertani (LB) medium containing 0.1 mg mL^{-1} ampicillin (Canvax Biotech, Córdoba, Spain) were inoculated with transformed cells and grown for 16 h at 37 °C. These cultures were diluted into 800 mL of LB containing 0.1 mg mL^{-1} ampicillin, grown at
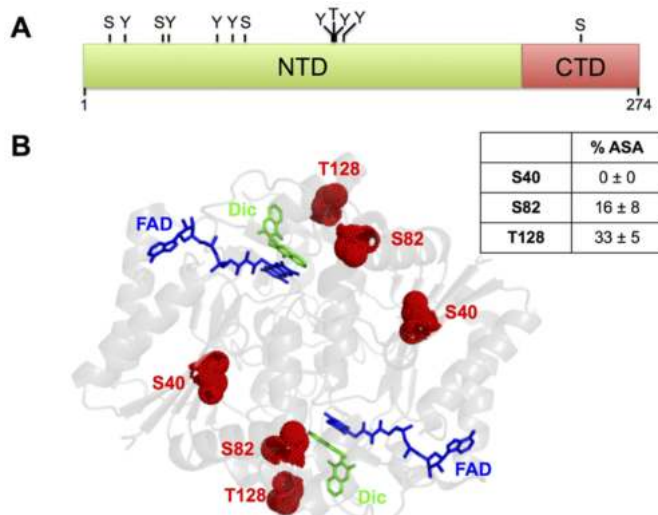
Fig. 1. Phosphorylation sites in human NQO1. A) Schematic representation of the different phosphorylation sites in the primary sequence of NQO1 (as described in Phosphosite Plus®; https://www.phosphosite.org/); B) Structural location of the phosphorylation sites at S40, S82 and T128 investigated in this work (PDB code 2F1O) [24]. The percentage of accessible surface area (% ASA) was determined usin GetArea (http://curie.utmb.edu/getarea.html) [34] and it is reported as the average ± s.d. from eight monomers found in this crystallographic structure.

37 °C for 3 h to reach an optical density of about 1.0 and then these were transferred to 25 °C and induced with 0.5 mM IPTG (Isopropyl β-*d*-1-thiogalactopyranoside, Canvax Biotech). After 6 h, cells were harvested by centrifugation and frozen at −80 °C for 16 h. Cells were resuspended in binding buffer, BB (20 mM Na-phosphate, 300 mM NaCl, 50 mM imidazole, pH 7.4) with 1 mM PMSF (phenylmethylsulfonyl fluoride, Sigma-Aldrich, Madrid, Spain), sonicated in an ice bath, and the extracts were centrifuged (21000 g, 30 min, 4 °C). Supernatants were loaded into IMAC (immobilized-metal affinity chromatography) columns (Ni-Sepharose, Cytiva, Barcelona, Spain), washed with 30 vol of BB and eluted in this buffer containing 500 mM imidazole. These eluates were buffer exchanged using PD-10 columns (Cytiva, Barcelona, Spain) to 50 mM HEPES-KOH, pH 7.4 and stored at −80 °C. 5–10 mg of protein from IMAC were thawed, centrifuged for 10 min at 21000 g and 4 °C and loaded into a HiLoad® 16/60 Superdex® 200 pg (Cytiva, Barcelona, Spain). Size-exclusion chromatography (SEC) was carried out using 20 mM HEPES-NaOH, 200 mM NaCl, pH 7.4 at 20 °C and at a 1.5 mL min^{-1} flow rate as described [16]. Fractions containing dimeric NQO1 were concentrated and exchanged to HEPES-KOH 50 mM, pH 7.4 using PD-10 columns. After centrifugation for 20 min at 21000 g at 4 °C, the amount of protein (in NQO1 subunit) and the FAD content was determined from the UV–visible spectrum using: $\varepsilon_{(280)} = 47900$ M^{-1} cm^{-1} for NQO1 and $\varepsilon_{(450)} = 11300$ M^{-1} cm^{-1} for FAD [25]. Spectra were collected in a Cary 50 spectrophotometer (Agilent Technologies, Waldbronn, Germany) using 0.3 cm path-length quartz cuvettes and ~30 μM of NQO1 in protein subunit. Each NQO1 variant was expressed and analyzed at least three times to check for reproducibility. The purity of NQO1 proteins was evaluated by SDS-PAGE in 12% acrylamide gels (Fig. S1). To obtain apo-proteins, FAD was removed by treatment with 2 M KBr and 2 M urea, 1 mM β-mercapto-ethanol in BB and 1 mM PMSF as previously described [14,25]. Apo-proteins were buffer-exchanged to 50 mM HEPES-KOH, pH 7.4 and contained <3% of FAD bound based on UV–visible spectra. All purified proteins were stored at −80 °C upon flash-freezing in N$_2$. For enzyme kinetic studies, purified proteins were mixed with a 1 mM FAD solution in 50 mM HEPES-KOH, pH 7.4 for 10 min at room temperature, and the excess of FAD was removed using PD-10 columns. In all cases, this procedure yielded NQO1 protein with more than 95% FAD bound per NQO1 monomer based on UV–visible spectroscopy.

2.2. Titrations with FAD

Fluorescence titrations were carried out at 25 °C using 1 × 0.3 cm path-length cuvettes in a Cary Eclipse spectrofluorimeter (Agilent Technologies). Experiments were carried out in 20 mM sodium phosphate, pH 7.4. 20 μL of a 12.5 μM NQO1 stock solution (in NQO1 subunit) was mixed with 0–500 μL of FAD 10 μM and completed with buffer up to a 1 mL volume. Samples were incubated at 25 °C in the darkness for at least 10 min before measurements. Fluorescence spectra were acquired in the 340–360 nm range upon excitation at 280 nm (slits 5 nm) and spectra averaged over 10 scans using a scan rate of 200 nm min^{-1}. Blanks were also measured similarly (containing only buffer) and subtracted. The intensity of the fluorescence at 350 nm (I) vs. total concentration of FAD ([FAD]) was used to determine the apparent dissociation constant $K_{d(FAD)}$ using equation (1):

where I$_{holo}$ and I$_{apo}$ are the fluorescence intensities of NQO1$_{holo}$ and NQO1$_{apo}$, respectively, and [NQO1] is the total protein concentration (250 nM). This model assumes that FAD binding sites in the NQO1 dimer are equivalent. Data from at least two independent titrations using different preparations of apo-proteins for each NQO1 were used for fittings. The apparent change in binding free energy, $\Delta G_{(FAD)}$, was determined using the following equation: $\Delta G_{(FAD)} = R \cdot T \cdot \ln K_{d(FAD)}$, where R is the ideal gas constant (1.987 cal mol^{-1} K^{-1}) and T is the absolute temperature (298.15 K) The associated error was determined by linear propagation from the errors of $K_{d(FAD)}$.

2.3. Partial proteolysis with thermolysin

Thermolysin from *Geobacillus stearothermophilus* (Sigma-Aldrich) was prepared in 50 mM HEPES-KOH, pH 7.4 with 100 mM CaCl$_2$. The concentration of thermolysin stock solutions was determined spectrophotometrically using $\varepsilon_{280} = 66086$ M^{-1} cm^{-1}. For proteolysis experiments, purified NQO1 proteins were prepared at ~15 μM NQO1 with 100 μM FAD or 100 μM FAD +100 μM Dic in 50 mM HEPES-KOH, pH 7.4 (all concentrations were final concentrations). Solutions were incubated for 5 min at 25 °C, and reactions were initiated by addition of thermolysin (to 0.1–0.5 μM, final concentration of 10 mM CaCl$_2$). Aliquots were withdrawn at different times, mixed with 25 mM EDTA (ethylenediaminetetraacetic acid), pH 8.0 as final concentration, Laemmli's buffer was added in a 1:1 vol ratio and final solutions were denatured at 95 °C for 5 min. Controls without thermolysin were prepared similarly as samples for time zero. Samples were resolved in SDS-PAGE gels (12% in acrylamide), scanned and analyzed by densitometry using ImageJ (http://rsb.info.nih. gov/ij/). Time-dependent degradation of the full-length protein was fitted using a single exponential function to yield the first-order kinetic constant k_{obs}. The second-order rate constant k_{prot} was determined from k_{obs}/[Thermolysin]. Changes in the local stability (of native vs. cleavable state) upon mutation $\Delta\Delta G_{prot(mut-WT)}$ were determined by equation (2) [26]:

$$\Delta\Delta G_{prot(WT-mut)} = R \cdot T \cdot ln \frac{k_{prot(mut)}}{k_{prot(WT)}} \qquad \text{Equation 2}$$

where R is the ideal gas constant, T is the absolute temperature (298.15 K) and $k_{prot(mut)}$ and $k_{prot(WT)}$ are the second-order rate constants for the mutant and the WT protein, respectively. Errors in $\Delta\Delta G_{prot(WT-mut)}$ were determined by linear propagation of the fitting errors from the k_{prot} values.

2.4. Thermal denaturation of NQO1 enzymes

Purified NQO1 proteins were prepared at a 2 μM final concentration in subunit in the presence 20 μM of FAD in 50 mM HEPES-KOH, pH 7.4. Samples were loaded into 3 × 3 mm path length quartz cuvettes. Thermal denaturation was carried out in a Cary Eclipse (Varian) spectrofluorimeter equipped with a Peltier element. Samples were equilibrated at 20 °C for 3 min and temperature was increased up to 70 °C at a 2 °C min^{-1} scan rate. Fluorescence emission was recorded at 350 nm (slit 10 nm) upon excitation at 280 nm (slit 5 nm). Experimental curves were normalized using pre- and post-transition linear baselines to provide the apparent half-denaturation temperatures (T$_m$). For each variant, 3–6 replicas were prepared and analyzed and T_m values were presented as

$$I = I_{apo} + \left(I_{holo} - I_{apo}\right) \cdot \left(\frac{[NQO1] + [FAD] + K_{d(FAD)} - \sqrt{([NQO1] + [FAD] + K_{d(FAD)})^2 - 4 \cdot [NQO1] \cdot [FAD]}}{2 \cdot [NQO1]} \right) \qquad \text{Equation 1}$$

2.5. Hydrogen-deuterium exchange (HDX) monitored by mass spectrometry

Backbone amide hydrogen/deuterium exchange (HDX) was monitored using mass spectrometry as NQO1$_{apo}$ and NQO1$_{holo}$ forms for WT, S40D, S82D and T128D NQO1 variants, and also as NQO1$_{dic}$ for WT and T128D. All procedures including sample preparation, sample measurement and data processing were performed almost identically as described previously [21]. Initial sample concentration was 20 μM and in case of holo-proteins, FAD was included at 200 μM concentration. Dic state was prepared from the holo-form by adding dicoumarol to final concentration of 200 μM. After a 10 min incubation at 25 °C, H/D exchange reaction was initiated by a 10-fold dilution of the protein solution into D$_2$O-based 50 mM HEPES-KOH, pD 7.4, 1 mM TCEP (tris (2-carboxyethyl)phosphine). Exchange was thus performed at 2 μM protein concentration. Aliquots were taken at 10 s, 30 s, 2 min, 5 min, 20 min, 1 h and 3 h and time points 10 s, 5 m, 3 h were replicated. The reaction was quenched by the addition of 0.5 M Glycine-HCl, pH 2.3 in a 1:1 ratio and flash frozen in liquid N$_2$. LC-MS analysis was started by rapid thawing of the sample and injection onto the custom-made coimmobilized nepenthesin-2 (Nep2) and pepsin column. Generated peptides were trapped and desalted on a SecurityGuard™ pre-column (ULTRA Cartridges UHPLC Fully Porous Polar C18, 2.1 mm, Phenomenex, Torrance, CA, USA). The digestion and desalting were driven by 0.4% formic acid (FA) in water pumped by 1260 Infinity II Quaternary pump (Agilent Technologies, Waldbronn, Germany) at a flow rate of 200 μL min^{-1}. Desalted peptides were then separated on an analytical column (LUNA® Omega Polar C18 Column, 100 Å, 1.6 μm, 100 mm × 1.0 mm, Phenomenex, Torrance, CA, USA) using a linear gradient (10–45% B in 7 min) followed by a quick step to 99% B lasting 5 min, where the solvent A was 0.1% FA/2% acetonitrile (ACN) in water, B was 0.1% FA/98% ACN in water. The gradient was delivered by the 1290 Infinity II LC System (Agilent Technologies, Waldbronn, Germany) at a flow rate 40 μL min^{-1}. The analytical column was coupled to an ESI source of 15T FT-ICR mass spectrometer (Solarix XR, Bruker Daltonics, Bremen, Germany) operating in broad-band MS mode. Acquired data were exported using DataAnalysis v. 5.3 (Bruker Daltonics, Bremen, Germany), processed by in-house designed software called DeutEx and handled as described previously [21]. Peptide identification was performed using separate LC-MS/MS analyses employing the same UPLC system and settings but coupled to a timsTOF Pro PASEF instrument (Bruker Daltonics, Bremen, Germany). Data were searched by MASCOT (version 2.4, Matrix Science, London, United Kingdom) against a custom-built database containing sequences of the proteases, NQO1 variants and common contaminations as listed in the cRAP databse. Decoy search was enabled, false-discovery ratio was set to 1% and ion score cut-off to 20. Fully deuterated proteins were prepared as described previously [21] and correction for back-exchange was applied according to Ref. [27]. HDX results are presented as Δ%D$_{av}$, a parameter that compares a given segment between two variants in a given ligation state (i.e. S40D$_{apo}$ vs. WT$_{apo}$) as the average of the three time points that show maximal difference in percentage of incorporated deuterium (% D) along a time course [21].

The HDX-MS data have been deposited to the ProteomeXchange Consortium via the PRIDE [28] partner repository with the dataset identifier PXD034600.

2.6. Pre-steady state enzyme kinetic analysis

Hydride-(HT) or deuteride-(DT) transfer reactions from NADH/D to NQO1 were carried out under anaerobic conditions using a stopped-flow spectrophotometer (SX.18 MV, Applied Photophysics Ltd., Leatherhead, UK) interfaced with a photodiode array detector, essentially as described [16]. Reactions were performed in 20 mM HEPES-KOH, pH 7.4. The reductive half-reaction was measured by mixing NQO1$_{holo}$ variants (7.5 μM) with NADH ranging from 7.5 to 100 μM. Multiple wavelength absorption data in the flavin absorption region were collected and processed as described [16]. Time-dependent spectral deconvolution was performed by global analysis and numerical integration methods using previously described procedures [16]. Deconvolution was carried out considering sequential and irreversible steps in the context of two (A→B→C) or three (A→B→C→D) step mechanisms, were A-D are spectral species, and allowed to determine observed rate constants (k_{obs}) for these steps as well as spectroscopic properties of intermediate A to D species. According to a recent study, catalytically relevant processes involved steps A→B (*Fast*) and B→C (*Slow*) [16]. Hyperbolic dependences of k_{obs} vs. NADH concentrations were fitted using equation (3):

$$k_{obs} = \frac{k_{HT} \cdot [NADH]}{K_{d(NADH)} + [NADH]} \qquad \text{Equation 3}$$

where k_{HT} is the limiting rate constant for HT and $K_{d(NADH)}$ is the equilibrium dissociation constant to a given active site.

Primary KIEs in the HT process were estimated as previously reported [16] in samples containing equimolar mixtures (7.5 μM of each component) by evaluating HT or DT k_{obs} (k_{obsHT} or k_{obsDT}) from NADH and [4R-^{2}H]-NADD to NQO1$_{ox}$ at different temperatures in the 6–20 °C range. Estimation of the activation parameters (frequency factor, A, and the activation energy, E$_a$) were determined using the Arrhenius equation as described [16].

2.7. Expression analyses in eukaryotic cells

Site-directed mutagenesis to S40A, S40D, S82A, S82D, T128A or T128D was carried out on the wild-type (WT) NQO1 cDNA cloned into the pCI-NEO plasmid by GenScript (Leiden, The Netherlands). Mutagenesis was confirmed by sequencing the entire cDNA.

HAP1 NQO1 knockout cells (HAP1 NQO1-KO; Horizon, Waterbeach, UK) were cultured in Iscove's modified Dulbecco's medium (IMDM, Lonza, Barcelona, Spain) supplemented with 10% fetal bovine serum (HyClone, GE Healthcare, Barcelona, Spain), 100 U mL^{-1} penicillin and 100 μg mL^{-1} streptomycin (Sigma Aldrich, Madrid, Spain) and cultured at 37 °C in a humidified incubator with 5% CO$_2$. Cells were transfected using Lipofectamine LTX with Plus Reagent (Thermo Fisher Scientific, Madrid, Spain) and selected using 1 mg mL^{-1} of G418 (Sigma Aldrich). For proteasomal inhibition studies, MG-132 (Calbiochem, Merck, Madrid, Spain) was added to the medium at 10 μg mL^{-1} for 6 h at 37 °C. After treatment with or without MG-132, cells were scrapped and lysed in RIPA buffer (50 mM Tris-HCl, 150 mM NaCl, 0.1% Triton X-100, 0.1% sodium dodecyl sulphate, 1 mM sodium orthovanadate, 1 mM NaF, pH 8.0) with protease inhibitors (COMPLETE, from Roche, Spain). Soluble extracts were denatured using Laemmli's buffer, resolved by SDS-PAGE and transferred to Polyvinylidene difluoride (PVDF) membranes (GE Healthcare). Immunoblotting was carried out using primary monoclonal antibodies anti-NQO1 (sc-393736, Santa Cruz Biotechnology, Dallas, TX, USA) and β-actin (sc-47778, Santa Cruz Biotechnology) from mouse at 1:200 and 1:5000 dilutions, respectively. As a secondary antibody, we used anti-mouse IgG-HRP (sc-516102, Santa Cruz Biotechnology) at 1:2000 dilution. Samples were visualized using luminol-based enhanced chemiluminescence (from BioRad Laboratories, Hercules, CA, USA) and analyzed using Image Lab (from BioRad Laboratories). NQO1 levels were normalized using the levels of β-actin.

3. Results and discussion

3.1. Effect of phosphomimetic mutations on FAD binding

We produced and purified WT, S40A/D, S82A/D and T128A/D variants of human NQO1 (Fig. S1). All variants seemed well-folded, as
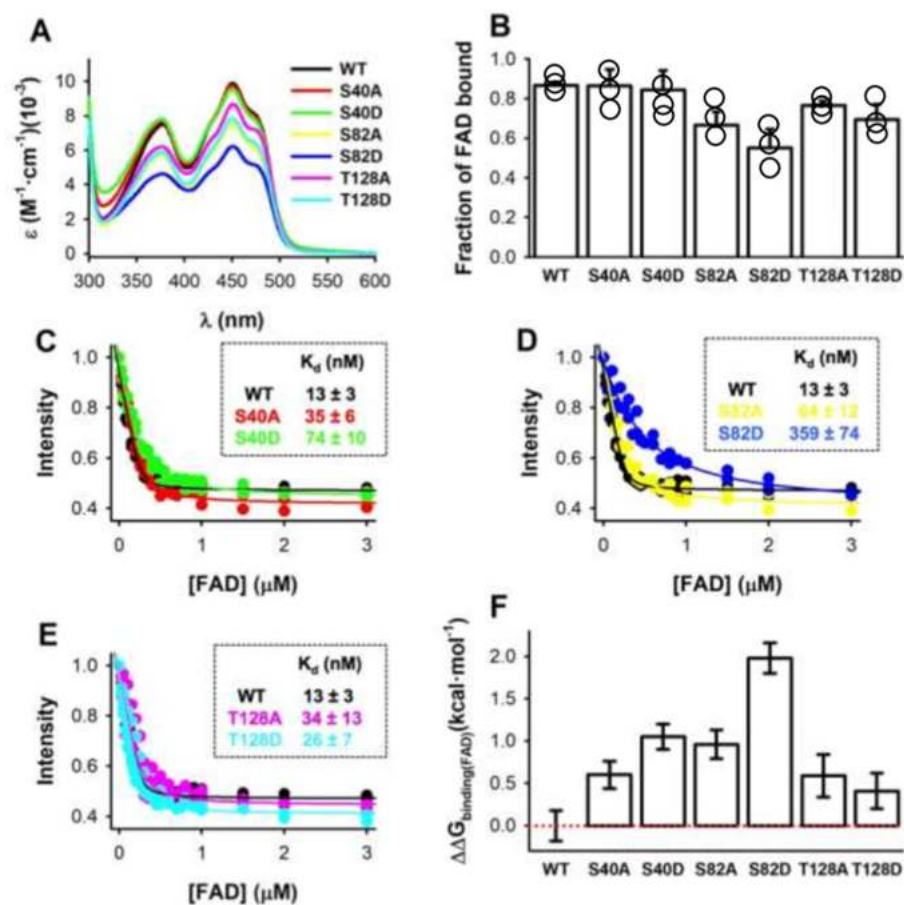
Fig. 2. FAD content and binding affinity. A) UV–vis spectra normalized per NQO1 monomer. B) The fraction of FAD bound determined from the spectra shown in A and considering $\varepsilon_{450} = 11300$ M^{-1} cm^{-1} for NQO1 fully saturated with FAD. Data were the mean $\pm$ s.d. from at least three independent purifications of each variant. Experiments in A-B were performed using $\sim$30 µM NQO1 monomer in 50 mM HEPES-KOH pH 7.4. C-E) Titrations of apo-proteins with FAD. K_d values were estimated from simultanenous fitting of data from 2 to 4 indepedent titrations for each variant; Titrations were carried out using 250 nM NQO1 monomer in 20 mM sodium phosphate, pH 7.4 at 25 °C; F) Changes in FAD binding free energy (mutant-WT). Errors are those from linear propagation from those in K_d values.
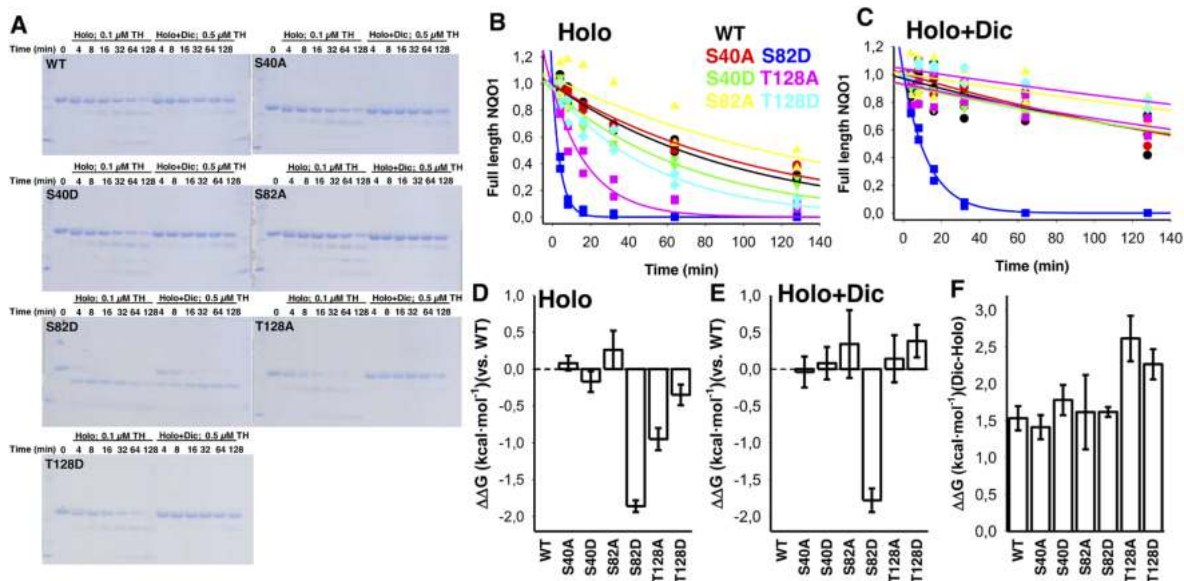
Fig. 3. Stability of the NTD evaluated by partial proteolysis. A) SDS-PAGE analysis of proteolysis kinetics with thermolysin. B–C) Protelysis kinetics of NQO1 proteins. Data are from at least 2 independent experiments for each NQO1 variant. D-E) Effect of different mutation on the thermodynamic stability of the cleavage site vs. WT ($\Delta\Delta G = - R \cdot T \cdot ln \frac{k_{mut}}{k_{WT}}$) as NQO1$_{holo}$ (D) and NQO1$_{dic}$ (E). F) Effect of Dic binding on the thermodynamic stability of the cleavage site for each variant ($\Delta\Delta G = - R \cdot T \cdot ln \frac{k_{dic}}{k_{holo}}$). Experiments were carried out using $\sim$15 µM NQO1 with 100 µM FAD or 100 µM FAD +100 µM Dic in 50 mM HEPES-KOH 10 mM CaCl$_2$, pH 7.4 at 25 °C and the thermolysin concentrations indicated. Errors in panels D–F are those lineally propagated from those of the second-order rate constants k_{prot}.

indicated by size-exclusion chromatography (SEC) and circular dichroism (CD) spectroscopy (Fig. S1). We then carried out" a detailed *in vitro* characterization of their functional properties. First, to test whether phosphomimetic mutations may affect FAD binding affinity, we measured the FAD content in NQO1 protein samples as purified, that qualitatively correlates with FAD binding affinity [25,29–31]. The most noticeable decrease in FAD content was found for S82D, whereas this decrease was milder in T128D (Fig. 2A and B). It must be noted that the mutated sites are not in close proximity to the FAD molecule: the minimal distance from S40 is of 14.5 Å, for S82 is 14 Å and for T128 is 12 Å (using the structure with PDB code 2F1O) [24] that supports long-range propagation of mutational effects to the FBS. To provide quantitative measurements on FAD binding affinity, we then prepared all variants in the NQO1$_{apo}$ state and carried out titrations with FAD monitored by fluorescence spectroscopy (Fig. 2C–F). The mutant S82D showed the largest effect, with a ~30-fold decrease in binding affinity (~2.0 kcal mol^{-1} in binding free energy), whereas the effects of S40D and T128D were milder (6-fold and 2-fold, corresponding to changes in binding free energy of ~1.1 kcal mol^{-1} and ~0.4 kcal mol^{-1}, respectively). These results highlight that the proximity of the phosphorylation site to the FBS or the solvent accessibility of the phosphorylated site does not solely determine their impact on FAD binding affinity. Since FAD binding affinity and intracellular flavin availability may be important for NQO1 intracellular stability [10], these analyses suggest that different phosphomimetic mutations may cause different response to riboflavin supplementation in terms of intracellular stability and activity.

3.2. Effect of the phosphomimetic mutations on the local stability of the NTD

Proteolysis by thermolysin is a useful tool for evaluating the effect of mutations and ligand binding on the thermodynamic local stability of the NTD, in particular due to initial cleavage between S72–V73 (the thermolysin cleavage site, TCS) [13]. Among the phosphomimetic mutants investigated, the most noticeable effect was found for S82D that destabilized by 2 kcal mol^{-1} both in the NQO1$_{holo}$ and NQO1$_{dic}$ states (Fig. 3). The effects of S40D and T128D were much weaker, and in the case of T128D, Dic binding essentially abolished the 0.5 kcal mol^{-1} destabilization caused by this mutation in the TCS of the NQO1$_{holo}$ state. Thus, destabilization of the TCS correlated well with the effect of the phosphomimetic mutations on FAD binding affinity on the NQO1$_{holo}$ state. It is interesting to note that the control variant T128A displays larger sensitivity towards proteolysis than T128D in the NQO1$_{holo}$ state, suggesting that at this residue, a decrease in side-chain causes more significant perturbations on the local stability of the NTD than introduction of a negative charge.

3.3. Effects of phosphomimetic mutations on conformational stability

The thermal stability of NQO1 provides information on the stability of the monomers and the MMI in the dimer, since the NQO1 dimers undergoes unfolding and dissociation prior to the rate-limiting step of irreversible denaturation [30]. We thus evaluated the effects of phosphomimetic mutations on the conformational stability as NQO1$_{holo}$ proteins (Fig. 4). The largest effects were observed for the mutations S40D and S82D, that decreased by ~7 °C the thermal stability of the enzyme. It is interesting to note that denaturation of S82D showed a wider profile, possibly affecting the denaturation mechanism (e.g. this mutation may affect the *coupling* between domains in the irreversible denaturation process). Consequently, the S40D and S82D mutations may destabilize the NQO1 monomer and/or the MMI in the dimer.

3.4. Effects of phosphomimetic mutations on the local stability from HDX

Hydrogen-deuterium exchange (HDX) has been proven to provide high-resolution information on the local stability effects on NQO1 due to mutations and ligand binding, including long-range propagation of these effects across the structure [18,21]. Since the EX1 regime was marginally detected (Fig. S2), it is reasonable to interpret our HDX data using a local thermodynamic stability approach (i.e. EX2) and, in this context, little information on protein dynamics can be inferred.

Despite their different solvent accessibilities (% ASA; S40 < S82 < T128), the impact of phosphomimetic mutations on the local stability of the protein, as determined by HDX, did not correlate with this feature and was strongly dependent on the ligation state (Fig. 5). The S40D mutant showed weak local destabilizing effects based on a EX2 mechanism, and also mildly affected residues of the CTD in the NQO1$_{apo}$ state (Figs. S2 and S3). HDX analyses also allow to suggest that, despite their solvent-accessibility in the crystal structure, these residues are exposed to the solvent (i.e. accessible for phosphorylation) due to local thermodynamic fluctuations (Fig. S2).

In the NQO1$_{apo}$ state, the destabilizing effects of S40D localized in the vicinity of the mutated residue and the CTD, whereas FAD binding restricted these effects to the local environment of S40 (Fig. 5A and Figs. S3–S4). The effects of S82D extended to almost the entire structure of NQO1$_{apo}$ and were of mild-to-moderate intensity (Fig. 5B and Figs. S3–S4). FAD binding restricted these destabilizing effects more locally around the mutated site. The effects of the T128D mutation were very local, destabilizing residues comprising 69–76, while stabilizing residues 126–131 in both the NQO1$_{apo}$ and NQO1$_{holo}$ states (Fig. 5C and Figs. S3–S4).

We then used our HDX studies to investigate different functional sites, namely the residues involved in FAD binding (i.e. the FBS, with implications for FAD binding affinity)(Fig. 6A), solvent-exposed residues (that could be involved in PPI)(Fig. 6B) and the MMI (whose stability is associated with the thermal stability of the enzyme)(Fig. 6C).
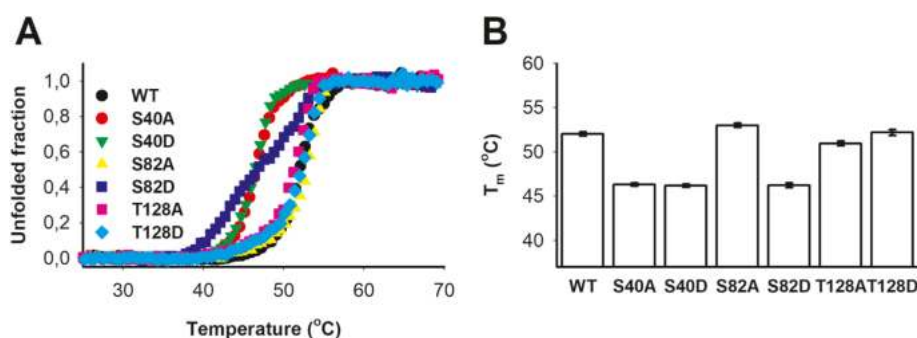
Fig. 4. Thermal stability of NQO1 variants. A) Thermal denaturation profiles of NQO1$_{holo}$ variants as monitored by fluorescence spectroscopy. B) Half-denaturation temperatures (T$_m$) from at least three different technical replicas (mean ± s.d.). Experiments were carried out using 2 μM NQO1 in the presence 20 μM of FAD in 50 mM HEPES-KOH, pH 7.4 and using a scan rate of 2 °C min^{-1}.
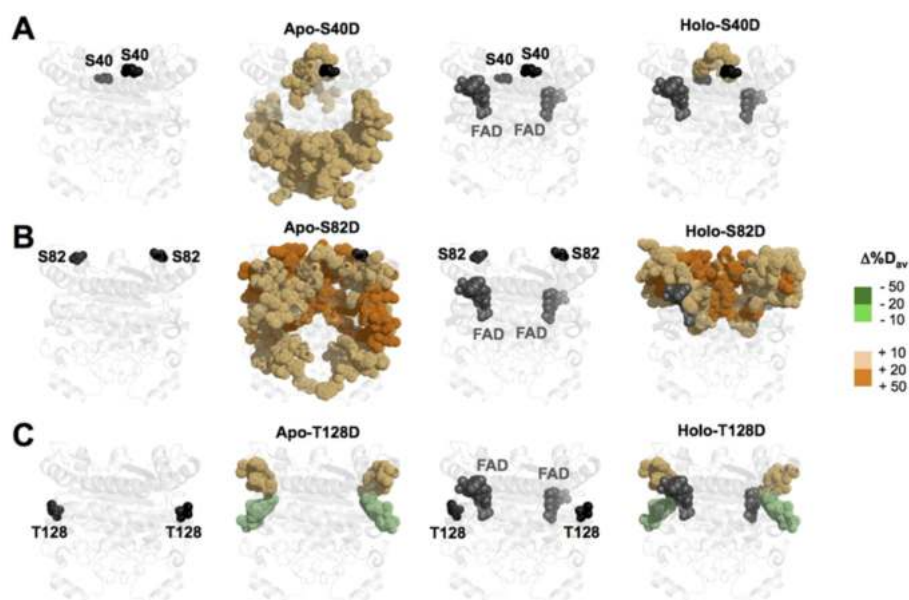
Fig. 5. Overall effects of phosphomimetic mutations on the local stability by HDXMS. A-C) Effect of the mutations S40D (A), S82D (B) and T128D (C) on the local stability of the NQO1$_{apo}$ and NQO1$_{holo}$ states. Changes in stability (as $\Delta\%D_{av}$ using the WT protein in a given ligation state as a reference) are colored according to the scale indicated in the right side of the figure (green tones indicate stabilization whereas orange tones indicate destabilization). Structural display was carried out using the structure with PDB code 2F1O [24]. Experiments were carried out using 20 μM of apo-protein (NQO1$_{apo}$) or purified protein with 200 μM of FAD (NQO1$_{holo}$) and the H/D exchange reaction was initiated by a 10-fold dilution of the protein solution into D$_2$O-based 50 mM HEPES-KOH, pD 7.4, 1 mM TCEP at 25 °C. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
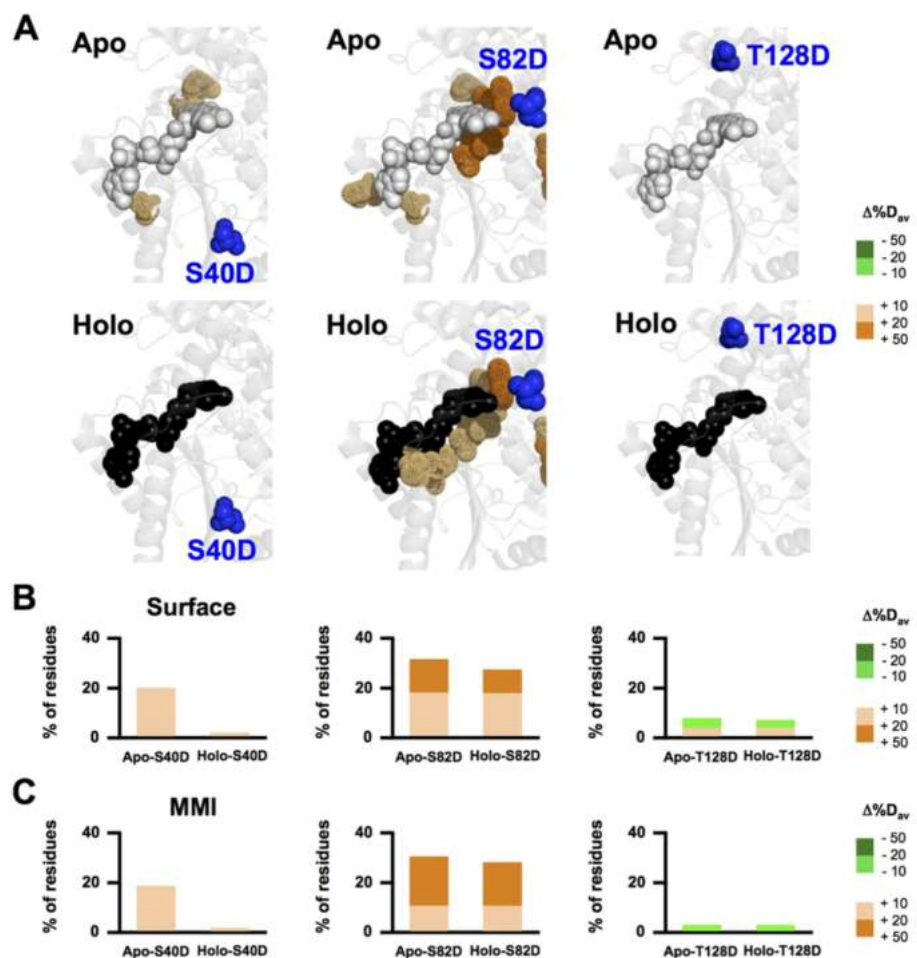
Fig. 6. Effect of phosphomimetic mutations on the local stability of the FBS, protein surface and the MMI. A-C) Effect of the mutations S40D, S82D and T128D mutation on the stability of the FBS in the NQO1$_{apo}$ and NQO1$_{holo}$ states. B) Percentage of residues in the protein surface (considering those with at least 20% solvent accessibility the software GETAREA; http://curie.utmb.edu/getarea.html; (7)); C) Effect of phosphomimetic mutations on the stability of residues belonging to the MMI (considering those determined according to Ref. [21]). Changes in stability (as $\Delta\%D_{av}$ using the WT protein as a reference) are colored according to the scale indicated in the right side of the figure (green tones indicate stabilization whereas orange tones indicate destabilization). Structural display and calculations were carried out using the structure with PDB code 2F1O [24]. Experiments were performed as indicated in the legend of Fig. 5. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Regarding the stability of the FBS, the mutation S82D caused large changes in stability of the FBS, both in the NQO1$_{apo}$ and NQO1$_{holo}$ states, thus explaining its large effect on FAD binding affinity (Figs. 2 and 6A). The mutation S40D only decreased the stability of small set of residues (A21, Y156 and H162) of the FBS in the NQO1$_{apo}$, thus explaining its milder effect on FAD binding affinity (Figs. 2 and 6A). The lack of noticeable effects of the T128D on the stability of the FBS explains its marginal effect on FAD binding affinity vs. the WT protein (Figs. 2 and

6A).

Residues at the surface of NQO1 may be important for PPI. When we analyzed the effects of the phosphomimetic mutations S40D, S82D and T128D on the stability of solvent-exposed regions of the protein, we also observed different behaviors within the phosphomimetic mutants (Fig. 6B). In the NQO1$_{apo}$ state, the mutation S40D reduced the stability of about 20% of solvent-exposed residues, whereas upon FAD binding, this effect was virtually abolished (Fig. 6B). The mutation S82D affected the stability of about 30% solvent-exposed residues in both the NQO1$_{apo}$ and NQO1$_{holo}$ states (Fig. 6B). The impact of the T128D mutation on solvent-exposed residues was minimal, with mild stabilizing and destabilizing effects (Fig. 6B).

Thermal stability analyses on NQO1$_{holo}$ proteins revealed significant destabilization due to S40D and S82D phosphomimetic mutations (Fig. 4). We then used our HDX studies to determine whether changes in the stability of the MMI correlated with thermal destabilization (Fig. 6C, S5-S6). In the case of S40D, a large destabilization of the MMI is observed in the NQO1$_{apo}$ state. For S82D, the MMI is largely destabilized in both the NQO1$_{apo}$ and NQO1$_{holo}$ states. Changes in the MMI stability were minimal for T128D, consistent with its thermal stability being close to that of the WT protein. Therefore, the large effects of S40D and S82D mutations on thermal stability correlated well with their effect on the stability of the MMI.
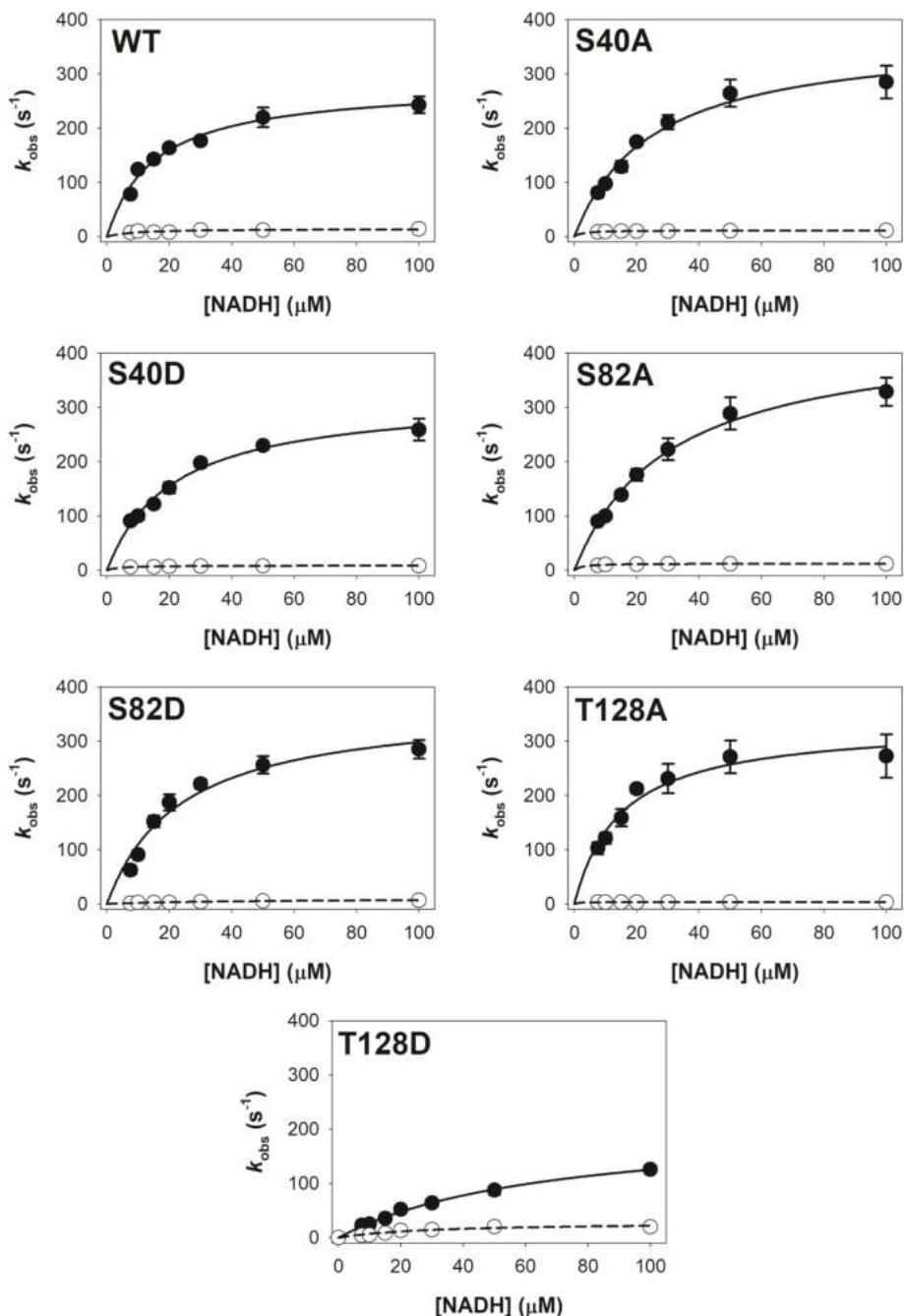
Fig. 7. Effect of phosphomimetic mutations on the reductive half-reaction of FAD using NADH. Data correspond to the fast reduction (closed circles) and slow reduction (open circles) steps and are displayed as mean $\pm$ s.d. from three replicates. Lines are fits using equation (3). Fitting parameters are compiled in Table 1. Experiments were carried out using 7.5 μM of NQO1$_{holo}$ in 50 mM HEPES-KOH, pH 7.4 and 6 °C.

Table 1
Kinetic parameters for the NADH-dependent reduction of FAD.

Variant	Fast reduction			Slow reduction		
	k_{HT} (s^{-1})	K_d (NADH) (µM)	$k_{HT}/K_{d(NADH)}$ (s^{-1}·µM^{-1})	k_{HT} (s^{-1})	K_d (NADH) (µM)	$k_{HT}/K_{d(NADH)}$ (s^{-1}·µM^{-1})
WT	281 ± 14	15 ± 2	19 ± 3	14 ± 2	8.2 ± 3.6	1.7 ± 1.0
S40A	372 ± 22	25 ± 4	15 ± 3	12 ± 1	2.3 ± 0.3	5.2 ± 1.1
S40D	320 ± 13	21 ± 2	15 ± 2	8.6 ± 0.3	3.8 ± 0.5	2.3 ± 0.4
S82A	441 ± 19	30 ± 3	15 ± 2	12 ± 1	2.3 ± 0.2	5.2 ± 0.5
S82D	368 ± 33	23 ± 5	16 ± 5	10 ± 1	43 ± 8	0.23 ± 0.06
T128A	333 ± 21	15 ± 3	22 ± 6	3.5 ± 0.1	1.1 ± 0.5	3.2 ± 1.5
T128D	210 ± 12	67 ± 7	3.1 ± 0.5	28 ± 4	29 ± 10	1.0 ± 0.5

3.5. Effects of phosphomimetic mutations on the reductive half-reaction of NQO1

The reduction of the two flavin molecules within the NQO1 dimer occurs through pathways with very different rates (*fast* and *slow* steps) [16]. We have evaluated the effect of phosphomimetic mutations on these two pathways by evaluating rates for HT from NADH by stopped-flow measurements (Fig. 7 and Table 1). Regarding the *fast* pathway of FAD reduction, most of the mutants did not affect significantly the k_{HT} or $K_{d(NADH)}$ values, although the mutation T128D reduced by 4.5-fold the affinity for NADH and by 6-fold the catalytic efficiency (Table 1). In the case of the *slow* reduction pathway, the mutation S82D reduced the affinity for NADH by 5-fold and the catalytic efficiency by 7-fold (Table 1). The mutation T128D increased both the k_{HT} and $K_{d(NADH)}$ for the slow step leading to little changes in catalytic efficiency (Table 1). Therefore, the functional cooperativity observed for the WT enzyme is affected differently by the phosphomimetic mutations S82D and T128D (about 4-fold increase and 6-fold decrease, respectively, as the ratio of the catalytic efficiency of the *fast/slow* steps vs. that of WT) (Table 1). It is interesting to note that based on HDX data ($\Delta\%D_{av}$), when we compare the NQO1$_{holo}$ and NQO1$_{dic}$ states of T128D and the WT proteins, some changes in the DBS (but not the FBS) are found for some residues, with different effects on these two states (Pro69 showed decreased stability whereas Tyr127, Tyr129 and Met132 were moderately stabilized)(Figs. S7–S9).

These effects of the mutations S82D and T128D on the enzyme catalytic efficiency for the fast and slow HT steps are likely associated with structural or energetic changes in the active site along the enzyme reaction coordinate. We have recently shown changes in the local stability of the active site (the FBS and DBS) in the NQO1$_{holo}$ and NQO1$_{dic}$ states due to the mutation S82D, particularly destabilization of these functional sites in the former state [18]. Here, we have carried out similar experiments with the mutant T128D (Figs. S7–S9). The mutation T128D affected locally and differently the NQO1$_{holo}$ and NQO1$_{dic}$ states. T128D mildly reduced the stability of a small set of residues in the vicinity of the NQO1$_{holo}$ active site (residues 69–74) whereas its effects were stabilizing, stronger and more extensive in the NQO1$_{dic}$ active site (affecting residues 69–74 and 122–132). A differential analysis of these effects on the local stability of the active site between NQO1$_{dic}$ and NQO1$_{holo}$ clearly showed that the overall effect of the T128D mutation on the active site stability is moderately stabilizing (Fig. S9). Considering that changes in the local stability of the active site between NQO1$_{holo}$ and NQO1$_{dic}$ states might reflect protein stability changes along the reaction coordinate for the HT process [16,18], our results suggest that the mechanisms associated with altered catalytic efficiency of S82D and T128D are different in structural and energetic terms. These results also

suggest that site-specific phosphorylation could affect the catalytic efficiency of NQO1 by altering the balance between structural, energetic and dynamic changes along the HT reaction.

To assess the impact of mutations on the dynamics in the active site of NQO1 along its catalytic function, we analyzed HT by using NADH and NADD at different temperatures (Figs. S10–S11). For all variants, *fast* and *slow* reduction steps were slowed down with NADD (Table S1), in general resulting, as reported for the WT, in kinetic isotope effects (KIEs) in the range 1.5–2.2 at the lower temperature evaluated. It is only worth noting a slight increase in KIE for the *fast* process of T128D and the lack of KIE for the *slow* process of T128A. KIEs for the *fast* process are essentially temperature-independent, as for the WT, with the value only slightly decreasing with temperature for S82A (Fig. S11, Table S1). Regarding the *slow* FAD reduction step, only the WT and S40D KIEs were temperature-independent. KIEs particularly decreased with temperature for S40A, S82A and S82D and slightly increased when mutating T128 (Fig. S11). This indicates that mutations hardly alter transitions under the energy barrier and tunnelling of both proton and deuterium in the *fast* HT step, but have a long-distance modulating effect in the *slow* step [16,18]. Analysis of data in the context of Arrhenius theory [activation energies (E$_a$) and pre-exponential factors (A)] indicated that for the *fast* step of all variants *passive dynamics* is the main source to achieve tunnel ready conformations (Fig. S10). The largest differences were found for *slow* steps. Particularly in S40A, S82A, S82D and T128A that showed remarkable changes in A values (orders of magnitude) and in E$_a$, *gating* (donor-acceptor-distance sampling) possibly contributes to engage more favourable conformations for a possible tunneling. Thus, replacements at sites S40, S82 and T128 particularly modulate the overall packing and general dynamics of the *slow* active site of NQO1 during HT, and therefore, also contribute to modulate cooperativity in the reductive half-reactions of both active sites in the protein dimer.

3.6. Effects of phosphomimetic mutations on steady-state protein levels upon expression in eukaryotic cells

To evaluate the effect of phosphomimetic mutants on the
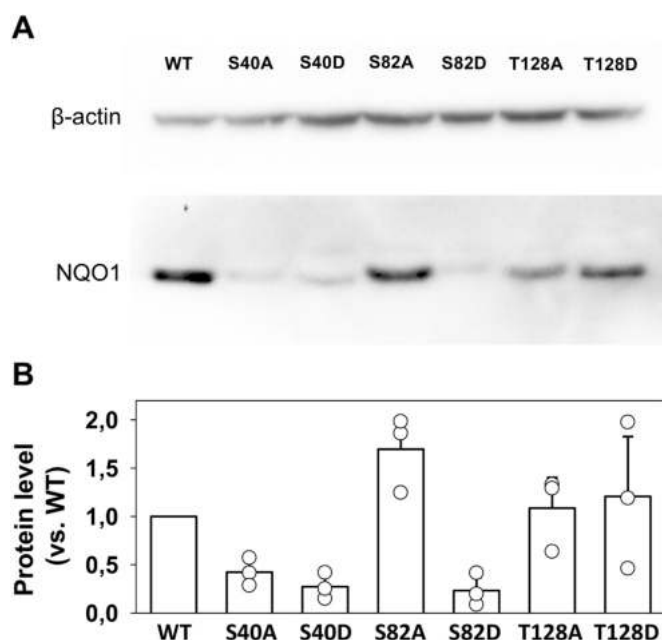
Fig. 8. Effect of phosphomimetic mutations on the NQO1 steady-state protein levels upon expression in stably-transfected HAP-1 KO cells. A) Representative Western-blot analyses; B) Protein levels expressed (considering WT as 1). Data are from three independent transfection experiments (mean ± s.d.) and were normalized by the levels of β-actin.

intracellular protein levels, we transfected HAP-1 KO cells with plasmids encoding WT as well as mutant forms of NQO1 (Fig. 8 and Figs. S12–18). Under these conditions, we considered that steady-state protein levels were determined to some extent by proteasomal degradation of the NQO1 variants (Figs. S12–S15) in agreement with previous studies [2, 10,18,32]. The behaviors observed were clearly site-specific. At the S40 site, either S40A or S40D mutants showed lower protein levels (about 30–40% of WT levels), suggesting that even a small perturbation at this site (e.g. S → A) causes significant perturbation of NQO1 intracellular stability. This result suggests that non-conservative mutations S40L and S40W (very rare mutations identified in large-scale sequencing projects of the human genome; i.e. gnomAD database) have dramatic effects on NQO1 intracellular stability. At the S82 site, only the S82D mutant showed decreased protein levels (about 20% of WT), while mutations at the T128 site had little or no effect on protein levels under these conditions. Therefore, our results support that phosphorylation at S40 and S82 sites could lead to reduced protein levels due to accelerated NQO1 degradation.

4. Conclusions

Protein phosphorylation is the most ubiquitous post-translational modification and serves to rapidly modify protein functionality [33]. In the particular case of human flavoproteins, our knowledge on the effects of site-specific phosphorylation is largely unknown [22]. In this work, we have carried out a comparative study of three phosphorylation sites (S40, S82 and T128) in human NQO1, a multifunctional flavoprotein whose dysregulation is associated with cancer and neurological disorders [1,2]. To this end, we used phosphomimetic mutation (S40D, S82D and T128D). These phosphorylation sites display different solvent accessibilities in crystallographic structures, although these solvent accessibilities do not correlate with their impact on the structure, stability and function of NQO1 (Table 2). For instance, the phosphomimetic

mutation S40D affects a fully buried serine residue in the protein crystal structure, and causes noticeable effects on *in vitro* and intracellular stability but mild effects on FAD binding affinity. The mutation S82D, affecting a more solvent-exposed residue, affects the conformational and intracellular stabilities, FAD binding and enzyme catalysis of NQO1, and its destabilizing effects propagate more extensively through the protein structure than any of the other sites. The effects of T128D are quite subtle, only causing local stability effects (stabilizing and destabilizing) in the vicinity of the mutated residue that moderately affects enzyme catalysis and cooperativity without a clear effect for the overall *in vitro* or intracellular stability in HAP-1 cells. It is important to realize that our extensive structure-function analyses were carried out using *pseudo-phosphorylating* mutations to isolate the effects of a given individual site. At this point, we realize that a full view of phosphorylation dynamics must include the characterization of a more extended set of phosphorylation sites, the kinases/phosphatase involved, the potential interaction between sites and the parallel or sequential phosphorylation state of different sites [2,22]. We anticipate that our experimental approach will help to improve our knowledge on this challenging task, and would be applied to investigate other human flavoproteins.

CRediT author statement

Juan Luis Pacheco-Garcia.- Data curation, formal analysis, investigation, methodology, resources, validation, visualization, writing – review & editing. **Ernesto Anoz-Carbonell**.- Data curation, formal analysis, investigation, methodology, resources, validation, visualization, writing – review & editing. **Dmitry S. Loginov**.- Data curation, formal analysis, investigation, methodology, resources, validation, visualization, writing – review & editing. **Pavla Vankova**.- Data curation, formal analysis, investigation, methodology, resources, validation, visualization, writing – review & editing. **Eduardo Salido**.- Funding acquisition, resources, writing – review & editing. **Petr Man**.- Data

Table 2

Summary of the effects of phosphomimetic mutations. For each protein variant and feature, these were clustered for semiquantitative comparison in four categories: (++++) mildly improved vs. WT; (+++) WT-like; (++) mildly-moderately impaired vs. WT; (+) largely impaired vs. WT.

Protein feature	NQO1 variant						
	WT	S40A	S40D	S82A	S82D	T128A	T128D
Conformation[a]	+++	+++	+++	+++	+++	+++	+++
FAD bound[b]	+++	+++	+++	++	++	+++	++
FAD binding affinity[c]	+++	+++	++	++	+	+++	+++
Proteolysis Holo[d]	+++	+++	+++	+++	+	++	+++
Proteolysis response to Dic[e]	+++	+++	+++	+++	+++	++++	++++
Thermal stability[f]	+++	+	+	+++	+	+++	+++
Global HDX Apo[g]	+++	N.Det.*	++	N.Det.*	+	N.Det.*	+++
Global HDX Holo[h]	+++	N.Det.*	+++	N.Det.*	+	N.Det.*	+++
FBS HDX[i]	+++	N.Det.*	+++	N.Det.*	+	N.Det.*	+++
MMI HDX[j]	+++	N.Det.*	++	N.Det.*	+	N.Det.*	+++
Fast FAD reduction[k]	+++	+++	+++	+++	+++	+++	+
Slow FAD reduction[l]	+++	++++	++	++++	++	++	++
Cooperativity[m]	+++	++++	+++	++++	+	+++	++
Intracellular abundance[n]	+++	+	+	+++	+	+++	+++

* N. Det.: not determined.
a Based on SEC and CD spectroscopy (Fig. S1).
b Based on visible absorption spectroscopy (Fig.2A and B).
c Based on FAD titrations (Fig. 2C–E).
d Based on proteolysis kinetics as holo-proteins (Fig. 3B and D).
e Based on the effects of Dic on proteolysis kinetics (Fig. 3D–F).
f Based on the effects on thermal stability (Fig. 4).
g Based on global HDX for the apo-state (Fig. 5A–C).
h Based on global HDX for the **holo**-state (Fig. 5A-C).
i Based on HDX results regarding the FBS (Fig. 6A).
j Based on HDX results regarding the MMI (Fig. 6C).
k Based on the results for the *fast* reduction pathway, including k_{HT}, $K_{d(NADH)}$ and $k_{HT}/K_{d(NADH)}$ (Table 1).
l Based on the results for the *slow* reduction pathway, including k_{HT}, $K_{d(NADH)}$ and $k_{HT}/K_{d(NADH)}$ (Table 1).
m Based on the results ($k_{HT}/K_{d(NADH)}$) for the *fast/slow* reduction pathways (Table 1).
n Based on Western-blot analysis of transfected cells grown without MG-132 (Fig. 8).

curation, formal analysis, funding acquisition, methodology, resources, validation, visualization, software, supervision, writing – review & editing. **Milagros Medina**.- Data curation, formal analysis, funding acquisition, methodology, resources, validation, visualization, software, supervision, writing – review & editing. **Rogelio Palomino-Morales**.- Data curation, formal analysis, methodology, resources, validation, visualization, supervision, writing – review & editing. **Angel L. Pey**.- Conceptualization, Data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, validation, visualization, software, supervision, writing - original draft, writing – review & editing.

Funding

Declaration of competing interest

The authors declare no competing, personal financial or other competing interests. Funding or employment sources played no role in the design, presentation or discussion of the presented research.

Data availability

Data will be made available on request.

Acknowledgments

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.abb.2022.109392.

References

[1] S.K. Beaver, N. Mesa-Torres, A.L. Pey, D.J. Timson, NQO1: a target for the treatment of cancer and neurological diseases, and a model to understand loss of function disease mechanisms, Biochim. Biophys. Acta, Proteins Proteomics 1867 (2019) 663–676, https://doi.org/10.1016/j.bbapap.2019.05.002.

[2] S. Luo, S.S. Kang, Z.H. Wang, X. Liu, J.X. Day, Z. Wu, J. Peng, D. Xiang, W. Springer, K. Ye, Akt phosphorylates NQO1 and triggers its degradation, abolishing its antioxidative activities in Parkinson's disease, J. Neurosci. 39 (2019) 7291–7305, https://doi.org/10.1523/JNEUROSCI.0625-19.2019.

[3] D. Ross, D. Siegel, The diverse functionality of NQO1 and its roles in redox control, Redox Biol. 41 (2021), https://doi.org/10.1016/j.redox.2021.101950.

[4] D. Ross, D. Siegel, NQO1 in protection against oxidative stress, Curr. Opin. Toxicol. 7 (2018) 67–72, https://doi.org/10.1016/j.cotox.2017.10.005.

[5] E. Salido, D.J. Timson, I. Betancor-Fernández, R. Palomino-Morales, E. Anoz-Carbonell, J.L. Pacheco-García, M. Medina, A.L. Pey, Targeting HIF-1α function in cancer through the chaperone action of NQO1: implications of genetic diversity of NQO1, J. Personalized Med. 12 (2022) 747, https://doi.org/10.3390/jpm12050747.

[6] G. Asher, P. Tsvetkov, C. Kahana, Y. Shaul, A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73, Genes Dev. 19 (2005) 316–321, https://doi.org/10.1101/gad.319905.

[7] G. Ben-Nissan, M. Sharon, Regulating the 20S proteasome ubiquitin-independent degradation pathway, Biomolecules 4 (2014) 862–884, https://doi.org/10.3390/biom4030862.

[8] A. di Francesco, C. di Germanio, A.C. Panda, P. Huynh, R. Peaden, I. Navas-Enamorado, P. Bastian, E. Lehrmann, A. Diaz-Ruiz, D. Ross, D. Siegel, J. L. Martindale, M. Bernier, M. Gorospe, K. Abdelmohsen, R. de Cabo, Novel RNA-binding activity of NQO1 promotes SERPINA1 mRNA translation, Free Radic. Biol. Med. 99 (2016) 225–233, https://doi.org/10.1016/j.freeradbiomed.2016.08.005.

[9] E.T. Oh, J.W. Kim, J.M. Kim, S.J. Kim, J.S. Lee, S.S. Hong, J. Goodwin, R. J. Ruthenborg, M.G. Jung, H.J. Lee, C.H. Lee, E.S. Park, C. Kim, H.J. Park, NQO1 inhibits proteasome-mediated degradation of HIF-1α, Nat. Commun. 7 (2016), 13593 https://doi.org/10.1038/ncomms13593.

[10] A. Martínez-Limón, M. Alriquet, W.H. Lang, G. Calloni, I. Wittig, R.M. Vabulas, Recognition of enzymes lacking bound cofactor by Protein quality control, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 12156–12161, https://doi.org/10.1073/pnas.1611994113.

[11] M. Faig, M.A. Bianchet, P. Talalay, S. Chen, S. Winski, D. Ross, L.M. Amzel, Structures of recombinant human and mouse NAD(P)H:quinone oxidoreductases: species comparison and structural changes with substrate binding and release, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 3177–3182, https://doi.org/10.1073/pnas.97.7.3177.

[12] R. Li, M.A. Bianchet, P. Talalayt, L.M. Amzel, The three-dimensional structure of NAD(P)H:quinone reductase, a flavoprotein involved in cancer chemoprotection and chemotherapy: mechanism of the two-electron reduction, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 8846–8850, https://doi.org/10.1073/pnas.92.19.8846.

[13] E. Medina-Carmona, R.J. Palomino-Morales, J.E. Fuchs, E. Padín-Gonzalez, N. Mesa-Torres, E. Salido, D.J. Timson, A.L. Pey, Conformational dynamics is key to understanding loss-of-function of NQO1 cancer-associated polymorphisms and its correction by pharmacological ligands, Sci. Rep. 6 (2016), 20331, https://doi.org/10.1038/srep20331.

[14] E. Medina-Carmona, J.L. Neira, E. Salido, J.E. Fuchs, R. Palomino-Morales, D. J. Timson, A.L. Pey, Site-to-site interdomain communication may mediate different loss-of-function mechanisms in a cancer-associated NQO1 polymorphism, Sci. Rep. 7 (2017), https://doi.org/10.1038/srep44532.

[15] W.D. Lienhart, V. Gudipati, M.K. Uhl, A. Binter, S.A. Pulido, R. Saf, K. Zangger, K. Gruber, P. Macheroux, Collapse of the native structure caused by a single amino acid exchange in human NAD(P)H:Quinone oxidoreductase, FEBS J. 281 (2014) 4691–4704, https://doi.org/10.1111/febs.12975.

[16] E. Anoz-Carbonell, D.J. Timson, A.L. Pey, M. Medina, The catalytic cycle of the antioxidant and cancer-associated human NQO1 enzyme: hydride transfer, conformational dynamics and functional cooperativity, Antioxidants 9 (2020) 1–22, https://doi.org/10.3390/antiox9090772.

[17] C.F. Megarity, H. Abdel-Aal Bettley, M.C. Caraher, K.A. Scott, R.C. Whitehead, T. A. Jowitt, A. Gutierrez, R.A. Bryce, K.A. Nolan, I.J. Stratford, D.J. Timson, Negative cooperativity in NAD(P)H quinone oxidoreductase 1 (NQO1), Chembiochem 20 (2019) 2841–2849, https://doi.org/10.1002/cbic.201900313.

[18] J.L. Pacheco-Garcia, E. Anoz-Carbonell, P. Vankova, A. Kannan, R. Palomino-Morales, N. Mesa-Torres, E. Salido, P. Man, M. Medina, A.N. Naganathan, A.L. Pey, Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue, Redox Biol. 46 (2021), 102112, https://doi.org/10.1016/j.redox.2021.102112.

[19] A. Martínez-Limón, G. Calloni, R. Ernst, R.M. Vabulas, Flavin dependency undermines proteome stability, lipid metabolism and cellular proliferation during vitamin B2 deficiency, Cell Death Dis. 11 (2020) 725, https://doi.org/10.1038/s41419-020-02929-5.

[20] D. Siegel, S. Bersie, P. Harris, A. di Francesco, M. Armstrong, N. Reisdorph, M. Bernier, R. de Cabo, K. Fritz, D. Ross, A redox-mediated conformational change in NQO1 controls binding to microtubules and α-tubulin acetylation, Redox Biol. 39 (2021), 101840, https://doi.org/10.1016/j.redox.2020.101840.

[21] P. Vankova, E. Salido, D.J. Timson, P. Man, A.L. Pey, A dynamic core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer, Biomolecules 9 (2019) 728, https://doi.org/10.3390/biom9110728.

[22] E. Medina-Carmona, B. Rizzuti, R. Martín-Escolano, J.L. Pacheco-García, N. Mesa-Torres, J.L. Neira, R. Guzzi, A.L. Pey, Phosphorylation compromises FAD binding and intracellular stability of wild-type and cancer-associated NQO1: insights into flavo-proteome stability, Int. J. Biol. Macromol. 125 (2019) 1275–1288, https://doi.org/10.1016/j.ijbiomac.2018.09.108.

[23] J.L. Pacheco-Garcia, D. Loginov, B. Rizzuti, P. Vankova, J.L. Neira, D. Kavan, N. Mesa-Torres, R. Guzzi, P. Man, A.L. Pey, A single evolutionarily divergent mutation determines the different FAD-binding affinities of human and rat NQO1 due to site-specific phosphorylation, FEBS Lett. 596 (2022) 29–41, https://doi.org/10.1002/1873-3468.14238.

[24] G. Asher, O. Dym, P. Tsvetkov, J. Adler, Y. Shaul, The crystal structure of NAD(P)H quinone oxidoreductase 1 in complex with its potent inhibitor dicoumarol, Biochemistry 45 (2006) 6372–6378, https://doi.org/10.1021/bi0600087.

[25] E. Medina-Carmona, J.E. Fuchs, J.A. Gavira, N. Mesa-Torres, J.L. Neira, E. Salido, R. Palomino-Morales, M. Burgos, D.J. Timson, A.L. Pey, Enhanced vulnerability of human proteins towards disease-associated inactivation through divergent evolution, Hum. Mol. Genet. 26 (2017) 3531–3544, https://doi.org/10.1093/hmg/ddx238.

[26] J.E. Fuchs, I.G. Muñoz, D.J. Timson, A.L. Pey, Experimental and computational evidence on conformational fluctuations as a source of catalytic defects in genetic diseases, RSC Adv. 6 (2016) 58604–58612, https://doi.org/10.1039/C6RA05499D.

[27] Z. Zhang, D.L. Smith, Determination of Amide Hydrogen Exchange by Mass Spectrometry: A New Tool for Protein Structure Elucidation, Cambridge University Press, 1993.

[28] Y. Perez-Riverol, J. Bai, C. Bandla, D. García-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D.J. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher, M. Walzer, S. Wang, A. Brazma, J.A. Vizcaíno, The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences, Nucleic Acids Res. 50 (2022), https://doi.org/10.1093/nar/gkab1038. D543–D552.

[29] A.L. Pey, Biophysical and functional perturbation analyses at cancer-associated P187 and K240 sites of the multifunctional NADP(H):quinone oxidoreductase 1, Int. J. Biol. Macromol. 118 (2018) 1912–1923, https://doi.org/10.1016/j.ijbiomac.2018.07.051.

[30] A.L. Pey, C.F. Megarity, D.J. Timson, FAD binding overcomes defects in activity and stability displayed by cancer-associated variants of human NQO1, Biochim. Biophys. Acta (BBA) - Mol. Basis Dis. (2014) 2163–2173, https://doi.org/10.1016/j.bbadis.2014.08.011, 1842.

[31] J.L. Pacheco-García, M. Cano-Muñoz, I. Sánchez-Ramos, E. Salido, A.L. Pey, Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein, J. Personalized Med. 10 (2020) 1–31, https://doi.org/10.3390/jpm10040207.

[32] D. Siegel, A. Anwar, S.L. Winski, J.K. Kepa, K.L. Zolman, D. Ross, Rapid polyubiquitination and proteasomal degradation of a mutant form of NAD(P)H: quinone oxidoreductase 1, Mol. Pharmacol. 59 (2001) 263–268.

[33] J.A. Ubersax, J.E. Ferrell Jr., Mechanisms of specificity in protein phosphorylation, Nat. Rev. Mol. Cell Biol. 8 (2007) 530–541, https://doi.org/10.1038/nrm2203.

[34] R. Fraczkiewicz, W. Braun, Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules, J. Comput. Chem. 19 (1998) 319–333. https://doi.org/10.1002/(SICI)1096-987X(199802)19:3<319::AID-JCC6>3.0.CO;2-W. https://doi.org/10.1002/(SICI)1096-987X(199802)19:3%3c319::AID-JCC6%3e3.0.CO;2-W.

MATERIAL SUPLEMENTARIO DE LA PUBLICACIÓN 4

Different phenotypic outcome due to site-specific phosphorylation in the cancer-associated NQO1 enzyme studied by phosphomimetic mutations.

Different phenotypic outcome due to site-specific phosphorylation in the cancer-associated NQO1 enzyme studied by phosphomimetic mutations

Juan Luis Pacheco-Garcia, Ernesto Anoz-Carbonell, Dmitry S. Loginov, Pavla Vankova, Eduardo Salido, Petr Man, Milagros Medina, Rogelio Palomino-Morales, and Angel L. Pey.

Supplementary methods

Far-UV circular dichroism (CD) spectroscopy

Far-UV CD spectroscopy was performed at 25 ºC in K-phosphate 20 mM at pH 7.4 using 5 µM (in monomer) of protein with 25 µM FAD. Spectra were collected in a Jasco J-710 spectropolarimeter in the 190-260 nm range, at 100 nm·min⁻¹, using 1 nm bandwidth, 1 s response time and 1 mm path length cuvette. Each spectrum was the average of six scans, and each sample was prepared in triplicate. To calculate mean residue ellipticities ($[\Theta]_{MRW}$), we used the following equation:

$$[\Theta]_{MRW} = \frac{MRW \cdot \Theta_{obs}}{10 \cdot l \cdot c}$$

where MRW was equal to the molecular weight of the NQO1 monomer (31691 g·mol⁻¹) divided by (N-1), being N = 280 the number of residues in the monomer, Θ_{obs} was the ellipticity (in mdeg), l was the path length (in cm) and c the concentration of protein in mg·mL⁻¹. Spectra were reported as mean ± s.d. from three replicates. Results can be found in Figure S1B.

Analysis of deuterium exchange kinetic mechanism

The exchange mechanism of NQO1 in solution was assessed through the presence of EX1 kinetics. Peak width at 1% of signal intensity was derived for each peptide (individual horizontal lines in the graphs in Figure S2) and all time points (captured through the red gradient) for all protein forms (mutants) and protein states (apo/holo). The values shown in these graphs were calculated as follows – peak widths of all protein forms/states across individual time point were averaged and each peak width was divided by this number. The final value thus represents indication if the peak width of a particular point deviates from the average value. In theory, for a fully EX2 exchanging region the value should be close to 1, whereas a shift toward higher numbers indicates presence of EX1 kinetics. It should be noted that the chosen intensity threshold of 1% is very sensitive and thus captures small contributions from EX1. Situations where the EX2 and EX1 are

1

somehow equal in intensity is however very rare for the NQO1 mutants/states studied here and limited to region 155-165 in $S40D_{apo}$ and $S82D_{apo}$ and to the ultimate C-terminus of these two protein/ligation states (based on detailed analysis of isotopic profiles).

Supplementary data

Figure S1. Purification and folding of NQO1 proteins. A) NQO1 proteins purified from IMAC (5-10 mg) were injected into a HiLoad 16/60 SuperDex 200 pg column. SEC was carried out in 20 mM HEPES-NaOH 200 mM NaCl, pH 7.4 at 20ºC and a 1.5 mL·min^{-1} flow rate. For each protein, the fractions indicated by vertical red dashed lines and corresponding to dimeric NQO1 were pooled, concentrated and stored. The SDS-PAGE gel shows the purity of the stored protein after SEC. B) Far-UV circular dichroism spectra for all holo-protein variants. Spectra were acquired as described in [1]. For each sample, 6 scans were carried out and the results for each variant is the average±s.d. for three independent preparations. The spectra for WT NQO1 are included for comparison.
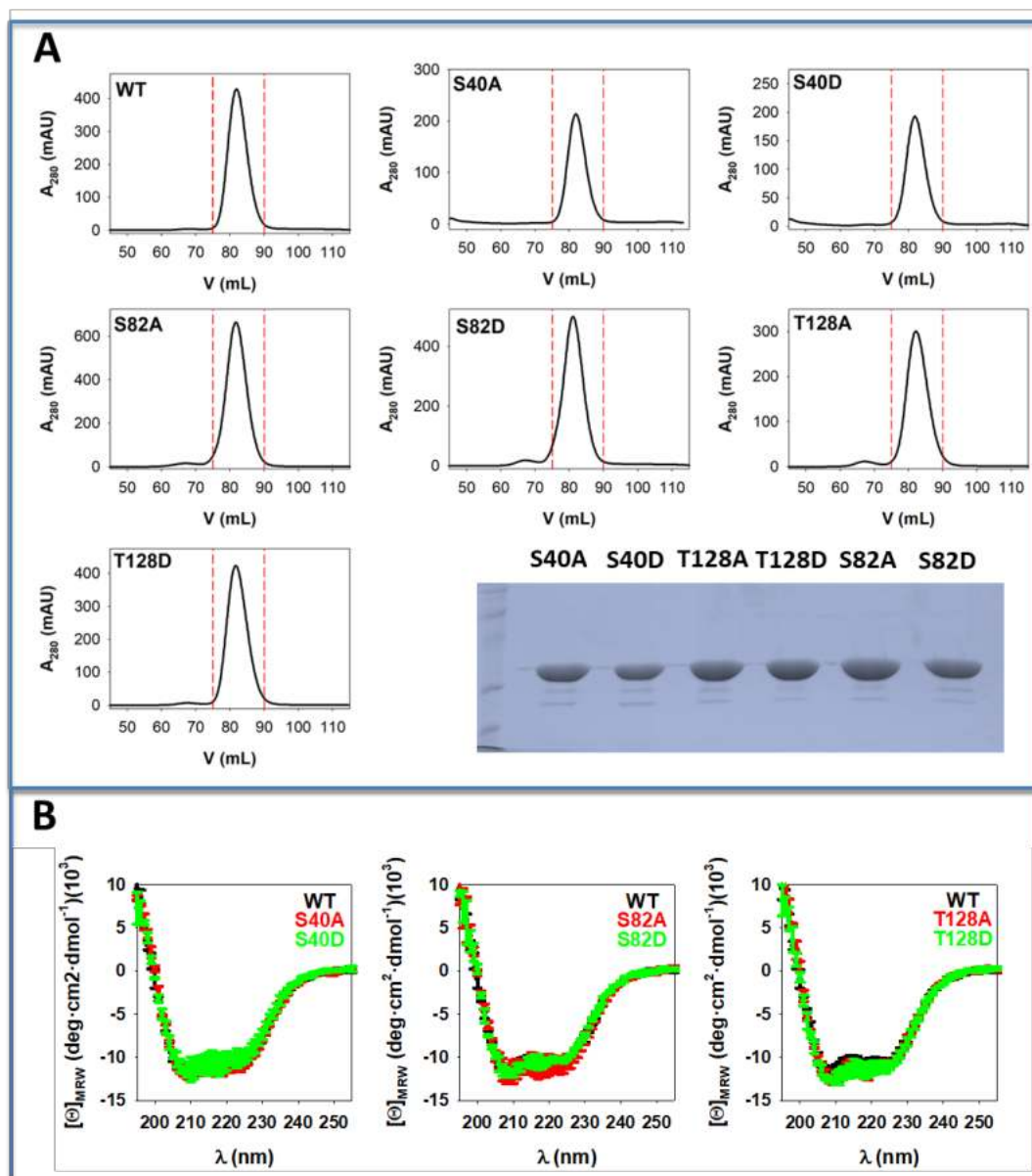
Figure S2. EX1 regime is not highly populated in the HDX of WT and phosphomimetic mutants. Details on these analyses can be found in supplementary methods. Left and right panels show data for apo-(a) and holo-(a)proteins, respectively. Individual lines are the peptides and the time-resolved kinetics is captured through the red gradient (shown above each graph, in seconds). Systematic shift toward values >1 indicates presence of EX1 exchange kinetics. Analysis was performed with high sensitivity and only two regions, 155-165 and the C-terminus, are showing more prominent EX1 in S40D$_{apo}$ and S82D$_{apo}$ states.
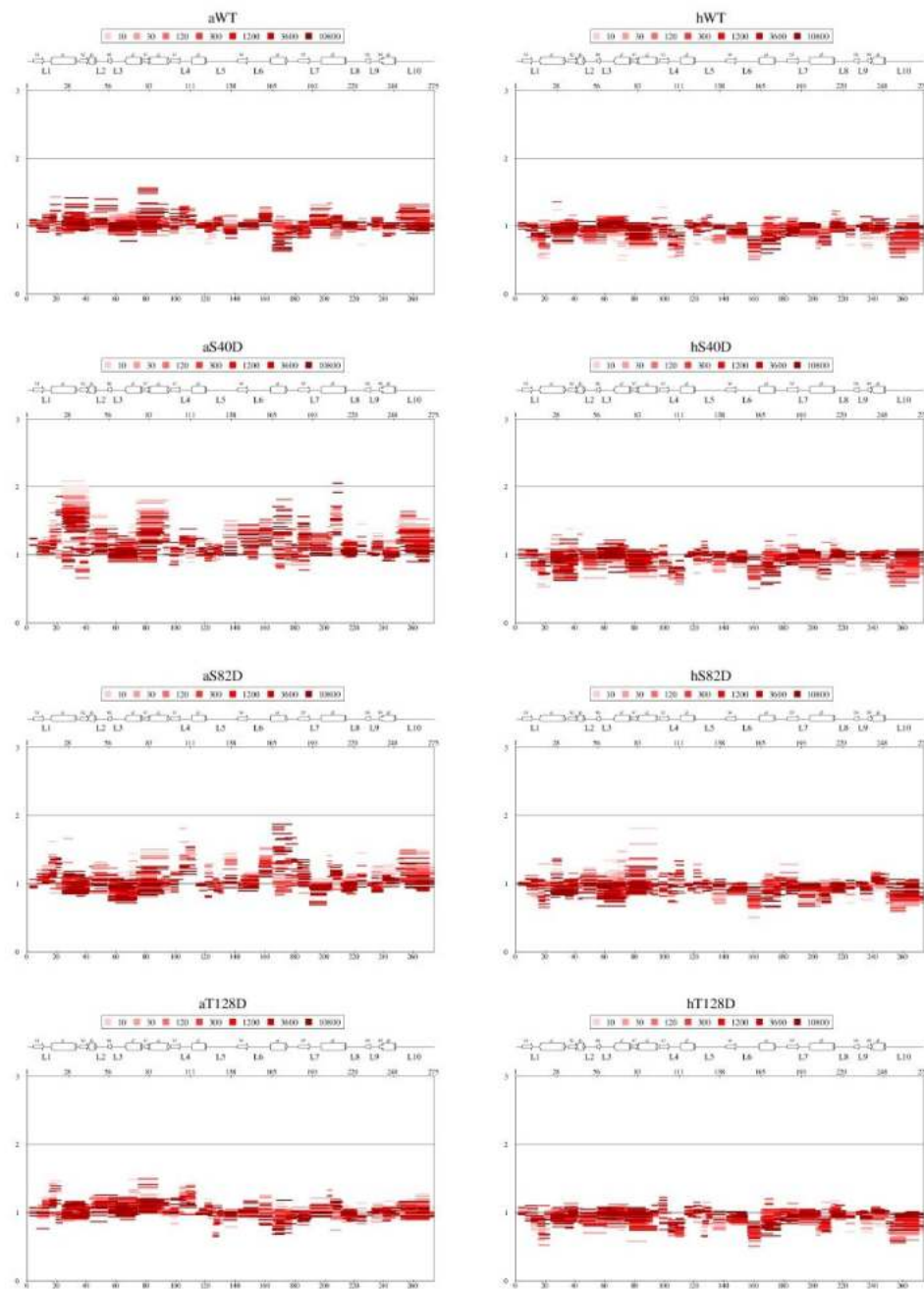
Figure S3. Rainbow heatmaps showing the deuteration profile for WT, S40D, S82D and T128D in their apo (a) and holo (h) forms. Each thick section represents individual protein state (indicated on the left) and is subdivided into the individual time points. The x-axis represents the primary sequence. The secondary structure elements are shown above the heatmap according to [2] (PDB: 1D4A). The residues forming the MMI, the DBS and the FBS are also depicted. Color scale in %D is displayed below the heatmap.
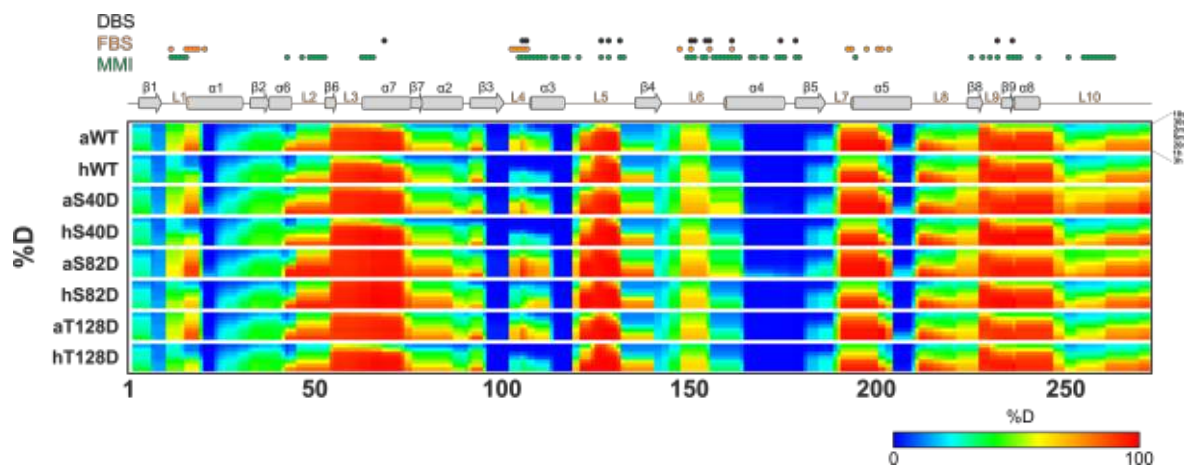
Figure S4. Structural display of HDX using a rainbow color representation. Two time points are shown for NQO1$_{apo}$ and NQO1$_{holo}$ states. HDX for the earliest (A) and latest (B) time points. The structure used for display was that reported by [3] (PDB code 2F1O).
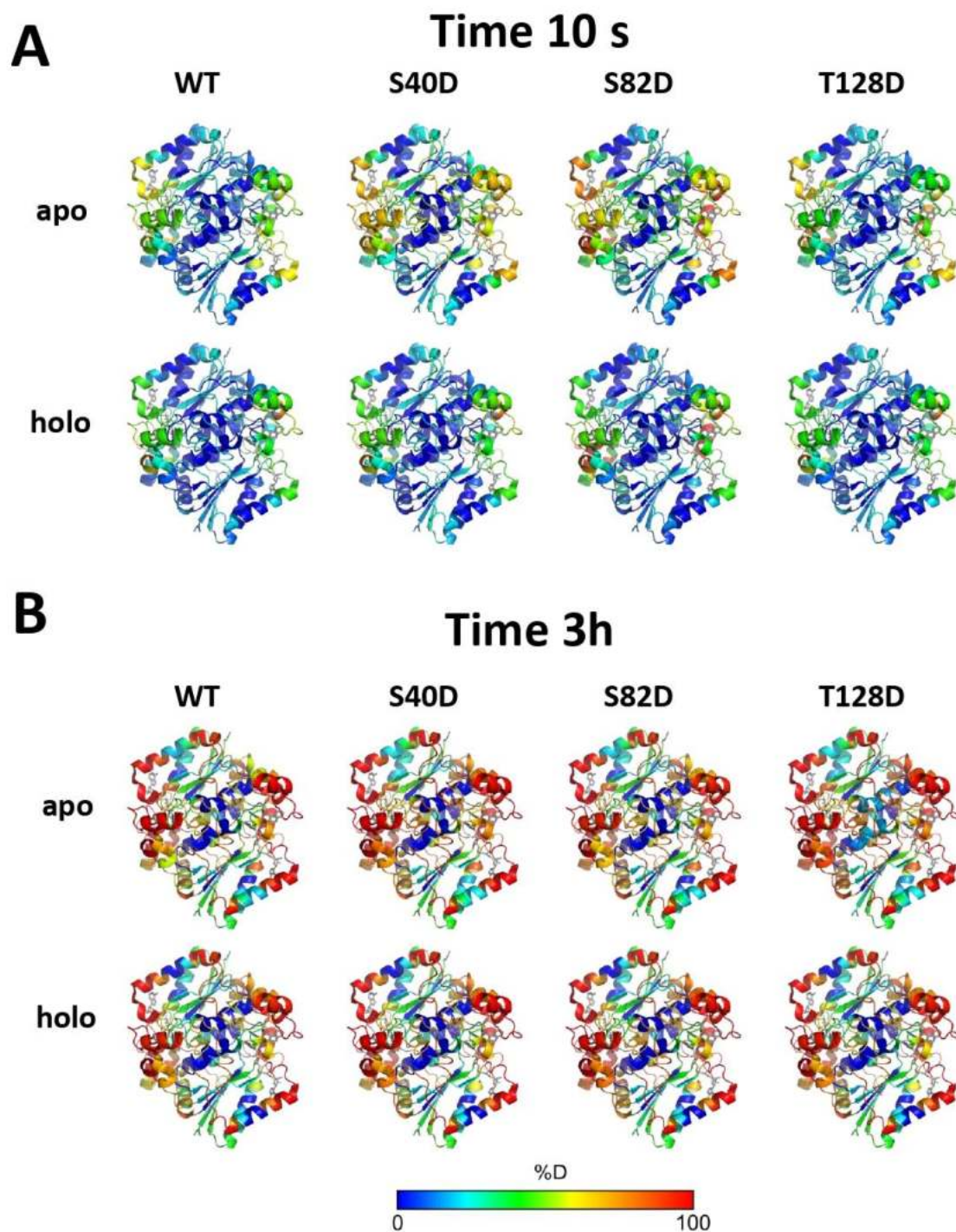
Figure S5. Effects of phosphomimetic mutations on the structural stability of the MMI in the NQO1$_{apo}$ state. A) Structural location of residues S40, S82 and T128 that have been mutated to D. B) Representation of the MMI; residues in black are those involved in the MMI based on [4] and displayed using the structure with PDB 2F1O [3]; C-E) Structural location of residues that show changes in local stability compared to WT NQO1. Residues are colored according to the residues with values of $\Delta\%D_{av}$ higher or lower than 10%. Two different views rotated by 90º are shown. Experimental details of HDX experiments can be found in the main text.
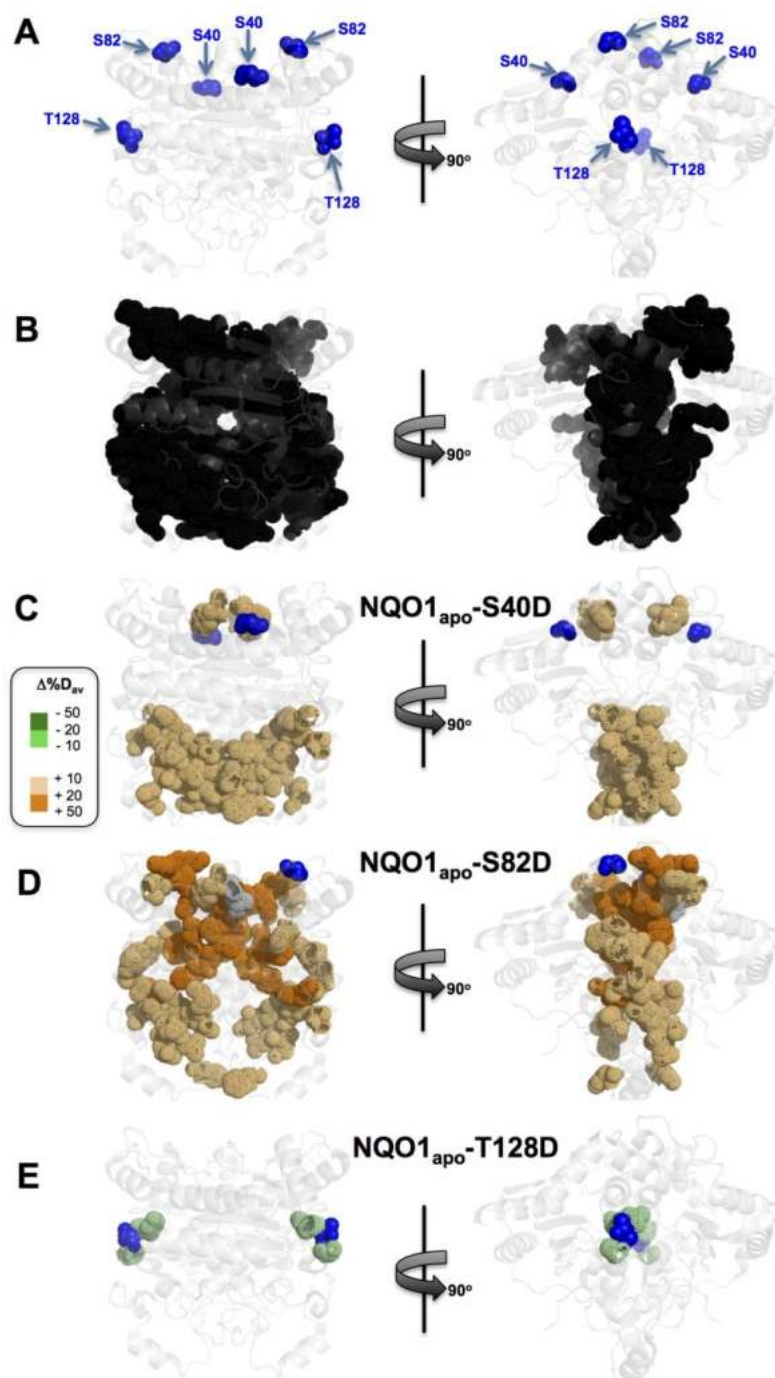
Figure S6. Effects of phosphomimetic mutations on the structural stability of the MMI in the NQO1$_{holo}$ state. A) Structural location of residues S40, S82 and T128 that have been mutated to D. B) Representation of the MMI; residues in black are those involved in the MMI based on [4] and displayed using the structure with PDB 2F1O [3]; C-E) Structural location of residues that show changes in local stability compared to WT NQO1. Residues are colored according to the residues with values of $\Delta\%D_{av}$ higher or lower than 10%. Two different views rotated by 90º are shown. Experimental details of HDX experiments can be found in the main text.
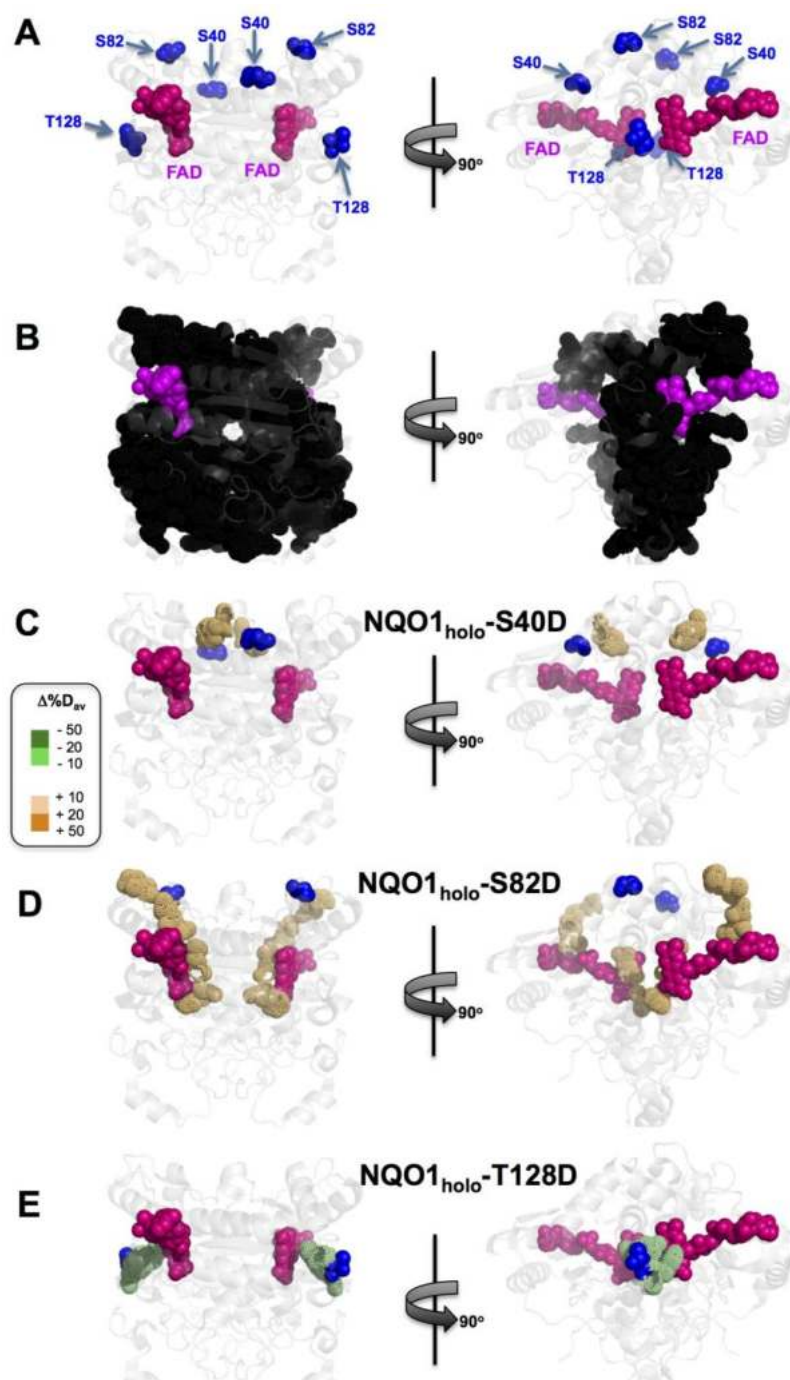
Figure S7. Effects of the T128D mutation on the structural stability of the active site (FBS and DBS) in the NQO1$_{holo}$ and NQO1$_{dic}$ states. Panels in the left side show the structural location of T128, FAD and Dic. Panels in the right side show residues that change in local stability compared to WT NQO1. Residues are colored according to the residues with values of $\Delta\%D_{av}$ higher or lower than 10%. Two different views rotated by 90º are shown. Experimental details of HDX experiments can be found in the main text. The plot in the right side shows $\Delta\%D_{av}$ values over the full sequence of NQO1. The structure with PDB code 2F1O [3] was used for display.
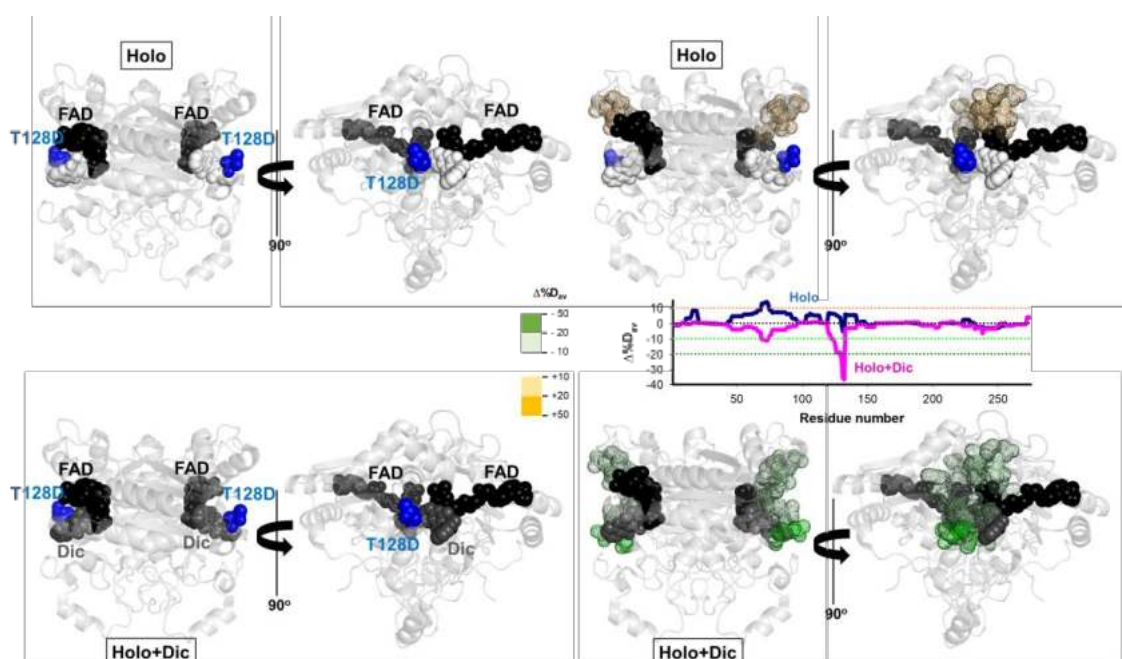
Figure S8. Rainbow heatmaps showing the deuteration profile for WT and T128D in their apo (a), holo (h) and dicoumarol (d) forms. Each thick section represents individual protein state (indicated on the left) and is subdivided into the individual time points. The x-axis represents the primary sequence. The secondary structure elements are shown above the heatmap according to [2] (PDB: 1D4A). The residues forming the MMI, the DBS and the FBS are also depicted. Color scale in %D is displayed below the heatmap.
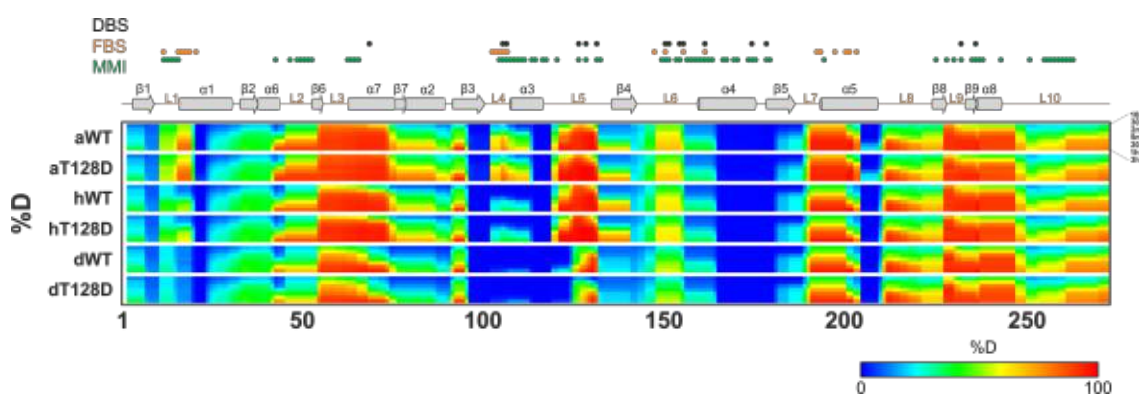
Figure S9. Difference of the effects of the T128D mutation on the structural stability of the active site (FBS and DBS) in the NQO1$_{dic}$ and NQO1$_{holo}$ states. Panel A shows the difference in stability between the mutant and the WT protein in the NQO1$_{dic}$ and NQO1$_{holo}$ states over the full sequence of NQO1 (i.e. the difference between $\Delta\%D_{av}$ between NQO1$_{dic}$ and NQO1$_{holo}$ states for T128D vs. WT). Panel B shows the structural location of T128, FAD and Dic (upper panels) and changes in this sum of local stability compared to WT NQO1 (lower panels). Residues are colored according to the residues with values of $\Delta\%D_{av}$ lower than 10%. Two different views rotated by 90º are shown. Experimental details of HDX experiments can be found in the main text. The structure with PDB code 2F1O was used for display [3].
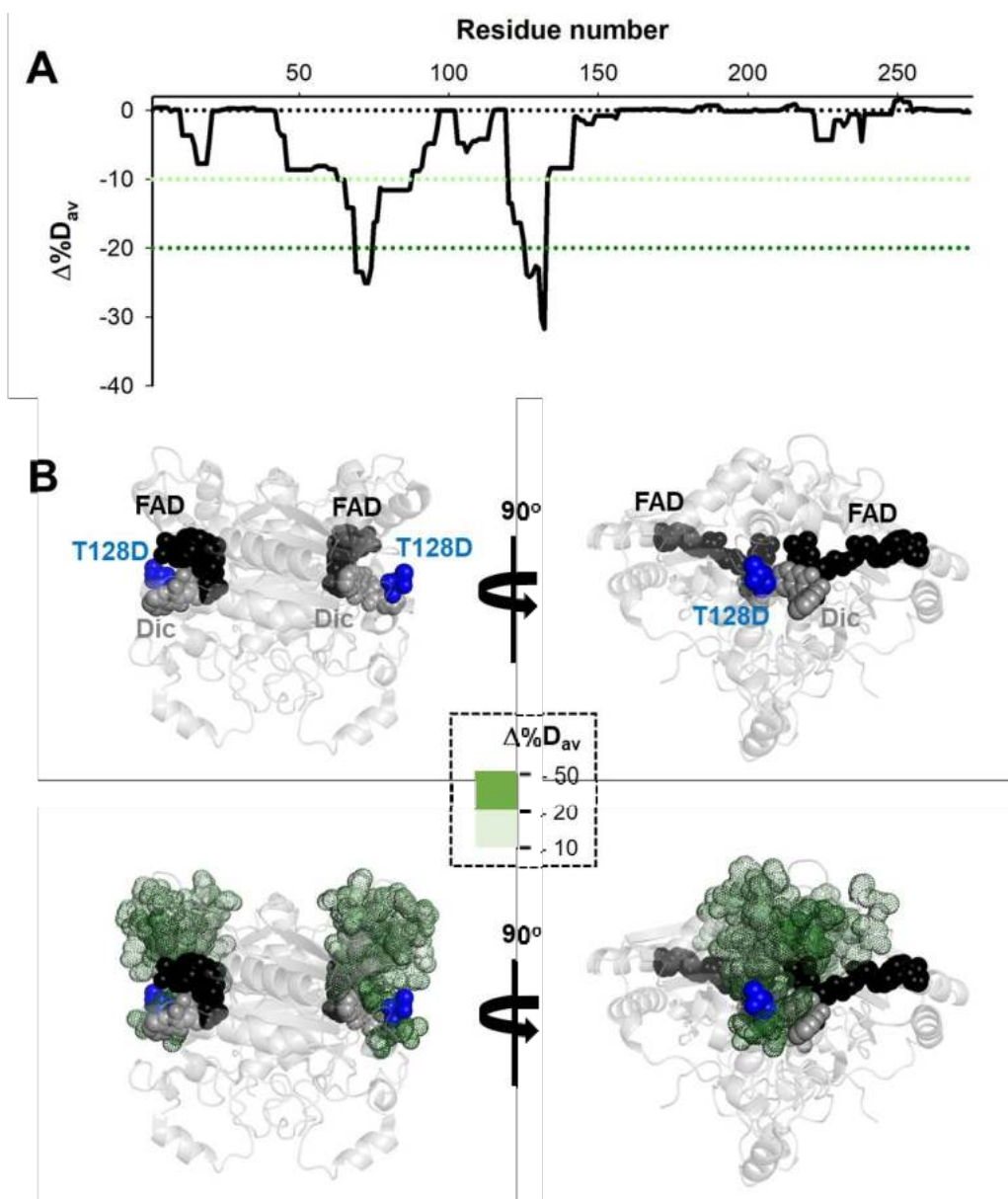
Figure S10. Temperature dependence of k_{obs} for the two (fast and slow) hydride/deuteride transfer (HT/DT) processes from NADH/D to NQO1. Data for HT and DT are displayed as circles and squares, respectively. Arrhenius plots were used to calculate the activation parameters compiled in Table S1. Data for WT and S82D are taken from [5,6].
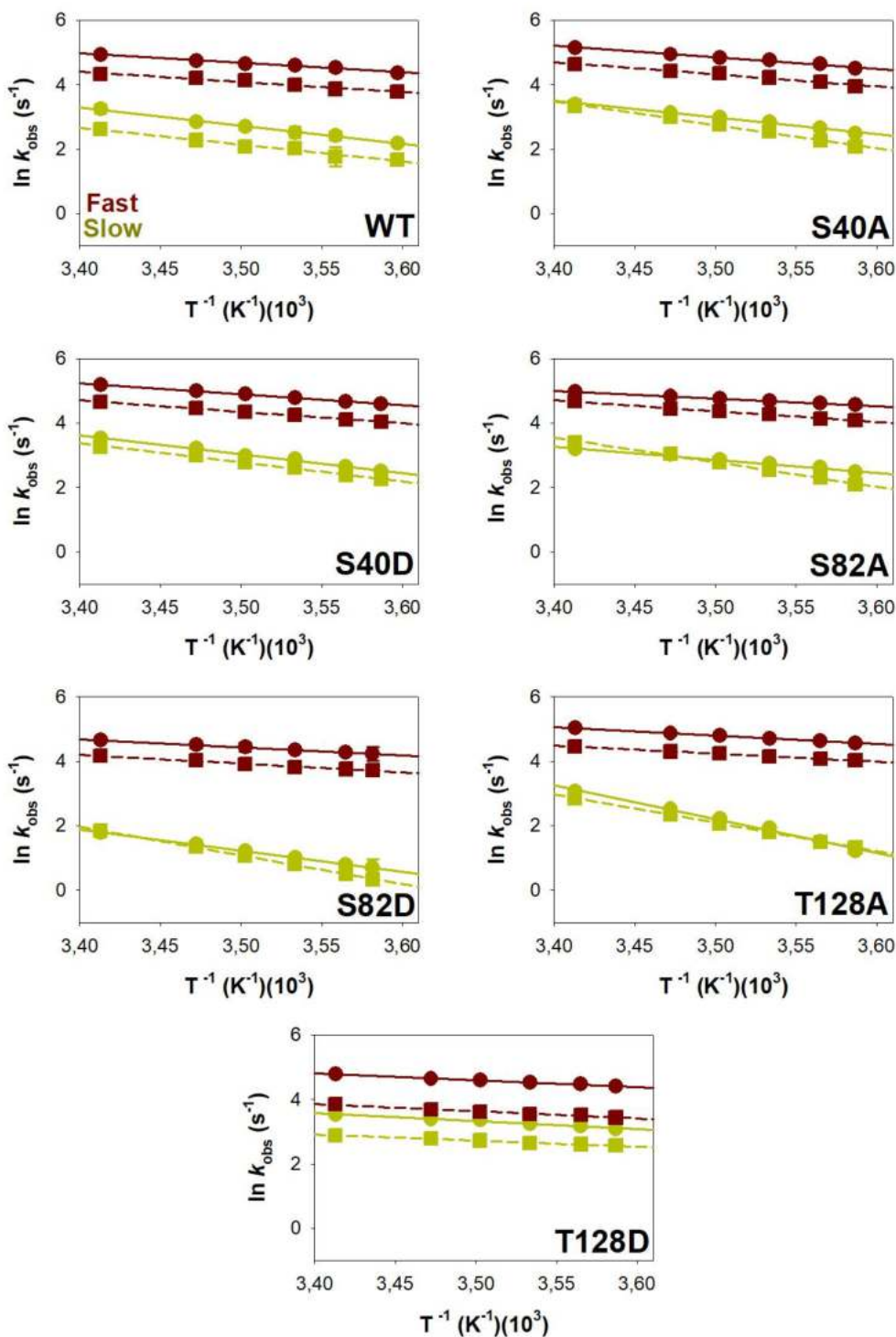
Figure S11. Temperature dependence of kinetic isotope effects (KIEs) for parameters for the two (fast and slow) hydride/deuteride transfer (HT/DT) processes from NADH/D to NQO1. KIEs were calculated from the quotient of rate constants for HT and DT. Data for WT and S82D are taken from [5,6].
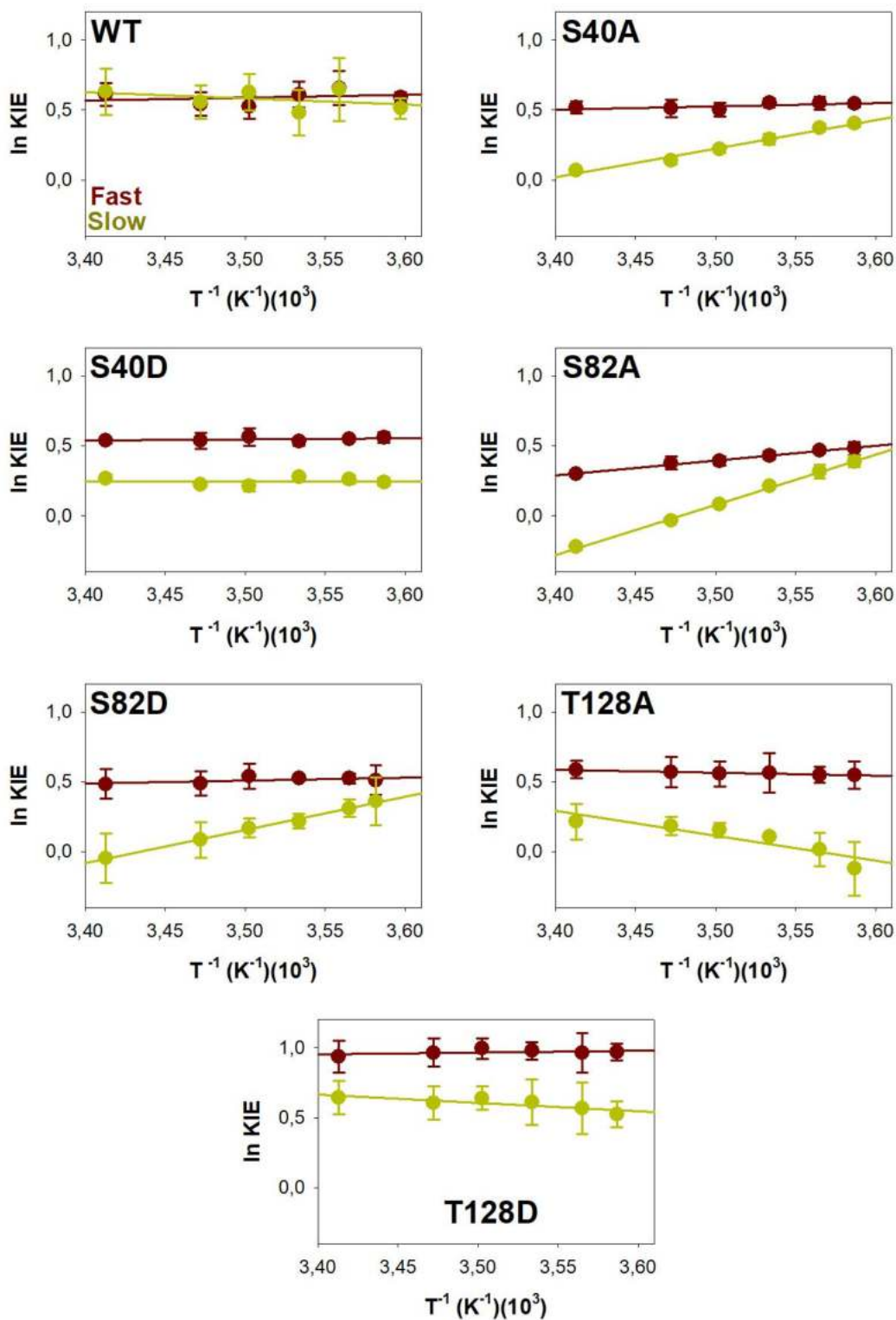
Figure S12. Western-blot analysis of HAP-1KO cells transfected with NQO1 variants in the absence (-) or presence (+) of the proteasomal inhibitor MG-132. As positive control for proteasomal degradation, the variant P187S was included [6-8]. Experimental details can be found in the main text.
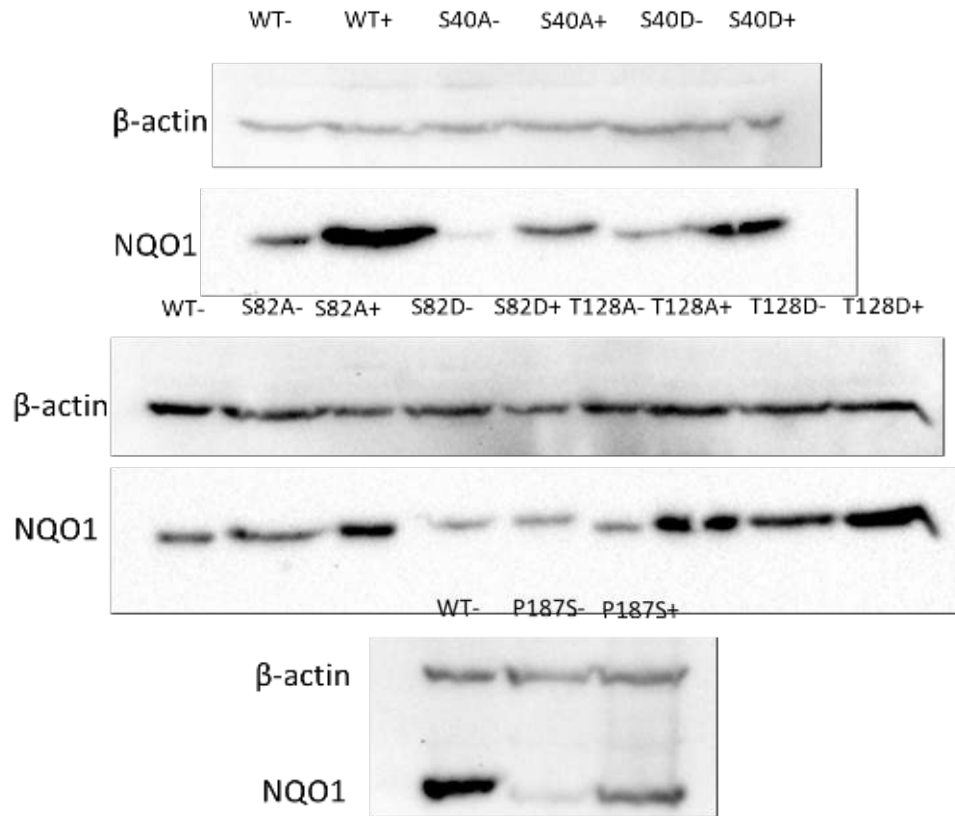
Figure S13. Full image of Western-blot analysis of HAP-1KO cells transfected with WT and S40A/D NQO1 variants in the absence (-) or presence (+) of the proteasomal inhibitor MG-132. The antibody used was anti-β-actin after using an anti-NQO1 antibody.

.
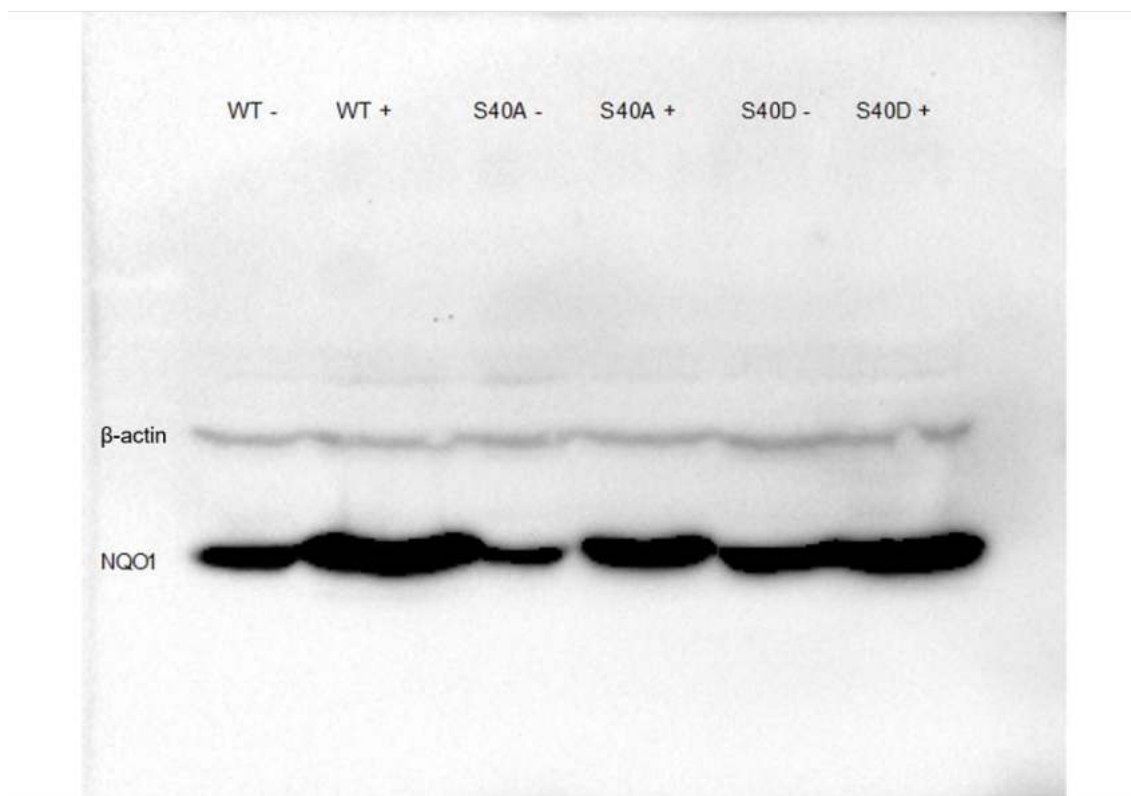
Figure S14. Full image of Western-blot analysis of HAP-1KO cells transfected with WT, S82A/D and T128A/D NQO1 variants in the absence (-) or presence (+) of the proteasomal inhibitor MG-132. The antibody used was anti-β-actin after using an anti-NQO1 antibody.
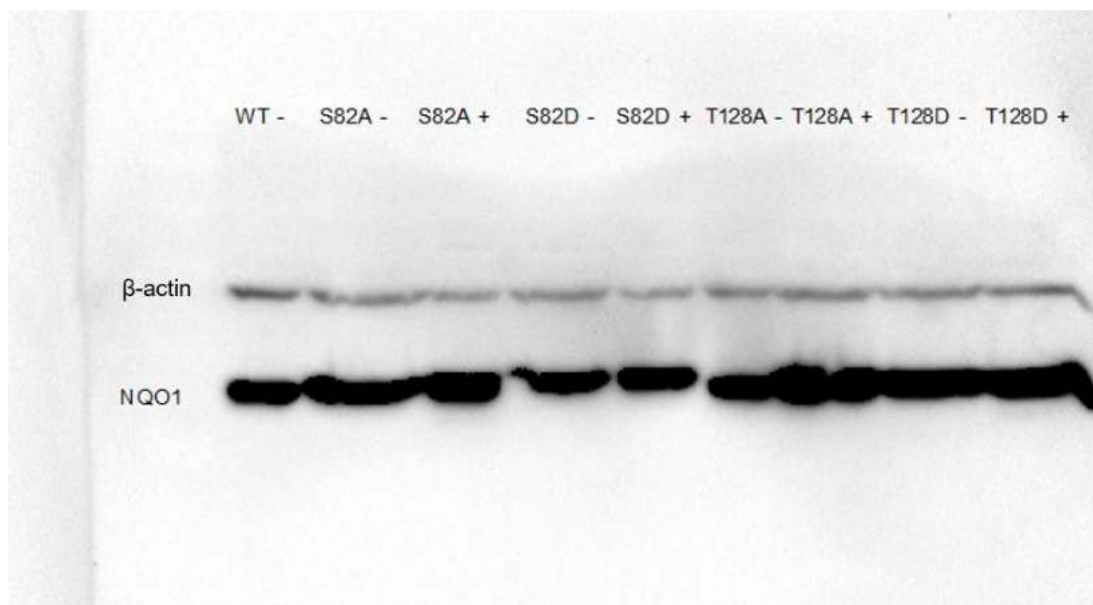
Figure S15. Full images of Western-blot analysis of HAP-1KO cells transfected with WT and P187S NQO1 variants in the absence (-) or presence (+) of the proteasomal inhibitor MG-132. The antibody used was anti-β-actin after using an anti-NQO1 antibody.
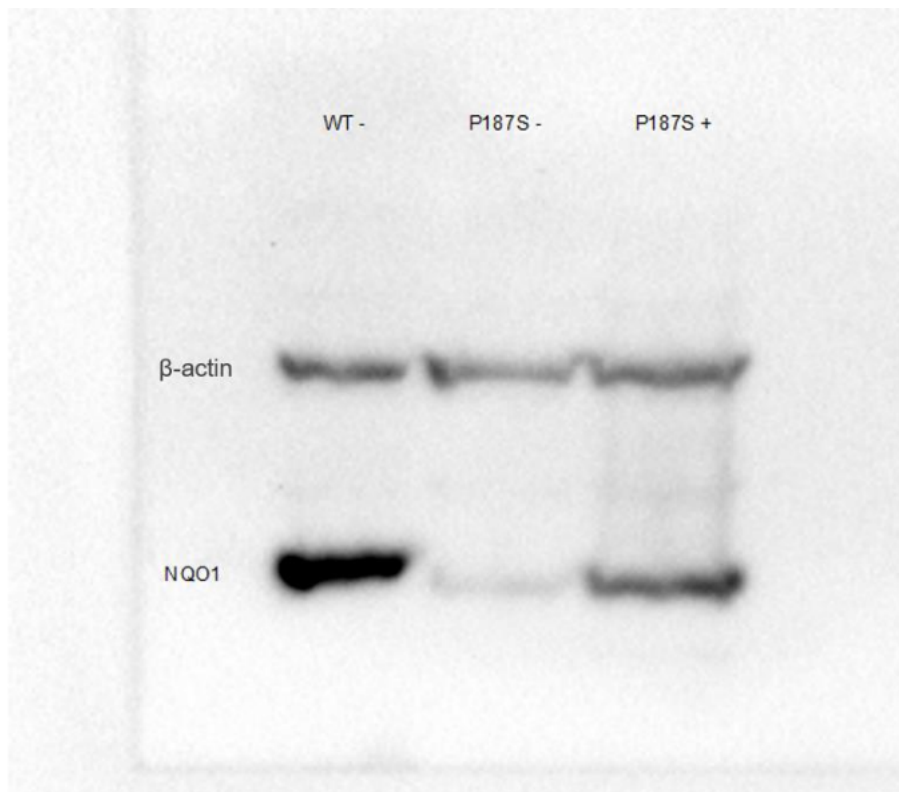
Figure S16. Full images of Western-blot analysis of HAP-1 KO cells transfected with WT, S40A/D, S82A/D and T128A/D when cultures were not exposed to MG-132. Panels A and B show the blots using anti-NQO1 and anti-β-actin, respectively. **Replica 1.**
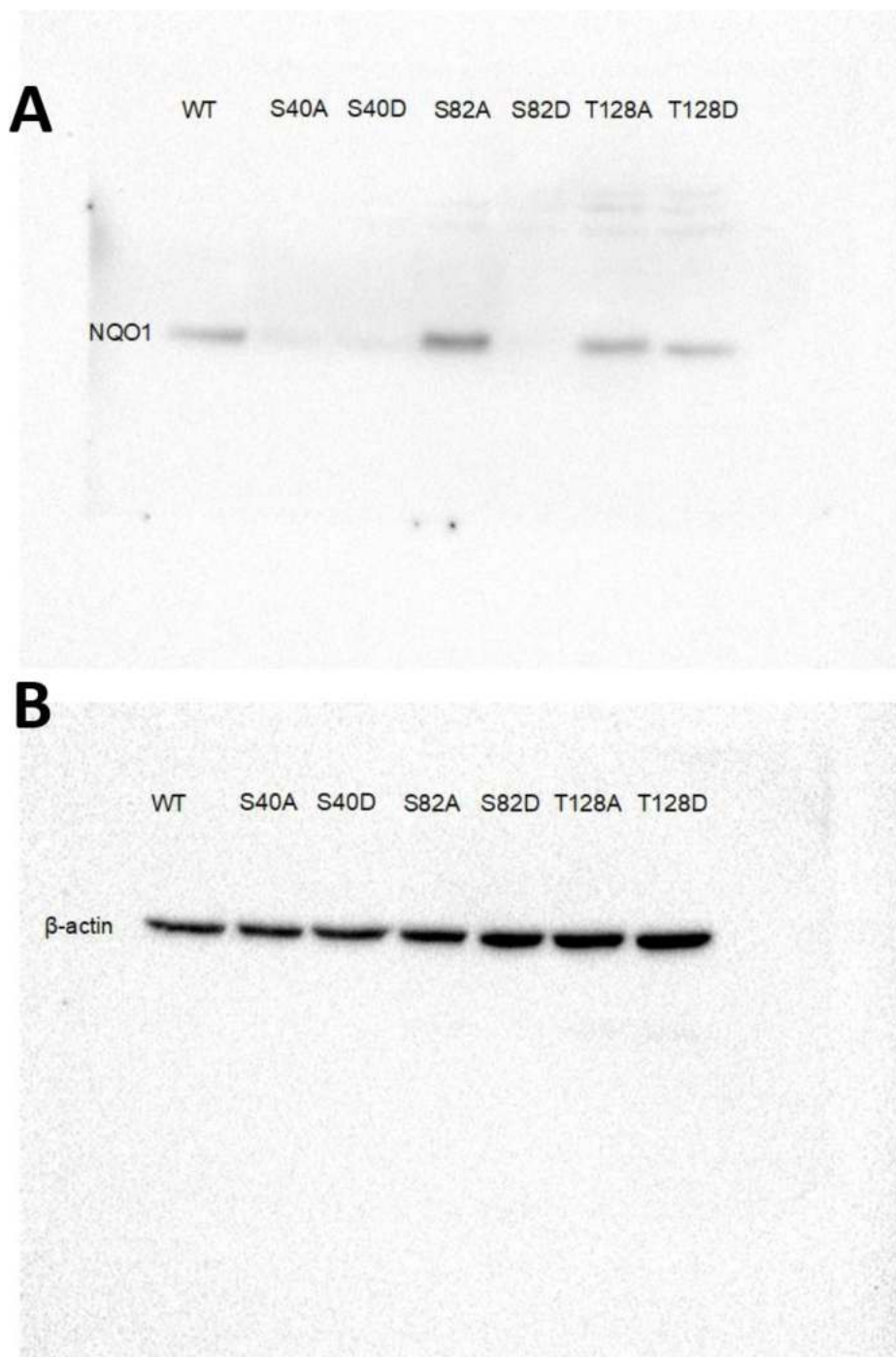
Figure S17. Full images of Western-blot analysis of HAP-1 KO cells transfected with WT, S40A/D, S82A/D and T128A/D when cultures were not exposed to MG-132. Panels A and B show the blots using anti-NQO1 and anti-β-actin, respectively. **Replica 2.**
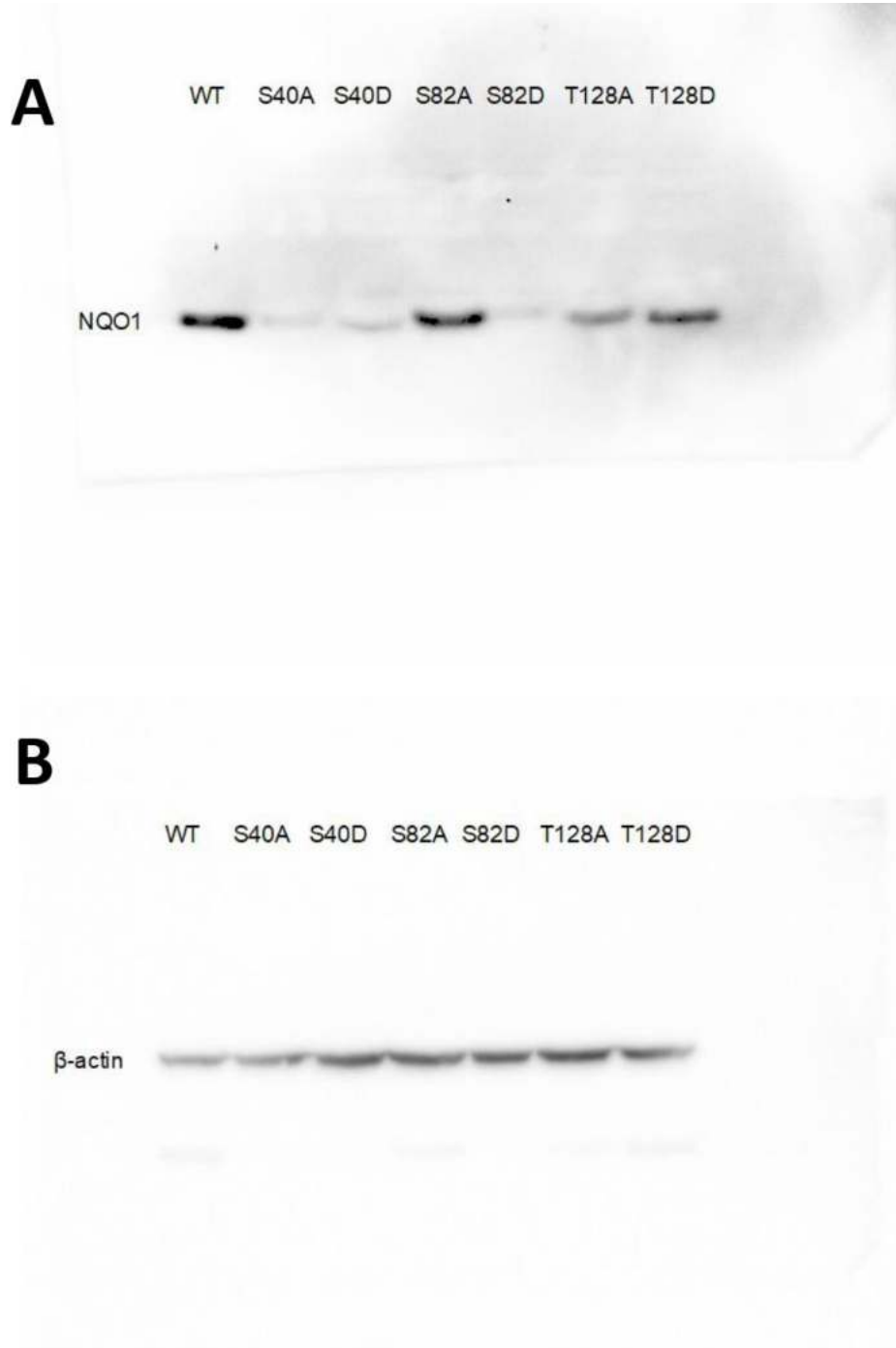
Figure S18. Full images of Western-blot analysis of HAP-1 KO cells transfected with WT, S40A/D, S82A/D and T128A/D when cultures were not exposed to MG-132. Panels A and B show the blots using anti-NQO1 and anti-β-actin, respectively. **Replica 3.**
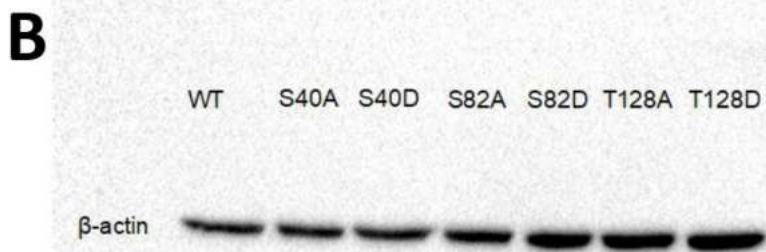
Table S1. Arrhenius parameters and KIEs for the HT/DT in the reduction of NQO1 variants by NADH/NADD. All values correspond to data obtained with equimolecular concentrations of the reactants in the stopped-flow equipment (n > 3, mean ± SD). Analyses were performed as previously described [5].

Step	HT			DT			KIE [a]	ΔE_{aDT-HT} (kcal·mol^{-1})	A_{HT}/A_{DT}
	k_{obsHT} [a] (s^{-1})	E_{aHT} (kcal·mol^{-1})	A_{HT} (s^{-1})	k_{obsDT} [a] (s^{-1})	E_{aDT} (kcal·mol^{-1})	A_{DT} (s^{-1})			
FAST									
WT[b]	78±1	6.1±0.2	(5.3±1.2)·10^6	44±2	6.3±0.4	(4.1±1.1)·10^6	1.8±0.1	0.2±0.1	1.3±0.6
S40A	90±2	7.0±0.1	(3.3±0.9)·10^7	52±1	7.5±0.1	(4.4±0.9)·10^7	1.8±0.1	0.5±0.1	(7±1)·10^{-1}
S40D	98±5	6.8±0.1	(2.3±0.8)·10^7	56±1	7.0±0.1	(1.9±0.5)·10^7	1.8±0.1	0.2±0.1	1.3±0.8
S82A	95±7	4.6±0.1	(4.0±0.9)·10^5	58±1	6.7±0.1	(1.0±0.4)·10^7	1.6±0.1	2.1±0.1	(4±2)·10^{-2}
S82D[b]	68±3	4.9±0.1	(4.9±1.0)·10^5	41±1	5.4±0.2	(6.4±3.3)·10^5	1.7±0.1	0.5±0.2	0.8±0.4
T128A	96±6	5.3±0.1	(1.4±0.9)·10^6	56±1	4.9±0.1	(3.8±1.0)·10^5	1.7±0.1	-0.5±0.1	3.8±0.7
T128D	82±2	4.1±0.1	(1.5±0.5)·10^5	31±2	4.4±0.1	(9.2±1.4)·10^4	2.6±0.1	0.3±0.1	1.6±0.7
SLOW									
WT[b]	8.9±0.9	10.9±0.5	(3.4±0.9)·10^9	5.3±0.2	9.8±0.5	(2.6±0.6)·10^8	1.8±0.3	-1.1±0.2	13±6
S40A	12±0.5	10.3±0.2	(1.6±0.5)·10^9	8.0±0.3	14.4±0.2	(1.6±0.5)·10^{12}	1.5±0.1	4.1±0.2	(10±6)·10^{-4}
S40D	12±2	11.7±0.1	(2.0±0.7)·10^{10}	9.5±0.2	11.7±0.2	(1.8±0.5)·10^{10}	1.3±0.1	0±0.1	1.3±0.9
S82A	12±1	8.0±0.1	(2.5±0.8)·10^7	8±0.3	15.1±0.1	(6.6±1.0)·10^{12}	1.5±0.2	7.1±0.2	(4±1)·10^{-6}
S82D[b]	2.0±0.8	13.0±0.1	(3.0±0.8)·10^{10}	1.4±0.3	17.7±0.3	(1.0±0.5)·10^{14}	1.4±0.6	4.7±0.3	(3±2)·10^{-4}
T128A	3.4±0.1	20.9±0.2	(1.1±0.3)·10^{17}	18±1	17.4±0.2	(2.0±0.5)·10^{14}	0.9±0.2	-3.5±0.1	(6±4)·10^2
T128D	22±1	4.9±0.1	(1.6±0.4)·10^5	13±2	3.8±0.1	(1.2±0.5)·10^4	1.7±0.2	-1.2±0.1	14±10

[a]Values at 6 °C. [b]Data from [5,6].

Supplementary references

[1]   E. Medina-Carmona, J.L. Neira, E. Salido, J.E. Fuchs, R. Palomino-Morales, D.J. Timson, A.L. Pey, Site-to-site interdomain communication may mediate different loss-of-function mechanisms in a cancer-associated NQO1 polymorphism, Scientific Reports. 7 (2017) 44532. https://doi.org/10.1038/srep44532.

[2]   M. Faig, M.A. Bianchet, P. Talalay, S. Chen, S. Winski, D. Ross, L.M. Amzel, Structures of recombinant human and mouse NAD(P)H:quinone oxidoreductases: species comparison and structural changes with substrate binding and release., Proc Natl Acad Sci U S A. 97 (2000) 3177–82. https://doi.org/10.1073/pnas.97.7.3177.

[3]   G. Asher, O. Dym, P. Tsvetkov, J. Adler, Y. Shaul, The crystal structure of NAD(P)H quinone oxidoreductase 1 in complex with its potent inhibitor dicoumarol., Biochemistry. 45 (2006) 6372–8. https://doi.org/10.1021/bi0600087.

[4]   P. Vankova, E. Salido, D.J. Timson, P. Man, A.L. Pey, A dynamic core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer, Biomolecules. 9 (2019) 728. https://doi.org/10.3390/biom9110728.

[5]   E. Anoz-Carbonell, D.J. Timson, A.L. Pey, M. Medina, The catalytic cycle of the antioxidant and cancer-associated human NQO1 enzyme: Hydride transfer, conformational dynamics and functional cooperativity, Antioxidants. 9 (2020) 1–22. https://doi.org/10.3390/antiox9090772.

[6]   J.L. Pacheco-Garcia, E. Anoz-Carbonell, P. Vankova, A. Kannan, R. Palomino-Morales, N. Mesa-Torres, E. Salido, P. Man, M. Medina, A.N. Naganathan, A.L. Pey, Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue, Redox Biology. 46 (2021) 102112. https://doi.org/10.1016/j.redox.2021.102112.

[7]   D. Siegel, A. Anwar, S.L. Winski, J.K. Kepa, K.L. Zolman, D. Ross, Rapid Polyubiquitination and Proteasomal Degradation of a Mutant Form of NAD(P)H:Quinone Oxidoreductase 1, Molecular Pharmacology. 59 (2001) 263-268. https://doi.org/10.1124/mol.59.2.263

[8]   A. Martínez-Limón, M. Alriquet, W.H. Lang, G. Calloni, I. Wittig, R.M. Vabulas, Recognition of enzymes lacking bound cofactor by Protein quality control, Proc Natl Acad Sci U S A. 113 (2016) 12156–12161. https://doi.org/10.1073/pnas.1611994113.

4. DISCUSIÓN

4.1 Las mutaciones naturales encontradas en NQO1 pueden afectar a su estabilidad y función

Actualmente, se conocen dos polimorfismos de NQO1 (P187S y R139W) que se han encontrado en líneas celulares de cáncer y que se han asociado con un mayor riesgo de desarrollar esta enfermedad[147,158]. Sin embargo, la variabilidad genética en NQO1 es mucho más que estas dos variantes relativamente comunes. Como se indicó anteriormente, se han identificado 163 mutaciones de cambio de sentido poco comunes en NQO1, y su efecto sobre la estabilidad y la función de la proteína no se conoce.

En esta tesis hemos llevado a cabo un estudio comparativo sobre el efecto en la estabilidad y función de NQO1 entre mutaciones de cambio de sentido procedentes de estudios de secuenciación masiva del exoma y del genoma humano (obtenidas de la base de datos gnomAD), que no están asociadas a cáncer, y de estudios de secuenciación de muestras de cáncer (obtenidas de la base de datos COSMIC). En estudios previos se ha observado que K240Q, una variante poco común de NQO1 detectada en tumores, produce un ligero impacto en la estabilidad de NQO1, lo que podría potenciar los mecanismos subyacentes en el desarrollo de la enfermedad debido a la pérdida de su función[155,187].

En primer lugar, analizamos un conjunto de mutantes localizados en el extremo N-terminal, cinco procedentes de gnomAD y tres de COSMIC (Figura 7A). Se seleccionaron las ocho primeras mutaciones por orden en la secuencia desde el extremo N-terminal. Estas mutaciones presentan distinto grado de conservación evolutiva y afectan a residuos pertenecientes a distintas regiones: G3 se encuentra en el primer bucle del extremo N-terminal, L7 y V9 se encuentran en una lámina β enterrada, T16 en un bucle expuesto al solvente y Y20 y A29 en una hélice α, estando Y20 expuesta al solvente y A29 enterrada en la estructura (Figura 7B). Según la predicción basada en su estructura, el impacto sobre la estabilidad de NQO1 entre estas mutaciones poco comunes sería bastante diferente.

Llevamos a cabo la caracterización de estas variantes en relación a su estabilidad y funcionalidad. Por un lado, estudiamos la estabilidad térmica, que muestra la estabilidad del dímero que se disocia previamente a la etapa limitante de desnaturalización[94,155]. Y

por otro, la estabilidad local en el extremo N-terminal por estudios de proteólisis limitada[188]. Además, comprobamos el efecto en su funcionalidad al evaluar la afinidad por el FAD, que es esencial para su función[119], y por el dicumarol, que aporta información sobre la estabilidad del extremo C-terminal[117], importante para la unión del sustrato NAD(P)H[119] y para la estabilidad intracelular[152,189]. Los resultados obtenidos muestran que la pérdida de estabilidad y/o función no está más asociada a las variantes procedentes de COSMIC que a las variantes procedentes de gnomAD (Publicación 1).
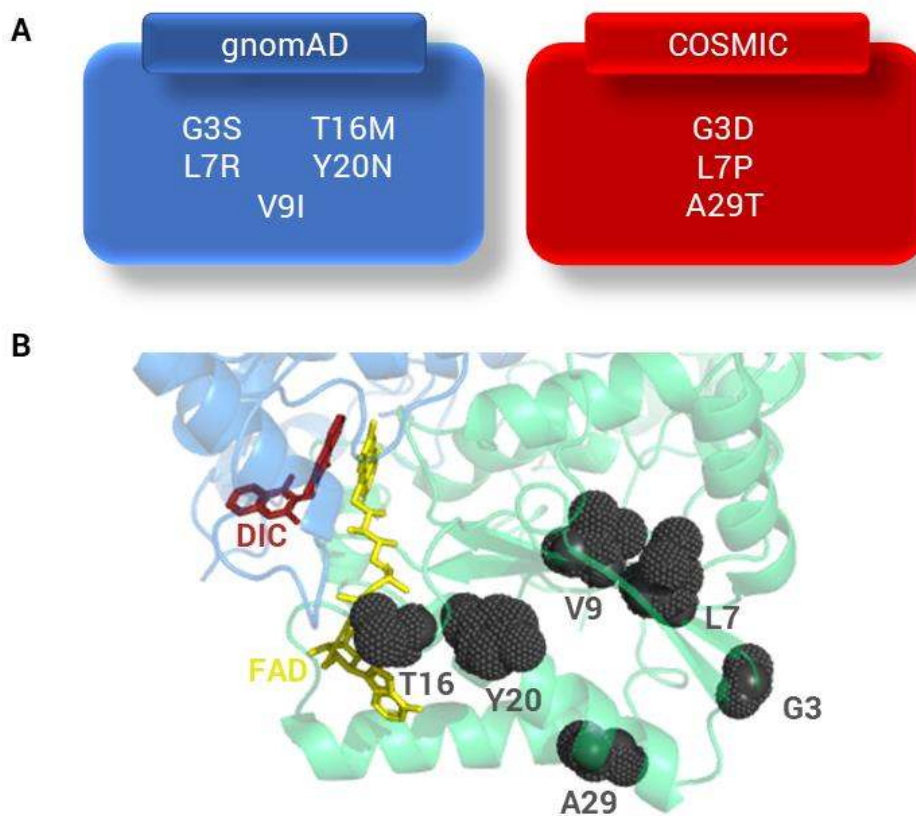
Figura 7. Representación de las mutaciones estudiadas en el extremo N-terminal. (A) Mutaciones obtenidas de cada base de datos. (B) Localización estructural de los residuos mutados en NQO1, así como de las moléculas de FAD y dicumarol (DIC) (PDB 2F1O)[117].

El efecto de las mutaciones G3S y G3D es prácticamente neutro, al igual que el de la mutación V9I, sin mostrar efectos significativos en su estabilidad y funcionalidad. A pesar de que G3S supone la introducción de un aminoácido polar en un ambiente hidrofóbico, y G3D de un aminoácido polar y cargado, los análisis de modelado

molecular con Rosetta[190] indican que hay una reacomodación en el extremo N-terminal que permite tolerar esas perturbaciones. V9I produce la introducción de un aminoácido más grande, pero según los análisis de modelado molecular esa alteración se ve compensada al favorecerse nuevas interacciones de Van Der Waals, dando como resultado un efecto neutro. Distintos estudios sugieren que una gran parte de las mutaciones de cambio de sentido son neutras o tienen un impacto leve en la función de la proteína[98,99]. En nuestro caso, podemos indicar que tienen ese efecto neutro sobre las propiedades que hemos estudiado. Sin embargo, es posible que estas mutaciones aún puedan involucrar algún efecto perjudicial, ya sea afectando a la actividad enzimática o a otras funciones de NQO1, como la interacción con otras biomoléculas[128], o desarrollando algún mecanismo de ganancia de función[187].

Los mutantes L7P y L7R disminuyeron fuertemente los niveles de proteína soluble de manera que no fue posible purificarlos. Esta leucina se encuentra enterrada en un ambiente hidrofóbico en el que introducir una prolina o una arginina supone una fuerte alteración de la estabilidad de esa región. Por un lado, la prolina es un aminoácido más pequeño que crea una cavidad en el interior de la proteína dificultando las interacciones hidrofóbicas de esa región, así como el puente de hidrógeno existente en la proteína WT entre dos hebras β de la que hoja plegada β en la que se encuentra esa leucina. Por otro lado, la arginina supone la introducción de un aminoácido cargado en un ambiente hidrofóbico, además de provocar impedimentos estéricos con los residuos cercanos al ser un aminoácido de mayor tamaño. L7P y L7R por lo tanto están perturbando el plegamiento de NQO1 y evitando la formación del dímero. Esto se ha observado en estudios anteriores con mutantes que afectan a residuos enterrados en la estructura de NQO1, donde la sustitución de la prolina 187 por aminoácidos cargados como una arginina (P187R) o un ácido glutámico (P187E), o por aminoácidos de mayor tamaño como la leucina (P187L) afectó fuertemente al plegamiento y estabilidad de NQO1[155]. Estos efectos tan drásticos se han observado del mismo modo en otras proteínas como la alanina:glioxilato aminotransferasa (AGT) o la transtiretina (TTR), donde mutaciones no conservativas en el interior de la proteína reducen fuertemente los niveles de proteína soluble[187].

De las mutaciones naturales que se pudieron purificar, los efectos más fuertes se observaron en las mutaciones T16M e Y20N. Ambas disminuyen la estabilidad térmica 4-5 ºC, posiblemente por la alteración de la interfaz monómero-monómero (MMI, del

inglés *monomer-monomer interface*) ya que esos residuos se encuentran muy próximos. De hecho, el residuo T16 forma parte de la propia interfaz monómero-monómero. El efecto desestabilizante de estas mutaciones se propaga a regiones alejadas como el sitio de corte de la termolisina (TCS, del inglés *thermolysin cleavage site*) (a 15-20 Å de distancia), localizado en el dominio N-terminal. Esta comunicación alostérica con el sitio de corte de termolisina muestra un aumento de la dinámica local que a menudo se correlaciona con una reducción de la afinidad por FAD[153]. Tanto T16M como Y20N y A29T presentan una afinidad más baja por FAD, especialmente T16M que tiene una afinidad 12 veces menor. Cabe destacar que la mutación A29T afecta al sitio de unión de FAD (FBS, del inglés *FAD binding site*), que se encuentra a 13 Å de distancia, sin afectar a la estabilidad local en torno al sitio de corte de la termolisina, igual que ocurre en algunas mutaciones estudiadas en el Anexo I, mostrando una desestabilización alostérica del FBS, pero no del TCS. Los estudios de modelado molecular indican que T16M, que introduce un aminoácido hidrofóbico en lugar de uno polar, lleva a la recolocación de las cadenas laterales de algunos residuos que desplaza el sitio de unión de FAD, mientras que Y20N, que supone un cambio por un aminoácido más pequeño, evita la formación de puentes de hidrógeno entre esos residuos y otros cercanos al sitio mutado que llevan a la reagrupación de las cadenas laterales cercanas alterando el modo de unión del FAD. Por el contrario, A29T introduce un aminoácido polar que posiblemente forma un nuevo puente de hidrógeno dentro de la hélice α en la que se encuentra, y el efecto se propaga afectando al sitio de unión del FAD. Esta propagación de los efectos mutacionales a larga distancia que provoca la fuerte desestabilización del FBS y TCS ha sido descrita anteriormente en NQO1. El polimorfismo P187S causa efectos pleiotrópicos sobre la estabilidad y función de NQO1 perturbando la dinámica del sitio unión del FAD y su actividad enzimática[154], que lleva a un drástico descenso en la afinidad por FAD[94,152]. Efectos similares se han observado en la mutación fosfomimética S82D, donde se ha demostrado un alteración de la dinámica del dominio N-terminal y un fuerte descenso en la afinidad por FAD[183]. Esta propagación de efectos mutacionales a larga distancia sobre la estructura de NQO1 se profundiza en la sección 4.2.

A raíz de los efectos alostéricos observados en algunas de las mutaciones naturales decidimos estudiar la comunicación alostérica en NQO1 (Anexo I). Una manera de estudiar ese alosterismo consiste en utilizar mutaciones que crean cavidades en el

interior de la proteína en distintas posiciones y ver su efecto. Dado el efecto tan drástico de las mutaciones L7P y L7R, decidimos estudiar la comunicación alostérica que existe entre esta zona y el resto de la proteína. Para ello, sustituimos la leucina por aminoácidos más pequeños como alanina, valina y glicina (L7A, L7V y L7G). L7G es una mutación muy disruptiva que impide el plegamiento de la proteína o su solubilidad, al igual que las mutaciones L7P y L7R. En cuanto a las variantes L7A y L7V, introducen una alteración parecida a L7P al introducir un residuo más pequeño que la leucina, pero su efecto en la solubilidad no fue tan drástico, lo que nos permitió purificarlas y estudiar su efecto en toda la proteína. Ambas mutaciones disminuyen la estabilidad térmica, es decir, su temperatura de desnaturalización (Tm) unos 5-6 ºC. Este descenso de la estabilidad térmica está relacionado a una menor estabilidad del dímero, cuya disociación ocurre previamente a la etapa limitante de la desnaturalización de NQO1[94,155,187]. Esta desestabilización puede suponer que las mutaciones en L7 favorezcan la población de estados parcialmente plegados que son más fácilmente reconocidos por el proteosoma 20S[139]. Por el contrario, ningún mutante tuvo un efecto notorio en la afinidad por FAD o en la actividad catalítica. El estudio de la estabilidad estructural por intercambio hidrógeno-deuterio acoplado a espectrometría de masas (HDX-MS, del inglés *Hydrogen-Deuterium Exchange Mass Spectrometry*) permite determinar cambios en la conformación y dinámica de una proteína por el efecto de mutaciones que provocan la sustitución de un aminoácido por otro, así como por la unión de algún ligando[191]. Los datos de HDX-MS mostraron que estas mutaciones desestabilizan principalmente la forma apo de la proteína. Aunque el efecto es mayoritariamente local, se propaga levemente a zonas más alejadas como el sitio activo, lo que explicaría el ligero cambio observado en la cinética enzimática. La comunicación alostérica se estudió también entre otros residuos enterrados (L10 y L30) y el resto de la proteína, mostrando un comportamiento similar al observado en L7, pero de mayor magnitud, sobre todo en términos de solubilidad, afinidad por FAD y estabilidad de la proteína en su forma apo. Estos resultados ayudan a mejorar nuestra comprensión sobre la propagación de efectos mutacionales en esta flavoproteína y servir como base para deducir las consecuencias fenotípicas de aquellas mutaciones que afecten de forma alostérica a esas regiones.

Por otra parte, estudiamos un conjunto de mutantes naturales localizados en el sitio activo de NQO1 (Figura 8), que previsiblemente podrían tener un gran impacto sobre la

estabilidad estructural y la funcionalidad de la proteína (Publicación 2). La alteración de la función proteica por mutaciones naturales que ocurren en el sitio activo de la proteína[192–194] o regiones muy cercanas[195] está descrita para muchas proteínas. En ocasiones, esas mutaciones no eliminan la reacción bioquímica al completo, pero sí alteran la cinética de la reacción[196].
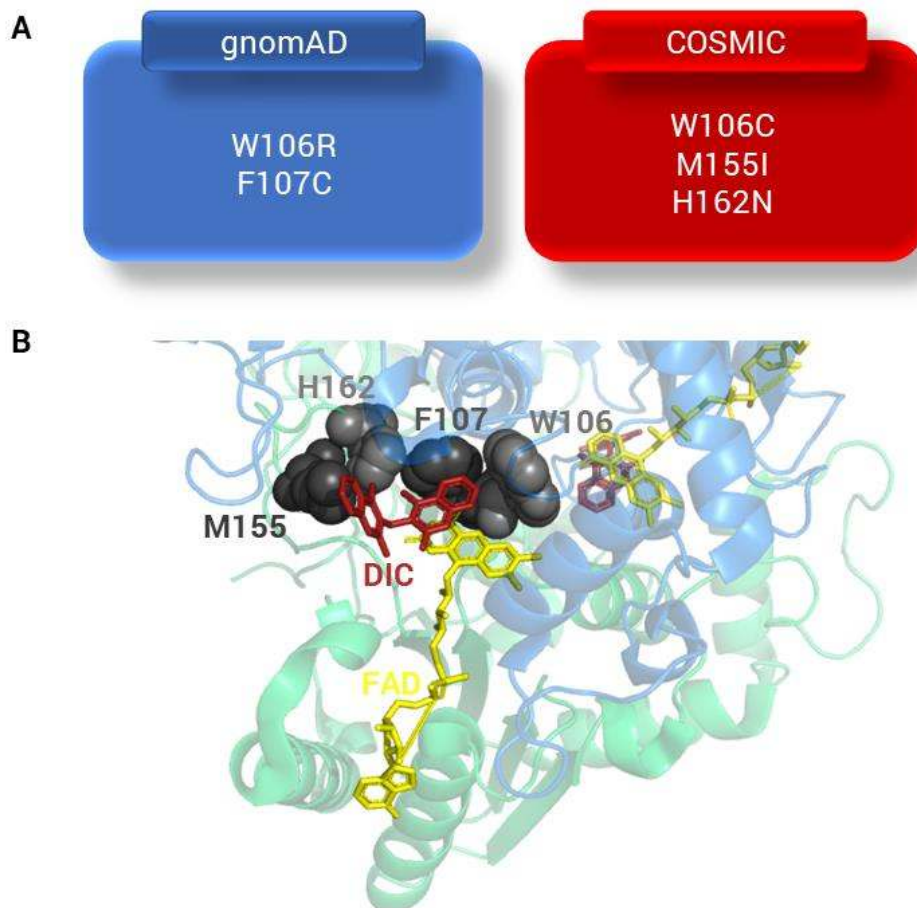
Figura 8. Representación de las mutaciones estudiadas en el sitio activo. (A) Mutaciones obtenidas de cada base de datos. (B) Localización estructural de los residuos mutados en NQO1, así como de las moléculas de FAD y dicumarol (DIC) (PDB 2F1O)[117].

La caracterización experimental de las mutaciones naturales en el sitio activo de NQO1 muestra efectos muy distintos, y al igual que en el conjunto de mutantes del extremo N-terminal, su efecto sobre la estabilidad y función de NQO1, es independiente de las muestras en las que se han encontrado.

Las variantes W106R y W106C afectan al triptófano 106 que está enterrado en la estructura de NQO1 y participa tanto en la interfaz entre los dos monómeros como en el

sitio de unión del FAD y del dicumarol. Únicamente el mutante W106R mostró una reducción en la afinidad por el FAD, aunque ambos mostraban un contenido en FAD menor que la proteína WT tras la purificación. La afinidad de W106R por FAD es 500 veces menor que la de la proteína WT. Esta es la afinidad más baja por FAD observada hasta ahora por una mutación de cambio de sentido en NQO1, muy inferior a la afinidad del polimorfismo P187S que es 10-40 veces inferior que la proteína WT[94,152]. Resulta interesante que W106C reduzca el contenido en FAD sin afectar a la afinidad, ya que el descenso en el contenido de FAD por el efecto de una mutación en NQO1, suele estar asociado a una menor afinidad por el FAD en esa variante[94,155,156,183]. El estudio de la estabilidad estructural por HDX-MS mostró que W106C causa una desestabilización leve en toda la proteína, mientras que W106R presenta un efecto desestabilizante mayor en distintos sitios funcionales: MMI, FBS y el sitio de unión del dicumarol (DBS, del inglés *dicoumarol binding site*). Estudios previos han demostrado que la unión de dicumarol a NQO1 provoca una estabilización estructural del DBS, que se propaga a regiones más alejadas estabilizando la MMI o el FBS[153,154,157], y se ha sugerido la existencia de una red de residuos que interactúan y conectan estructural y energéticamente el FBS y el DBS a través de la MMI[157]. Dado que la desestabilización causada por W106R se mantiene, e incluso es más fuerte cuando el dicumarol está presente, podemos sugerir que W106R altera la comunicación entre esos residuos provocando una pérdida de esa comunicación alostérica. Esta diferencia en la desestabilización causada por ambas mutaciones explica el efecto previamente descrito sobre la afinidad por el FAD, así como la alteración de su actividad catalítica. Mientras que W106R reduce 6000 veces la velocidad de la reacción de reducción del FAD, W106C sólo reduce levemente la velocidad de la reacción de oxidación.

El mutante F107C afecta a un residuo que se encuentra también enterrado en la estructura de la proteína formando parte de la MMI, del FBS y del DBS. En contra de lo esperado, F107C no redujo la afinidad por FAD. Resulta llamativo que no afecte a la afinidad a pesar de ser una mutación no conservativa que cambia tanto la polaridad como el tamaño[197] de un residuo que forma parte del FBS[117]. De hecho, el análisis por HDX-MS mostró solamente una leve desestabilización del FBS y del DBS que desaparecía en presencia del dicumarol. No obstante, el análisis de las cinéticas enzimáticas mostró que tanto la reacción de reducción como la de oxidación se ven

afectadas reduciendo la eficiencia catalítica considerablemente, posiblemente por una alteración de la dinámica del mecanismo de reacción.

Al contrario que en el caso de F107C, la mutación M155I (forma parte de la MMI, del DBS y está cerca del FBS), que parece ser poco disruptiva, mostró una fuerte desestabilización en NQO1, afectando tanto a la MMI, el FBS y el DBS. El efecto de esta mutación se traduce en una drástica reducción de la afinidad por el FAD, siendo 50 veces inferior a la de la proteína WT. Además, esta desestabilización se traduce en un gran efecto en la eficiencia catalítica de NQO1, afectando fuertemente a la reacción de reducción tanto en la etapa rápida como lenta, y evitando que se pueda estudiar la reacción de oxidación del FAD al no poder alcanzar el estado de NQO1 completamente reducido.

Por último, la variante H162N, que pertenece a la MMI, el FBS y el DBS, mostró un comportamiento parecido a M155I. A pesar de que la predicción parecía indicar un efecto leve, experimentalmente se observó una desestabilización estructural en los principales sitios funcionales de la proteína, reduciendo la afinidad por el FAD 30 veces. Por el contrario, las cinéticas enzimáticas apenas se ven alteradas y no afecta a la eficiencia catalítica de la enzima.

En general, hemos observado que tanto las mutaciones obtenidas de gnomAD como de COSMIC pueden tener efectos muy deletéreos en NQO1. La reducción en el contenido y afinidad por FAD observada en distintas variantes, como T16M o W106R, sugiere que la proteína puede encontrarse en su forma apo intracelularmente[91]. Este estado disminuye la estabilidad intracelular de NQO1, favoreciendo su degradación por el proteosoma[91,139]. Esos efectos se asemejan a los observados en P187S, que presenta niveles de proteína celulares muy bajos y una actividad enzimática prácticamente nula[149,198], en parte debido a su baja afinidad por el FAD[94,152]. Como se ha sugerido su posible asociación a una mayor susceptibilidad de desarrollar distintos tipos de cáncer[147], podríamos pensar que estas mutaciones, o aquellas que conlleven efectos similares, también podrían incrementar el riesgo de desarrollar algún tipo de cáncer, o alguna de las enfermedades con las que NQO1 está relacionada[163].

4.2 La mutación fosfomimética S82D desestabiliza NQO1 de manera similar a la mutación de cambio de sentido P187S

Las mutaciones de cambio de sentido[94,153] y las modificaciones postraduccionales[199,200] pueden afectar a la naturaleza multifuncional de las flavoproteínas humanas. Entre esas modificaciones postraduccionales, se ha visto que la fosforilación de una proteína puede tener importantes consecuencias sobre su estructura y función[55]. Sin embargo, el efecto de la fosforilación específica en cada sitio está muy poco estudiado. Los estudios existentes muestran que la fosforilación en diferentes residuos de una proteína puede tener efectos distintos[201]. Por ejemplo, la fosforilación específica de distintas serinas de un dominio del receptor de glucocorticoides provoca cambios diferentes en la estructura local[202]. Del mismo modo, la fosforilación específica de serinas de la proteína C de unión a la miosina cardíaca tiene efectos diferentes sobre su función[203].

Apoyándonos en los análisis de estabilidad basados en la estructura de distintas flavoproteínas realizados con FoldX[204], parece ser que las mutaciones asociadas a enfermedad y la fosforilación puede que desestabilicen la estructura proteica de una forma parecida, sobre todo cuando afectan a residuos enterrados (Material Suplementario de Publicación 3). Para estudiar el efecto de la fosforilación específica de sitio elegimos el residuo S82 de NQO1, ya que su fosforilación intracelular está validada por 13 estudios proteómicos diferentes y está bien conservada entre mamíferos. Para ello, utilizamos la mutación fosfomimética S82D, que introduce una carga negativa que imita al grupo fosfato, y comparamos el efecto estructural y funcional frente al polimorfismo asociado a enfermedad P187S (Figura 9) (Publicación 3). Aunque ya se había visto anteriormente como afectan P187S[94,153,154] y S82D[183] a algunas propiedades de NQO1, el estudio por HDX-MS nos permite determinar cómo esas variantes cambian en la estabilidad y dinámica en cada región de una proteína, y como ese efecto se propaga por la estructura de la proteína[191]. Los análisis de HDX-MS y mecánico estadísticos de NQO1 indican que la proteína adopta diferentes subestados conformacionales[157] que pueden tener un papel importante en funciones como la unión de FAD o la degradación proteosomal[91,153]. Las cinéticas de intercambio de hidrogeno por deuterio muestran un núcleo estable en la proteína apo que mantiene el dímero unido[157]. Este núcleo se ve afectado por las mutaciones S82D y P187S, siendo la desestabilización causada por P187S más fuerte que la de S82D. Esta desestabilización de la MMI en P187S ya se había sugerido previamente, al observar el descenso de su

estabilidad térmica[94], ya que la disociación del dímero ocurre previamente a la etapa limitante de la desnaturalización de NQO1[94,155,187]. En cambio, el efecto desestabilizante de S82D en la MMI no se traduce en un descenso de su estabilidad térmica (de hecho, aumenta ligeramente)[183], posiblemente por ser una desestabilización más leve y que afecta a menos residuos. La desestabilización causada por P187S y S82D se propaga a distintas regiones. Los estudios previos de P187S indican que este polimorfismo provoca cambios locales en la estructura y dinámica en torno al FBS en la apoproteína[156] asi como en el extremo C-terminal[153,154]. Esta desestabilización en torno al FBS lleva a que P187S presente una afinidad por FAD 10-40 veces inferior[94,152]. Los análisis de HDX-MS corroboran la fuerte desestabilización de P187S como apoproteína tanto en el extremo N-terminal como en el C-terminal. La unión de FAD reduce la desestabilización observada en el domino N-terminal de P187S, pero sigue manteniendo una fuerte desestabilización en el domino C-terminal y la MMI. Por otra parte, se ha observado que S82D aumenta la dinámica del dominio N-terminal disminuyendo la afinidad por FAD 20 veces[183]. A diferencia de P187S, la perturbación parece ser que no se propaga al domino C-terminal[183]. El estudio por HDX-MS muestra una desestabilización global de S82D como apoproteína similar a la de P187S, pero de menor intensidad. El efecto observado en el dominio N-terminal, afectando al FBS explica el descenso en la afinidad por FAD descrito para S82D[183]. La unión de FAD a S82D tiene efectos diferentes a los de P187S. Por una parte, disminuye la desestabilización de la MMI y estabiliza fuertemente el extremo C-terminal. Sin embargo, la desestabilización en el dominio N-terminal se mantiene y es más fuerte.
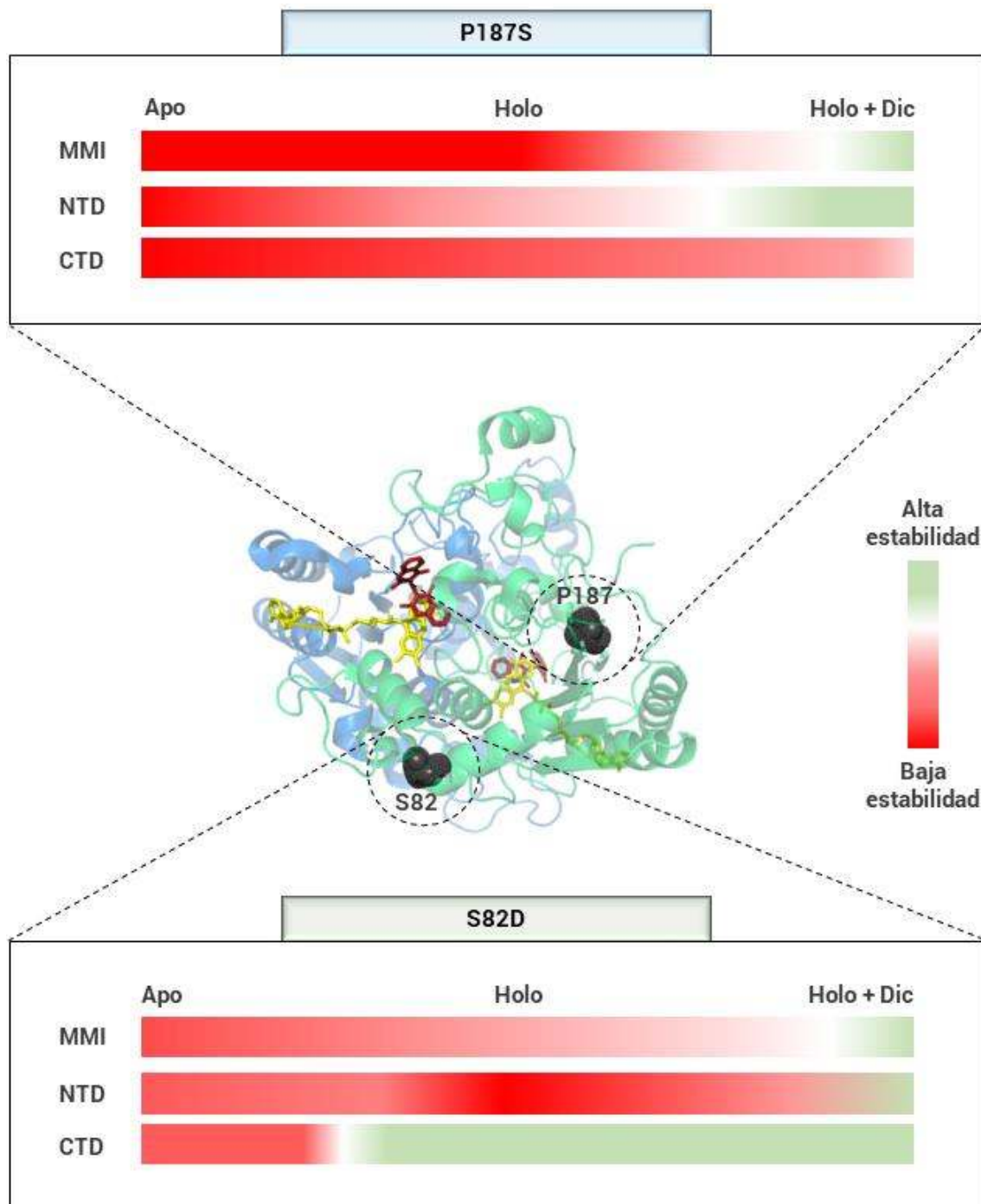
Figura 9. Efecto del polimorfismo P187S y la mutación fosfomimética S82D sobre la estabilidad estructural de NQO1. El efecto se muestra en distintas regiones (MMI, interfaz monómero-monómero; NTD, dominio N-terminal; CTD, dominio C-terminal) y en distintos estados de ligación (Apo, proteína sin FAD; Holo, proteína más FAD; Holo+Dic, proteína más FAD y dicumarol). Las mutaciones están representadas en la estructura cristalina de NQO1 (PDB 2F1O)[117].

La desestabilización del domino C-terminal tiene implicaciones muy importantes para NQO1 debido a su papel en la unión del NAD(P)H y en la actividad catalítica[153], así como para su estabilidad intracelular[152,153,189]. Estudios previos han mostrado que NQO1 en su forma apo tiene un dominio C-terminal muy dinámico que se estabiliza tras la unión de FAD[152,153]. Sin embargo, en P187S el domino C-terminal sigue siendo muy flexible incluso como holoproteína reduciendo su actividad catalítica[152,153]. Además, la desestabilización termodinámica del dominio C-terminal por P187S lleva a una reducción en su afinidad por el dicumarol de 6-10 veces[154,155]. En base a la diferente estabilidad del dominio C-terminal observada entre P187S y S82D, es de esperar que presenten diferencias tanto en afinidad por el dicumarol, así como en los cambios de estabilidad estructural tras la unión del dicumarol. De hecho, la unión de dicumarol a S82D tuvo efectos muy leves en la estabilidad estructural observada por HDX-MS, al igual que ocurre en la proteína WT. En cambio, la unión de dicumarol a P187S contrarrestó la desestabilización estructural observada por toda la proteína, menos en el dominio C-terminal donde los efectos se redujeron, pero aún eran evidentes. La estabilización estructural observada en el resto de la proteína tras la unión de dicumarol apoya una vez más la comunicación descrita entre los distintos dominios de NQO1[154]. La diferencia de estabilidad del dominio C-terminal entre S82D y P187S coincide con los datos obtenido por proteólisis limitada con tripsina que muestran un dominio C-terminal moderadamente desestabilizado en el caso de S82D en su forma holo y fuertemente desestabilizado para P187S, que en ambos casos se estabilizaba bastante tras la unión de dicumarol. Además, esta diferencia de estabilidad del dominio C-terminal se refleja en la reducida afinidad por dicumarol de P187S frente a S82D observada los ensayos de calorimetría de titulación isotérmica.

La desestabilización estructural observada en el dominio N-terminal y en el domino C-terminal sugieren que S82D, y sobre todo P187S, pueden afectar a la actividad enzimática, ya que ambos participan en el sitio activo, siendo esenciales para a unión del FAD, del NAD(P)H y de los sustratos[117,119]. Al estudiar la reacción de reducción, observamos que estas mutaciones no afectaban a la velocidad de la reacción de reducción de la primera flavina (etapa rápida), pero disminuían considerablemente la velocidad de reducción de la segunda flavina (etapa lenta), sobre todo P187S. Por lo tanto, ambas mutaciones disminuyen la eficiencia catalítica de la proteína, potenciando la cooperatividad negativa que existe entre los dos sitios activos de NQO1[118], siendo el

caso de P187S mucho más fuerte, ya que uno de los sitios activos queda prácticamente inactivo.

Como se ha mencionado anteriormente, la estabilidad del dominio C-terminal es clave para la estabilidad intracelular de NQO1. La flexibilidad de este dominio en P187S favorece su poliubiquitinación y rápida degradación por el proteosoma[152,189], que se ve favorecida además por el descenso en afinidad por FAD[94,152]. La reducida afinidad por FAD, consecuencia de la desestabilización en torno al FBS, favorece su estado como apoproteína intracelularmente que lleva a su reconocimiento y degradación proteosomal[90,91]. Por ello, la suplementación con riboflavina (precursor del FAD) produce un leve incremento de los niveles de proteína[153]. Este descenso en la estabilidad intracelular también se ha observado en S82D[183], y al igual que ocurre en P187S hemos demostrado que el descenso en los niveles de proteína intracelulares se debe a una mayor degradación proteosomal. Dado que S82D desestabiliza estructuralmente el domino N-terminal en torno al FBS, reduciendo la afinidad por FAD, esta mutación está favoreciendo el estado apo en el interior de la célula que es más fácilmente reconocido por el proteosoma[90,91]. Como hemos visto, la desestabilización estructural es más fuerte en P187S que en S82D, y ese efecto está directamente relacionado con los niveles de proteína: 5% vs. 15-20% respecto a los de NQO1 WT. Como la inhibición del proteosoma muestra un mayor aumento de los niveles de P187S frente a los S82D, podemos considerar que esa diferencia se debe a un mayor reconocimiento de P187S frente a S82D por el proteosoma.

Previamente, se ha mostrado que la suplementación con dicumarol es capaz de estabilizar a P187S evitando su degradación proteosomal al disminuir la flexibilidad del dominio C-terminal[153,154]. Puesto que la unión de dicumarol a S82D también produce cierta estabilización del dominio C-terminal, aunque con un efecto menor que P187S, resulta coherente el aumento de los niveles de proteína de S82D tras la tras la suplementación con dicumarol, siendo menos notable que en P187S.

Tanto la estabilización ejercida por el FAD como por el dicumarol, sugieren que podrían utilizarse como chaperonas farmacológicas en NQO1 en una posible terapia frente a polimorfismos como P187S, por ejemplo, con la suplementación en la dieta de riboflavina (precursor del FAD)[120]. De hecho, la suplementación con riboflavina de la línea celular tumoral Caco-2 (homocigota para P187S), que presenta niveles de proteína muy bajos, produce un aumento de estos niveles[153]. Sin embargo, en el caso del

dicumarol, cuyo efecto anticoagulante[205] y desacoplante de la cadena de transporte de electrones[206] impide su aplicación directa, una estrategia interesante es desarrollar análogos que ejerzan un efecto equivalente, pero que carezcan de efectos secundarios[207]. De hecho, ya se han desarrollado compuestos basados en dicumarol que han reducido esos efectos adversos[208], pero actualmente no existe ninguna terapia basada en análogos del dicumarol.

4.3 La "fosforilación" específica en distintos sitios de NQO1 tiene consecuencias fenotípicas diferentes

Los efectos de la fosforilación específica de sitio en NQO1 están poco estudiados. De los 12 sitios de fosforilación que se han identificado, solamente existen estudios sobre 3 de ellos. El efecto de la fosforilación de S82 ha sido descrito en la sección 4.2, así como en estudios anteriores[183], mediante el uso de mutaciones fosfomiméticas. Por otra parte, otros estudios han mostrado el efecto intracelular de la fosforilación directa de S40 y T128, que desencadenan la ubiquitinación y degradación de NQO1[184]. Sin embargo, se desconocen las consecuencias estructurales y funcionales de estos eventos de fosforilación.

Siguiendo la estrategia utilizada en la sección anterior, hemos comparado el efecto estructural y funcional de la fosforilación en esos residuos utilizando las mutaciones fosfomiméticas S40D, S82D y T128D y las mutaciones no fosforilables S40A, S82A y T128A (Publicación 4). En general, observamos que las mutaciones fosfomiméticas afectan de distinta manera a la estructura, estabilidad y función de NQO1 y que los efectos eran independientes de lo enterrados que están esos residuos (% ASA: S40 < S82 < T128) (Figura 10). Por una parte, las tres mutaciones redujeron la afinidad por FAD, siendo el efecto de S82D mucho más fuerte. Mientras que S82D presenta una afinidad 30 veces menor, la afinidad de S40D y T128D es 6 y 2 veces más baja, respectivamente. Tanto los resultados obtenidos de los experimentos de proteólisis con termolisina como de HDX, muestran en S82D una desestabilización mucho mayor del extremo N-terminal, y concretamente del FBS, frente a la leve desestabilización observada en S40D y T128D, que coincide con el efecto observado en la afinidad por FAD. En todos los casos, el efecto se está propagando al FBS, una región alejada del

sitio mutado, como se ha descrito en otros mutantes de NQO1 analizados en las secciones 4.1 y 4.2.

Además de la desestabilización del FBS, los análisis de HDX mostraron que S40D y S82D desestabilizan la MMI en ausencia de FAD unido (como apoproteína), y en S82D este efecto se mantiene cuando tiene FAD unido. En cambio, T128D parece no afectar a la estabilidad de esta región. Eso explica el efecto observado sobre la estabilidad térmica de NQO1.Como se ha indicado anteriormente, la estabilidad térmica se asocia con la estabilidad de la MMI, ya que la disociación del dímero ocurre previamente a la etapa limitante de desnaturalización[94,155,187]. Por ello, la desestabilización observada por HDX-MS en la MMI se traduce en que S82D y S40D reducen la Tm alrededor de 7 ºC, mientras que T128D no tiene un efecto apreciable.

	S40D	S82D	T128D
Estabilidad MMI	↓	↓↓	=
Estabilidad térmica	↓	↓	=
Estabilidad FBS	↓↓	↓↓	↓
Afinidad por FAD	↓	↓↓	↓
Reducción del FAD			
Etapa rápida	=	=	↓
Etapa lenta	=	↓	=
Estabilidad intracelular	↓↓	↓↓	=
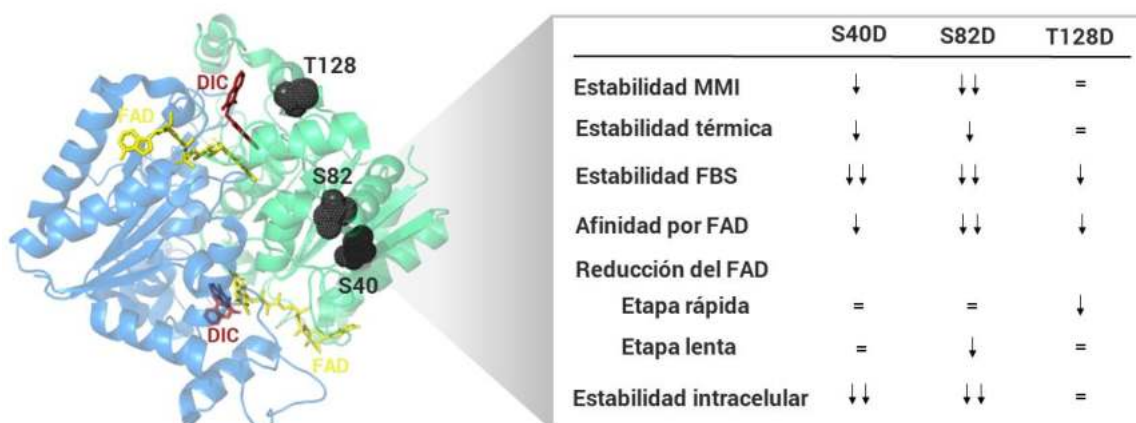
Figura 10. Sitios de fosforilación estudiados en la proteína NQO1 humana. Localización estructural de los sitios de fosforilación estudiados en este trabajo (PDB 2F1O)[117] y resumen de los efectos observados usando mutaciones fosfomiméticas. Los efectos pueden ser neutros (=), levemente perjudiciales (↓) o fuertemente perjudiciales (↓↓) en comparación a la proteína WT.

En cuanto a la actividad catalítica, en estudios anteriores se ha demostrado que la NQO1 WT y las variantes que no pueden ser fosforiladas tienen mayor actividad enzimática que las variantes fosforiladas[183,184] proporcionando un mayor efecto citoprotector[184]. En nuestro caso, observamos que la etapa rápida de la reacción de reducción del FAD solo se vio afecta por T128D que redujo la afinidad por NADH y la eficiencia catalítica, mientras que la etapa lenta sólo se vio afectada por S82D como habíamos discutido en la sección anterior. Por lo tanto, estos resultados demuestran que las mutaciones fosfomiméticas afectan de forma diferente a la cooperatividad funcional de NQO1. Esta

diferencia en el efecto sobre la actividad catalítica de la proteína, se observa también en la distinta estabilidad estructural del sitio activo que muestran ambas mutaciones a través de los resultados de HDX.

La estabilidad intracelular se ve afectada de forma específica en cada sitio de fosforilación. Como se ha mostrado anteriormente, la mutación S82D provoca una disminución en los niveles intracelulares de proteína del 80% respecto a la proteína WT. Una disminución similar (30-40%) se observa en el caso de S40D, aunque esta reducción también se observa en la mutación S40A, lo que sugiere que incluso pequeños cambios en este sitio disminuyen la estabilidad intracelular de NQO1. Sin embargo, T128D no tuvo ninguna repercusión en los niveles intracelulares de proteína. Cabe destacar que S40D, y sobre todo S82D, tienen el mayor efecto sobre la afinidad por FAD. Por lo tanto, en ausencia de suplementación con el cofactor (o su precursor riboflavina), ambas mutaciones posiblemente están favoreciendo que la proteína se encuentre en su forma apo. En ausencia de cofactor se ha demostrado que las flavoproteínas son más inestables y son dirigidas a su degradación por el proteosoma[91]. De hecho, la inhibición del proteosoma supuso un incremento de los niveles de proteína de estas variantes, corroborando esa degradación proteosomal acelerada.

Aunque el efecto de T128D parece ser menor, el estudio de la fosforilación directa de ese residuo parece indicar que la fosforilación de T128 desencadena la fosforilación de S40[184]. Dado que ambos eventos de fosforilación estarían coocurriendo sobre NQO1, sería interesante estudiar el impacto estructural y funcional de su efecto mediante mutantes fosfomiméticos que incluyan las mutaciones S40D y T128D simultáneamente.

5. CONCLUSIONES

A partir de los resultados obtenidos, las conclusiones más relevantes extraídas de esta tesis son las siguientes:

1. Las variantes naturales de NQO1 encontradas en la población tanto en muestras de pacientes sanos como en muestras de cáncer pueden afectar a la estabilidad y función de NQO1.

2. Los efectos desestabilizantes de las mutaciones localizadas en el extremo N-terminal pueden propagarse a regiones alejadas como la interfaz monómero-monómero, disminuyendo la estabilidad térmica del dímero de NQO1, y al sitio de unión del FAD, reduciendo la afinidad por este.

3. Las mutaciones localizadas en el sitio activo de NQO1 tienen los efectos más fuertes observados hasta el momento sobre la afinidad por FAD y sobre la actividad enzimática de NQO1, incluso por encima del polimorfismo P187S.

4. La desestabilización estructural del sitio de unión del FAD causada por las variantes naturales y fosfomiméticas reduce la afinidad de la proteína por el cofactor. Eso se traduce en una disminución de la estabilidad intracelular de la proteína al estar favorecido su estado como apoproteína, que es más fácilmente poliubiquitinada y degradada por el proteosoma.

5. La "fosforilación" de S82 provoca una desestabilización local que se propaga a distintos sitios funcionales, al igual que ocurre con el polimorfismo asociado a enfermedad P178S. Esta desestabilización favorece estados no competentes para la unión del cofactor disminuyendo la afinidad por FAD, la actividad catalítica y la estabilidad intracelular.

6. El efecto de la fosforilación es específico en cada sitio de NQO1. La "fosforilación" de S40 afecta principalmente a la estabilidad conformacional e intracelular. La "fosforilación" de S82 afecta a la estabilidad conformacional e intracelular, así como a afinidad por FAD y la actividad enzimática. Y la "fosforilación" de T128 sólo afecta a la actividad enzimática.

5. CONCLUSIONS

Based on the results obtained, the most relevant conclusions drawn from this thesis are as follows:

1. Natural variants of NQO1 found in the population, both in samples from healthy individuals and in cancer samples, can impact the stability and function of NQO1.

2. The destabilizing effects of mutations located at the N-terminal end can propagate to distant regions, such as the monomer-monomer interface, reducing the thermal stability of the NQO1 dimer, and to the FAD binding site, decreasing its affinity for FAD.

3. Mutations located in the active site of NQO1 have the most pronounced effects observed so far on the affinity for FAD and on the enzymatic activity of NQO1, even surpassing the P187S polymorphism.

4. The structural destabilization of the FAD binding site caused by natural and phosphomimetic variants reduces the protein's affinity for the cofactor. This results in a decrease in the protein's intracellular stability, favouring its state as an apoprotein, which is more easily polyubiquitinated and degraded by the proteasome.

5. Phosphorylation of S82 causes local destabilization that propagates to various functional sites, similar to what happens with the disease-associated P187S polymorphism. This destabilization promotes states that are not competent for cofactor binding, reducing the affinity for FAD, catalytic activity, and intracellular stability.

6. The effect of phosphorylation is specific at each site of NQO1. Phosphorylation of S40 mainly affects conformational and intracellular stability. Phosphorylation of S82 affects conformational and intracellular stability, as well as FAD affinity and enzymatic activity. Phosphorylation of T128 only affects enzymatic activity.

6. PERSPECTIVAS FUTURAS

Durante esta tesis hemos estudiado distintas variantes naturales de NQO1 que se han encontrado en la población, y hemos observado que algunas de ellas parecen tener un efecto neutro mientras que otras afectan fuertemente a la estabilidad y función enzimática de la proteína. Esta caracterización *in vitro* se ha llevado a cabo únicamente desde el punto de vista biofísico y bioquímico. El siguiente paso en este proyecto sería la transfección de líneas celulares para expresar estas variantes y ver cómo afectan a su estabilidad intracelular. Dentro de la célula eucariota, NQO1 interacciona con una red de chaperonas diferentes a las que hay en bacterias, que pueden ayudar a la proteína en su plegamiento y eso puede favorecer que algunas variantes adopten la conformación correcta[13]. Por el contrario, variantes que parece que no afectan a la conformación global de la proteína, pueden disminuir la estabilidad de la proteína dentro de la célula dirigiéndola a su degradación[26,32]. Además, hasta ahora sólo nos hemos fijado en el efecto de estas mutaciones sobre la función enzimática. Sin embargo, NQO1 es una proteína multifuncional que, aparte de su función enzimática, es capaz de interactuar con otras biomoléculas, como algunas proteínas supresoras de tumores estabilizándolas frente a su degradación[121]. Por lo tanto, es posible que aquellas mutaciones que hemos observado que parece que no afectan a la estabilidad y/o función enzimática de NQO1, provoquen cambios en los sitios de unión con esas biomoléculas impidiendo esa interacción. Por lo tanto, su estudio intracelular también implicaría ver si estas mutaciones afectan a la interacción con esas biomoléculas y si disminuye la estabilidad intracelular de esas proteínas que son estabilizadas por NQO1.

Por otra parte, sería interesante contar con herramientas bioinformáticas que nos ayuden a predecir las consecuencias fenotípicas de todas las variantes que existen en el genoma humano, ya que caracterizarlas individualmente de forma experimental para poder aplicar una medicina personalizada no es factible. Como puede observarse, la predicción realizada tanto en la Publicación 1 como en la Publicación 2, falla en algunos casos (por ejemplo, las mutaciones F107C y H162N). En colaboración con el Dr. Kresten Lindorff-Larsen estamos trabajando en la mejora de algoritmos predictivos, comparando la caracterización experimental con la computacional (Anexo II). Para ello, continuar con la caracterización del resto de variantes naturales encontradas es esencial ya que los datos experimentales obtenidos ayudan a mejorar esas herramientas

predictivas. De hecho, recientemente se ha desarrollado AlphaMissense, una herramienta predictiva basada en predicción de la estructura con AlphaFold2, que es capaz de predecir el efecto de variantes de cambio de sentido de forma muy precisa[209].

Respecto a la fosforilación de NQO1, por un lado, sería interesante seguir la estrategia utilizada en esta tesis para caracterizar los 9 sitios de fosforilación restantes y ver que efecto tienen sobre la proteína y su posible papel biológico. Por otro lado, sería interesante identificar las quinasas y fosfatasas involucradas en la fosforilación de S82 y el resto de sitios. Dado que NQO1 está sobreexpresada en algunos tumores[210] y su inhibición ha sido propuesta como una posible terapia frente al cáncer[168], identificar esas quinasas y fosfatasas, y conseguir su activación e inhibición respectivamente, de forma específica en células tumorales podría reducir los niveles intracelulares de NQO1 y provocar la muerte de esas células cancerosas.

CONTRIBUCIONES PERSONALES

En esta sección resaltaré mis contribuciones a las publicaciones que se han usado en esta tesis:

Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein.

- Expresión y purificación de proteínas.
- Western-blot.
- Determinación del contenido de FAD espectrofotométricamente.
- Estudios de desnaturalización térmica por fluorescencia.
- Experimentos de proteólisis limitada con termolisina.
- Titulaciones por fluorescencia.

Counterintuitive structural and functional effects due to naturally occurring mutations targeting the active site of the disease-associated NQO1 enzyme.

- Expresión y purificación de proteínas.
- Determinación del contenido de FAD espectrofotométricamente.
- Experimentos de dicroísmo circular.
- Análisis de dispersión dinámica de la luz.
- Ensayos de fluorescencia.
- Titulaciones por fluorescencia.
- Titulaciones por dicroísmo circular.

Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD (P) H: quinone oxidoreductase 1 and their pharmacological rescue.

- Expresión y purificación de proteínas.
- Análisis de expresión en células eucariotas.
- Estudios de suplementación con riboflavina y dicumarol en células.
- Estudios de inhibición del proteosoma.

Different phenotypic outcome due to site-specific phosphorylation in the cancer-associated NQO1 enzyme studied by phosphomimetic mutations.

- Expresión y purificación de proteínas.
- Determinación del contenido de FAD espectrofotométricamente.
- Titulaciones por fluorescencia.
- Experimentos de proteólisis limitada con termolisina.
- Estudios de desnaturalización térmica por fluorescencia.
- Análisis de expresión en células eucariotas.
- Estudios de inhibición del proteosoma.

ANEXO I

Allosteric Communication in the Multifunctional and Redox NQO1 Protein Studied by Cavity-Making Mutations

Allosteric Communication in the Multifunctional and Redox NQO1 Protein Studied by Cavity-Making Mutations

Juan Luis Pacheco-Garcia [1,*,†], Dmitry S. Loginov [2,†], Ernesto Anoz-Carbonell [3,†], Pavla Vankova [4,5], Rogelio Palomino-Morales [6], Eduardo Salido [7], Petr Man [2], Milagros Medina [3], Athi N. Naganathan [8] and Angel L. Pey [9,*]

[1] Departamento de Química Física, Universidad de Granada, Av. Fuentenueva s/n, 18071 Granada, Spain
[2] Institute of Microbiology—BioCeV, Academy of Sciences of the Czech Republic, Prumyslova 595, 252 50 Vestec, Czech Republic; dmitry.loginov@biomed.cas.cz (D.S.L.); pman@biomed.cas.cz (P.M.)
[3] Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias, Instituto de Biocomputación y Física de Sistemas Complejos (BIFI) (GBsC-CSIC Joint Unit), Universidad de Zaragoza, 50009 Zaragoza, Spain; eanoz@unizar.es (E.A.-C.); mmedina@unizar.es (M.M.)
[4] Institute of Biotechnology—BioCeV, Academy of Sciences of the Czech Republic, Prumyslova 595, 252 50 Vestec, Czech Republic; pavla.vankova@ibt.cas.cz
[5] Department of Biochemistry, Faculty of Science, Charles University, Hlavova 2030/8, 128 43 Prague, Czech Republic
[6] Departamento de Bioquímica y Biología Molecular I, Facultad de Ciencias y Centro de Investigaciones Biomédicas (CIBM), Universidad de Granada, 18016 Granada, Spain; rpm@ugr.es
[7] Center for Rare Diseases (CIBERER), Hospital Universitario de Canarias, Universidad de la Laguna, 38320 Tenerife, Spain; edsalido@gmail.com
[8] Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras (IITM), Chennai 600036, India; athi@iitm.ac.in
[9] Departamento de Química Física, Unidad de Excelencia en Química Aplicada a Biomedicina y Medioambiente e Instituto de Biotecnología, Universidad de Granada, Av. Fuentenueva s/n, 18071 Granada, Spain
* Correspondence: juanlupacheco@correo.ugr.es (J.L.P.-G.); angelpey@ugr.es (A.L.P.); Tel.: +34-958243173 (A.L.P.)
† These authors contributed equally to this work.

Abstract: Allosterism is a common phenomenon in protein biochemistry that allows rapid regulation of protein stability; dynamics and function. However, the mechanisms by which allosterism occurs (by mutations or post-translational modifications (PTMs)) may be complex, particularly due to long-range propagation of the perturbation across protein structures. In this work, we have investigated allosteric communication in the multifunctional, cancer-related and antioxidant protein NQO1 by mutating several fully buried leucine residues (L7, L10 and L30) to smaller residues (V, A and G) at sites in the N-terminal domain. In almost all cases, mutated residues were not close to the FAD or the active site. Mutations L→G strongly compromised conformational stability and solubility, and L30A and L30V also notably decreased solubility. The mutation L10A, closer to the FAD binding site, severely decreased FAD binding affinity (≈20 fold vs. WT) through long-range and context-dependent effects. Using a combination of experimental and computational analyses, we show that most of the effects are found in the apo state of the protein, in contrast to other common polymorphisms and PTMs previously characterized in NQO1. The integrated study presented here is a first step towards a detailed structural–functional mapping of the mutational landscape of NQO1, a multifunctional and redox signaling protein of high biomedical relevance.

Keywords: antioxidant defense; flavoprotein; FAD binding; structural perturbation; protein core; allosterism; cavity-making mutation

1. Introduction

Protein residues are coupled to each other by hydrogen-bond networks and packing interactions causing correlated motions and fluctuations in equilibrium that are critical for protein function [1]. These networks enable the *transmission* of signals due to mutations or PTMs, playing critical roles in multiple protein functions through complex mechanisms [1,2]. Since mutational effects can propagate to long distances [1–4] and many flavoproteins are multifunctional [2], our understanding of the mutational effects in structure–function relationships is limited by detailed experimental and computational analysis. Generally, the transmission of these mutational effects to distant functional effects can be referred to as allosteric communication (the basis of allosterism).

In this work, we carried out a detailed biochemical, biophysical, structural and computational analysis of the NAD(P)H quinone:oxidoreductase 1 enzyme (NQO1), a flavoprotein that displays multiple functions (reduction of quinones, xenobiotic detoxification, superoxide scavenging, and interaction with other macrobiomolecules) and shows allosteric communication of mutational and ligand binding effects to different functional sites [5–8]. NQO1 is a homodimeric flavoprotein that contains a tightly bound FAD molecule per subunit, and it is involved in the two-electron reduction of substrates, including cancer prodrugs, vitamins and superoxide radicals [5,9]. Additionally, NQO1 is associated, due to polymorphisms, mutations and altered expression levels, with several diseases such as cancer, Alzheimer's disease and Parkinson's disease [5,9]. Structurally, the protein is divided into two different domains: a large N-terminal domain (NTD; approximately residues 1–225) that harbors most of the active site residues and a tightly bound FAD molecule and a C-terminal domain (CTD; approximately residues 225–274) that completes the active site (i.e., the NAD(P)H and substrate binding sites) and the monomer:monomer interface [10–14]. The NQO1 catalytic cycle follows a ping-pong mechanism: in the reductive half reaction, a NAD(P)H molecule binds to the active site and reduces FAD to $FADH_2$; in the oxidative half reaction, the substrate binds and it is reduced by $FADH_2$, regenerating the FAD in the oxidized state. Remarkably, the two active sites (one per protomer, with FAD bound and the NADH/NADPH and substrate binding sites) are non-equivalent: both the oxidative and reductive half reactions in the first active site are one order of magnitude faster than the same half reaction in the second active site, thus explaining the negative cooperativity observed in steady-state experiments [7,15]

To test the presence and extent of allosteric communication between different functional sites in NQO1, we have generated in this work cavity-making mutations of fully buried Leu residues (at L7, L10 and L30) to Val, Ala and Gly (Figure 1). This approach has been largely used to investigate the contribution of the hydrophobic effect, the role of packing interactions on conformational stability and the propagation of mutational effects on protein structure and function [1,16]. To the best of our knowledge, this approach has never been applied to a human flavoprotein. The Leu residues selected in NQO1 are in most of the cases far from the FAD or NADH/dicoumarol binding pockets in the active site (Figure 1) [dicoumarol (Dic) is represented because it is a competitive inhibitor of NAD(P)H and crystal structures are available]. Our results establish that these leucine residues are in long-range allosteric communication with active site residues in a highly context-dependent manner, leading in some cases to counterintuitive effects on several of the multiple functional features of this flavoenzyme.

2. Materials and Methods

2.1. Protein Expression and Purification

Mutations were introduced by site-directed mutagenesis in the wild-type (WT) NQO1 cDNA cloned into the pET-15b vector (pET-15b-NQO1) by GenScript (Leiden, Netherlands). Codons were optimized for expression in *Escherichia coli* and mutagenesis was confirmed by sequencing the entire cDNA. The plasmids were transformed in *E. coli*

BL21(DE3) cells (Agilent Technologies, Santa Clara, CA, USA) for protein expression. These constructs contain a hexa-his N-terminal tag for purification.

To determine the amount of soluble NQO1, 5 mL of LB medium containing 0.1 mg·mL^{-1} ampicillin (LBA)(purchased from Canvax Biotech, Córdoba, Spain) was inoculated with transformed cells and grown for 16 h at 37 °C. A volume of 0.5 mL of these cultures was diluted into 10 mL of LB containing 0.1 mg·mL^{-1} ampicillin and grown at 37 °C for 3 h. After that, cultures were induced with 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG, Canvax Biotech, Córdoba, Spain) at 37 °C for 4 h. Cells were harvested by centrifugation at 2900× g in a bench centrifuge at 4 °C and frozen at −80 °C for 16 h. Cells were resuspended in binding buffer (20 mM Na-phosphate, 300 mM NaCl, 50 mM imidazole, pH 7.4) with 1 mM phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich, Madrid, Spain) and sonicated in an ice bath. A volume of 1 mL were taken as *total extracts* and 1 mL was centrifugated (24,000× g, 30 min, 4 °C in a bench centrifuge) to obtain the *soluble extracts* (Figure 2). The amount of NQO1 in total and soluble extracts was determined by Western blot analysis (Figure 2). Samples were denatured using Laemmli's buffer and resolved in 12% acrylamide SDS-PAGE gels and transferred to PVDF membranes (GE Healthcare, Chicago, IL, USA) using standard procedures. Immunoblotting was carried out using primary monoclonal antibody against NQO1 (sc-393736, Santa Cruz Biotechnology, Dallas, TX, USA) at 1:500 dilution and, as secondary antibody, we used an anti-mouse IgGκ BP-HRP (sc-516102, Santa Cruz Biotechnology) at 1:2000 dilution. Samples were visualized using Luminol-based enhanced chemiluminescence (from BioRad Laboratories, Hercules, CA, USA), scanned (using a Chemidoc XRS+ system from BioRad Laboratories, Hercules, CA, USA) and analyzed using Image Lab (from BioRad Laboratories).

For large-scale purifications, a preculture (40 mL) was prepared from a single clone for each variant and grown for 16 h at 37 °C in LBA and diluted into 2.4–4.8 L. After 3 h with shaking (200 rpm) at 37 °C, NQO1 expression was induced by the addition of 0.5 mM IPTG for 6 h at 25 °C. Cells were harvested by centrifugation at 8000× g and frozen overnight at −80 °C. NQO1 proteins were purified using immobilized nickel affinity chromatography columns (IMAC, GE Healthcare) and size-exclusion chromatography (SEC) as described [15] (Figure S1). Isolated dimeric fractions of NQO1 variants were exchanged to HEPES-KOH buffer 50 mM pH 7.4 using PD-10 columns (GE Healthcare). The UV–visible spectra of purified NQO1 proteins were registered in a Cary spectrophotometer (Agilent Technologies, Waldbronn, Germany) and used to quantify the content of FAD as described in [15]. For the samples for pre-steady state kinetic analyses, NQO1 proteins were incubated with 1 mM FAD and excess FAD was removed using PD-10 columns, obtaining a saturation fraction (FAD:NQO1 monomer) higher than 90% based on UV–visible spectra. Apo proteins were obtained by treatment with 2 M urea and 2 M KBr as described [6], obtaining samples with less than 2% saturation fraction of FAD based on UV–visible spectra. Samples were stored at −80 °C upon flash freezing in liquid N$_2$. Protein purity and integrity were checked by polyacrylamide gel electrophoresis in the presence of sodium dodecylsulphate (SDS-PAGE) (Figure S1).

2.2. Thermal Stability

Thermal denaturation of NQO1 proteins, as holo proteins (2 μM in monomer + 100 μM FAD), was monitored by following changes in tryptophan emission fluorescence in HEPES-KOH 50 mM at pH 7.4 as described in [17]. T$_m$ values are indicated as the mean ± s.d. of four replicates.

2.3. Partial Proteolysis by Thermolysin

NQO1 samples (10 μM in monomer) were prepared in HEPES-KOH 50 mM at pH 7.4 in the presence of 100 μM FAD (NQO1$_{holo}$) and 100 μM FAD + 100 μM Dic (NQO1$_{dic}$) in a volume of 135 μL and incubated at 25 °C for 5 min. Thermolysin (from *Geobacillus stearothermophilus*, Sigma-Aldrich, St. Louis, MO, USA) was prepared at 1 and 5 μM

(protease concentration) in HEPES-KOH 50 mM at pH 7.4 and 100 mM CaCl₂. To trigger the reaction, solutions of thermolysin were added to those of NQO1 (previously preincubated at 25 °C for 5 min) at a 1:10 ratio. Samples were collected over time and the reaction quenched by adding EDTA pH 8 (final concentration of 20 mM) and Laemmli's buffer (2×). Controls (time 0) were prepared likewise but with no added thermolysin. Experiments were carried out at 25 °C. Samples were resolved by SDS-PAGE under reducing conditions in gels containing 12% acrylamide as resolving gel and 4% acrylamide as stacking gel. Gels were stained with Coomassie blue G-250. Densitometry was carried out using ImageJ. Proteolysis reactions were analyzed using an exponential function to provide the apparent rate constant (k_{obs}).

2.4. FAD Content

FAD content was determined spectroscopically as described in [18]. Briefly, the UV–visible spectra of purified NQO1 proteins were registered in a Cary 100 spectrophotometer (Agilent). The UV–visible spectra were normalized considering that the UV–visible spectra can be deconvoluted using a $\varepsilon_{280} = 47,900$ M⁻¹·cm⁻¹ for the apo-NQO1 and $\varepsilon_{280} = 22,000$ M⁻¹·cm⁻¹ and $\varepsilon_{450} = 11,300$ M⁻¹·cm⁻¹ for free FAD as previously described [17].

2.5. FAD Binding Affinity

Fluorescence titrations were carried out at 25 °C using 1 × 0.3 cm path-length cuvettes in a Cary Eclipse spectrofluorimeter (Agilent Technologies, Waldbronn, Germany). Experiments were carried out in 20 mM K-phosphate, pH 7.4, essentially as described in [19]. Briefly, 20 μL of a 12.5 μM NQO1 stock solution (in subunit) was mixed with 0–500 μL of FAD 10 μM and the corresponding volume of buffer was added to yield a 1 mL final volume. Samples were incubated at 25 °C in the dark for at least 10 min before measurements. Spectra were acquired in the 340–360 nm range upon excitation at 280 nm (slits 5 nm), and spectra were averaged over 10 scans registered at a scan rate of 200 nm·min⁻¹. FAD binding fluorescence intensities at 350 nm were fitted using single and identical types of binding sites as described in [19].

2.6. Enzyme Kinetics for the Reductive Half Reaction with NADH

Fast hydride- and deuteride-transfer reactions (HT and DT, respectively) were carried out under anaerobic conditions using a stopped-flow spectrophotometer as described [15]. Briefly, the reductive half reaction was measured by mixing NQO1holo variants (7.5 μM) with NADH ranging from 7.5 to 100 μM (these refer to final concentrations). Reactions were performed in 20 mM HEPES-KOH, pH 7.4. Data were collected using either NADH or NADD, but using only one of these reducing species in a given experiment. Multiple-wavelength absorption data in the flavin absorption region were collected and processed as described [15]. Time-dependent spectral deconvolution was performed by global fitting analysis and numerical integration using previously described procedures [15]. This deconvolution procedure was carried out considering sequential and irreversible steps in the context of a two-step mechanism (A→B→C) and was used to determine observed rate constants (k_{obs}) for these steps as well as the spectroscopic properties of these species (A, B and C). According to our recent study, catalytically relevant NQO1 processes involved steps A→B→C [15]. Despite practical limitations preventing these measurements reaching pseudo-first-order conditions (such as minimum amount of protein sample for detection of flavin reduction and processes over 75 μM NADH becoming too close to the instrumental death time), hyperbolic dependences of k_{obs} vs. NADH concentrations were fitted using Equation (1):

$$k_{obs} = \frac{k_{HT} \cdot [NADH]}{K_d^{NADH} + [NADH]} \tag{1}$$

where k_{HT} is the limiting rate constant for HT and K_d^{NADH} is the apparent equilibrium dissociation constant to a given active site.

To determine primary kinetic isotopic effects (KIEs) in the HT process [20], the k_{obs} for HT and DT was determined by mixing NADH/D with NQO1$_{holo}$ using equimolar concentrations of NQO1$_{holo}$ and NADH or [4R-^{2}H]-NADD (7.5 μM of each component again experiments at higher NADH concentrations where limited but reactions becoming too fast for detection upon increasing temperature). These apparent KIEs were determined as the ratio of k_{obs} values using NADH and NADD. Experiments were carried out at temperatures ranging 6–20 °C. Activation parameters (frequency factor, A, and the activation energy, E$_a$) were determined using the Arrhenius equation as described [15].

2.7. Hydrogen/Deuterium Exchange Mass Spectrometry (HDXMS)

The structural impact of cavity-making mutations (L7V, L7A, L10V, L10A, L30V, and L30A) in NQO1 was evaluated using hydrogen/deuterium exchange (HDX) coupled to mass spectrometry. Proteins were monitored in their NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$ states with the only exception of L30A mutant being analyzed solely as NQO1$_{holo}$ and NQO1$_{dic}$. To start the HDX reaction, 20 μM protein was diluted 10 fold into a D$_2$O-based buffer [50 mM HEPES-KOH, 0.5 mM TCEP (tris(2-carboxyethyl)phosphine hydrochloride, Sigma-Aldrich, St. Louis, MO, USA), pD 7.4]. Under NQO1$_{holo}$ and NQO1$_{dic}$ conditions, 20 μM protein variants were pre-incubated at least for 10 min with 200 μM FAD (NQO1$_{holo}$) or with 200 μM FAD + 200 μM dicoumarol (NQO1$_{dic}$). After 10, 50, 250, 1250 and 6250 s of HDX the reaction was quenched by mixing with 0.5 M Glycine-HCl, pH 2.3 in ratio 1:1 and samples were frozen in liquid nitrogen. The time points 10, 250 and 6250 s were prepared as duplicates. Subsequently, each sample was thawed and injected onto the LC system including serially coupled immobilized protease columns (nepenthesin-2 and pepsin) where it was digested for 3 min by flow of 0.4% formic acid (FA) in water, delivered at a flow rate of 200 μL min^{-1} (1260 Infinity II Quaternary pump, Agilent Technologies, Waldbronn, Germany). Generated peptides were online trapped and desalted on SecurityGuard™ pre-column (ULTRA Cartridges UHPLC Fully Porous Polar C18, 2.1 mm, Phenomenex, Torrance, CA, USA). Next, peptides were separated on a reversed-phase analytical column (LUNA® Omega Polar C18 Column, 100 Å, 1.6 μm, 100 mm × 1.0 mm, Phenomenex, Torrance, CA, USA) at a flow rate of 40 μL min^{-1} using a 10–40% linear gradient of solvent B (A: 0.2% acetonitrile/0.1% FA in water; B: 98% acetonitrile/0.1% FA in water) (1290 Infinity II LC system, Agilent Technologies, Waldbronn, Germany) as described [6]. Digestion and separation were performed at 0 °C and pH 2.3 to minimize deuterium loss. The LC system was directly interfaced with an ESI source of 15T FT-ICR mass spectrometer (SolariX XR, Bruker Daltonics, Bremen, Germany). Data were exported and processed using the Data Analysis v. 5.3 (Bruker Daltonics, Bremen, Germany) and in-house developed DeutEx software [21]. Peptide identification for each variant was performed by data-dependent LC–MS/MS measurement using the same LC system and gradient elution but with ESI-timsTOF Pro PASEF instrument (Bruker Daltonics, Bremen, Germany). Peptides were identified by a MASCOT (Matrix Science, London, UK) search against a custom-built database combining sequences of NQO1 variants and contaminants from the cRAP database. Decoy search was enabled, the false discovery ratio was set to 1% and the ion score cut-off to 20. Exchange rates were corrected for back exchange as described previously [6,22]. To evaluate the effect of mutations, the difference in kinetics of deuterium incorporation (% D vs. time) of mutants and the WT protein was calculated for a given ligation state and for each protein segment experimentally determined, and the average of the two most different time points (mutant-WT) was determined (i.e., Δ%D$_{av}$). This procedure enabled determining the difference in HDX kinetics between two given NQO1 states as recently described [2,6].

2.8. Statistical Mechanical Model Predictions

The block version of the Wako–Saitô–Muñoz–Eaton model (bWSME) was employed by fixing the parameters to those from a recent analysis of NQO1 folding thermodynamics [23]. A detailed description of the model and model parameterization has been provided in a recent work [24]. Briefly, the residue-level coarse graining in the conventional Gō- and Ising-like WSME model is extended to included stretches of consecutive residues or blocks (b, with the most probable block size of 3). The model includes contributions from microstates with single stretches of folded residues (single sequence approximation, SSA), two stretches of folded residues (double sequence approximation, DSA) and DSA, allowing for interactions across the folded islands if they interact in the native structure (DSAw/L). The structure of NQO1 (PDB code 2F1O) [25] is employed to calculate the number and nature of van der Waals interactions (with a 5 Å heavy-atom cut-off), charge–charge interactions (protonation state of pH 7.0) and simplified solvation free energies for every microstate that numbers to 5,002,535, at 310 K and 100 mM ionic strength. The statistical weights of each of these microstates are summed up to calculate the overall partition function and partial partition functions, from which the free energy profiles are predicted.

Once the probabilities of every microstate are calculated, the ensemble is split into two sub-ensembles, which account for the probability of residue *i* to be folded with respect to the folding probability of residue *j*, to calculate coupling free energies [26]. Specifically, $\sum p_{i_f j_f}$ sums over all states in which both residue *i* and residue *j* are folded while $\sum p_{i_u j_f}$ accounts for all states in which residue *i* is unfolded and *j* is folded. Following this, the positive coupling free energy is calculated for every residue as:

$$\Delta G_+ = RTln(K_+) = RTln\left(\frac{\sum p_{i_f j_f}}{\sum p_{i_u j_f}}\right) \qquad (2)$$

The resulting two-dimensional matrix is termed the positive coupling matrix. The calculation is repeated for the mutant proteins (L7A, L10A, and L30A; mutations introduced via PyMol [27]) with the same set of parameters.
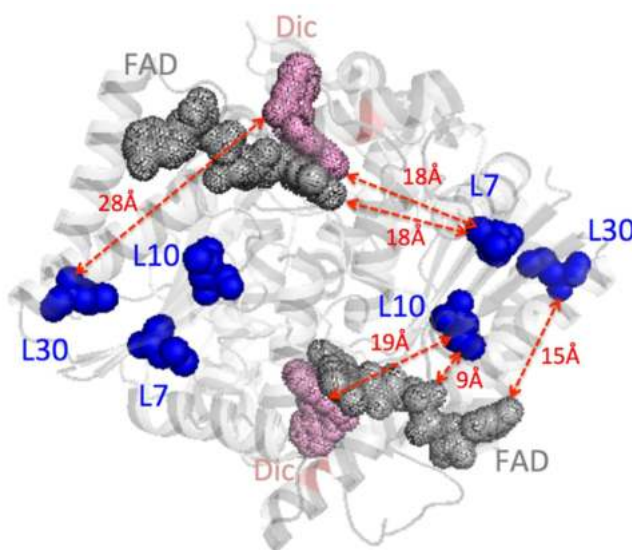
Figure 1. Structural location of the leucine residues mutated in this work. Leucine residues are indicated in blue. FAD is indicated in grey and Dic in pink. The figure was generated using the structure with PDB code 2F1O [25]. Solvent-accessible surface areas for L7, L10 and L30 are 0.3 ± 0.2%, 0.4 ± 0.2% and strictly 0%, respectively, using GetArea (http://curie.utmb.edu/getarea.html: accessed on 1 October 2019). Shortest distances between Leu residues and FAD and Dic molecules are indicated with red dashed lines.

3. Results and Discussion

3.1. Expression and Solubility of Cavity-Making Mutants

We mutated three buried Leu residues (L7, L10 and L30) to Val, Ala and Gly (Figure 1). These Leu residues are far from the FAD and Dic molecules, with only L10 being slightly closer than 10 Å to the FAD binding site (Figure 1). Even though codon optimization was used to generate the mutants, we found large differences in total protein and soluble protein levels among them. Western blot analysis of total *E. coli* extracts revealed decreased total protein levels in L7A, L30A and L30G mutants, and particularly for L7G that essentially abolished expression (Figure 2). Additionally, the ratio of the soluble-to-total protein amounts (S/T) showed that L7G, L30V, L30A, and L30G reduced the solubility of NQO1 (Figure 2). These results were confirmed by SEC and SDS-PAGE using large-scale purifications (Figure S1). Therefore, mutations to glycine were not further investigated due to extremely low yields and instability. Overall, these results also indicated that the effects of cavity-making mutations on expression levels and solubility largely depend on the location of the mutated site rather than its burial in the structure.
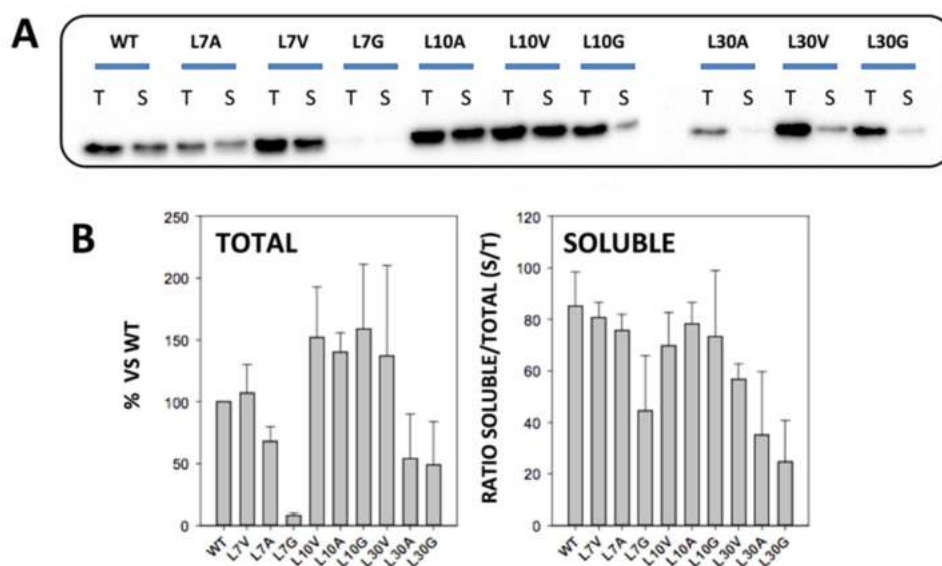
Figure 2. Expression levels and solubility of NQO1 variants. (**A**) Representative Western blot analysis of NQO1 total (T) and soluble (S) extracts from *E.coli* cells; (**B**) Quantitative analysis of NQO1 from expression levels (from six independent experiments; mean ± s.d.) solubility is displayed as the ratio of soluble/total protein. Expression was carried out at 37 °C for 4 h upon induction with IPTG 0.5 mM.

3.2. Thermal Stability of Cavity-Making Mutants

For those mutants that yielded good expression levels, we carried out further biophysical characterization. This set of mutants were analysed, first by thermal denaturation, which in the case of NQO1 is an irreversible process involving unfolding and dissociation of the protein dimer [18]. Therefore, these analyses do not provide thermodynamic information on mutational effects, but allows a qualitative comparison on their kinetic stability [18]. The experiments were carried out as NQO1$_{holo}$ proteins (purified protein + 100 μM FAD). The mutation L10A had the largest effect, decreasing the T_m by 8 °C. Mutations L7V, L7A and L10V showed a more modest destabilization (decreasing the T_m by 5–6 °C) (Figure S2). Intriguingly, the mutants L30V and L30A, that showed large effects on protein expression/solubility, only decreased the stability by 2–3 °C, suggesting that these mutants affect the expression/solubility more than thermal stability of the NQO1 dimer.

3.3. Local Stability of the NTD Investigated by Partial Proteolysis with Thermolysin

Partial proteolysis with thermolysin has been shown to inform on the local stability of the FAD binding site, with a primary cleavage site between S72 and V73 [14]. The proteolysis pattern is conserved across all mutants investigated (Figure S3). Enhanced sensitivity towards this proteolytic activity often correlates with reduced FAD binding affinity [2,14]. Thus, we carried out proteolysis experiments at two different concentrations of protease for all variants in the NQO1$_{holo}$ and NQO1$_{dic}$ states (i.e., protein as purified + 100 μM FAD + 100 μM dicoumarol) (Figure S3). Interestingly, none of the mutants cause large effects on proteolytic sensitivity (Figure S4 and Table S1). Therefore, these mutations did not seem to affect largely the local stability of the loop 57–66 close to the FAD binding site.

3.4. FAD Binding Affinity

To test potential effects on FAD binding, we first measured FAD content by UV–visible spectroscopy (Figure 3). Three variants (L10V, L10A and L30A) showed lower FAD content than the WT protein. Interestingly, the mutant L30A showed an abnormal spectra, with similar absorption at 375 and 450 nm, suggesting severe effects on the FAD binding mode (Figure 3A). Due to the correlation between FAD content of NQO1 protein variants and their affinity for FAD [14,18,19,28], we carried out direct titrations of apo-proteins with FAD to corroborate potential effects on FAD binding affinity (Figure 4). Mutants at L30 were too unstable to yield enough quantity of pure apo-protein. Nevertheless, we found that the mutant L10A reduced the affinity for FAD by 18 fold compared to the WT protein, which corresponds to a Gibbs free energy penalty of ~1.7 kcal·mol^{-1}. This result is also consistent with the expectation based on distance considerations—L10 is the closest to the FAD binding site among the Leu mutants studied (Figure 1).
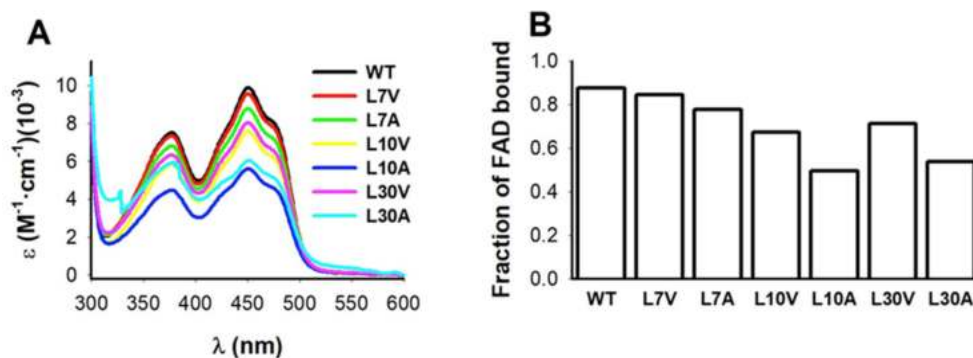
Figure 3. FAD content of purified NQO1 variants upon purification using IMAC plus SEC. (A) Absorption spectra in the region of absorption of FAD; (**B**) FAD content was calculated using the ratio of absorbance at 280 and 450 nm as described in materials and methods. Data are the average from two different purifications.
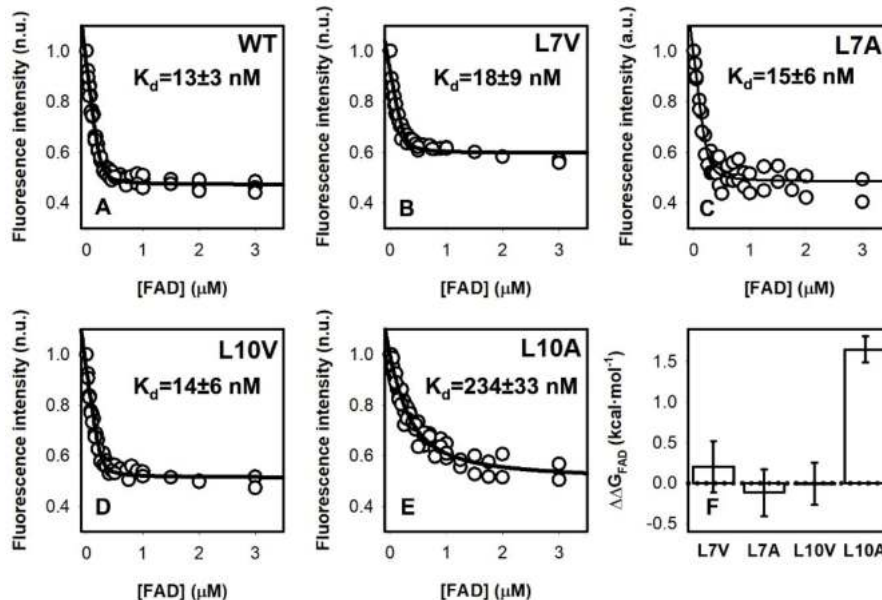
Figure 4. Fluorescence titrations of NQO1$_{apo}$ variants with FAD. (**A–E**) Experimental data are shown as circles and lines are best fits to a single type of independent binding sites. Data are from at least two independent experiments (N = 4 for WT and N = 2–3 for the mutants). (**F**) The changes in apparent binding free energy including errors from linear propagation. Experiments were performed at 25 °C.

3.5. Effect of Cavity-Making Mutations on Enzyme Kinetics

The reduction kinetics by NAD(P)H of the FAD bound to WT NQO1 is complex [15] (Figure S5). We have observed two different pathways with a difference in apparent rate constants of at least one order of magnitude, supporting the existence of functional negative cooperativity, named as *fast* and *slow* paths [2,15]. The cavity-making mutations had some effects on the catalytic cycle (Figure 5 and Table 1). Mutations L7A and L10A increased the K_d for NADH by 2 fold in the fast reaction pathway, and the L10A also increased the K_d for NADH by 3 fold in the slow reaction pathway (Table 1). However, it must be noted that these changes in the K_d for NADH are often associated with modest increases in the k_{HT}; therefore, catalytic efficiency is not largely affected in most cases (Table 1).
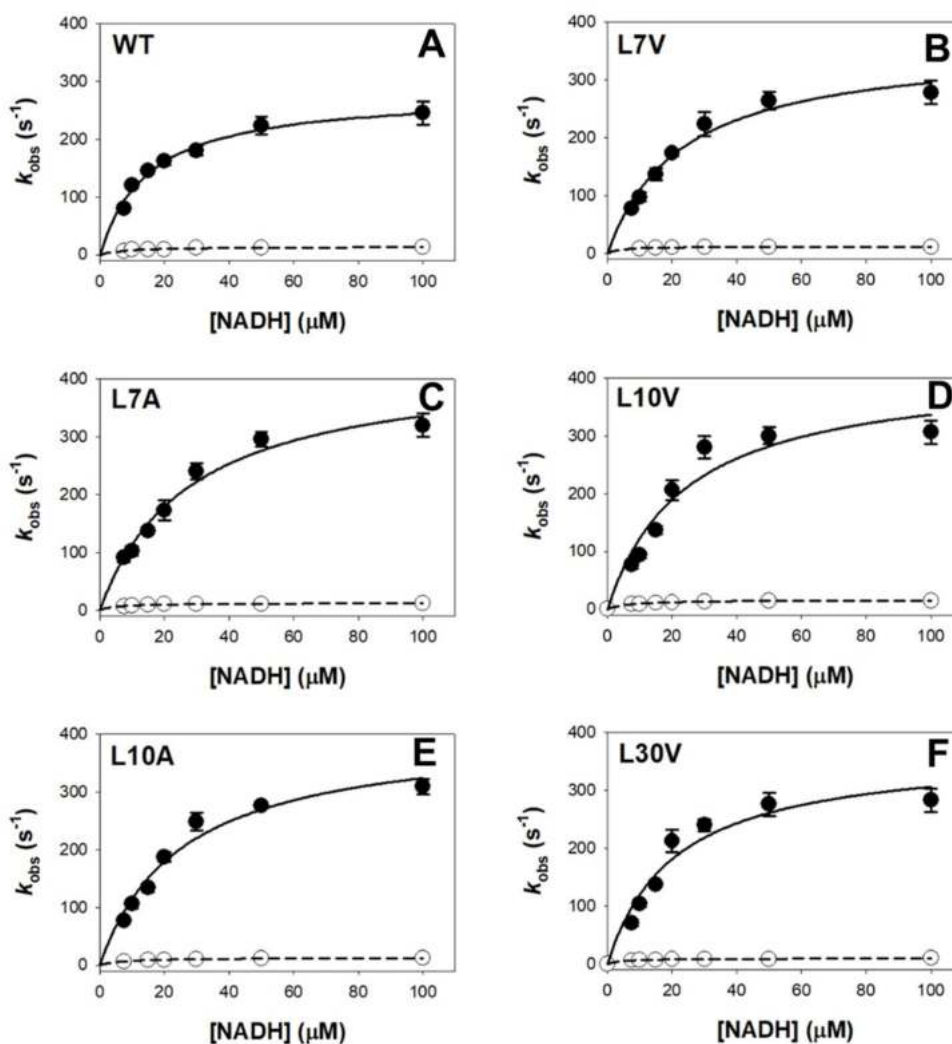
Figure 5. Dependence of k_{obs} on NADH concentration for NQO1 variants. (**A–F**) Enzyme kinetics for different variants (the variant identity is displayed in each panel). Experiments were carried out using 7.5 μM NQO1 at 6 °C. Data are from at least three replicates and displayed as the mean ± s.d. Black circles correspond to the fast pathway, and open circles show data for the slow pathway. Lines are best fits to Equation (1). Each panel shows the data for the given variant.

Table 1. Enzyme kinetic analysis for the reductive half reaction of NQO1$_{holo}$ variants with NADH. Primary data are shown in Figure 5.

Variant	$k_{HT\ FAST}$ (s^{-1})	$K_{d\ FAST}$ (μM)	$k_{HT}/K_{d\ FAST}$ (μM^{-1}·s^{-1})	$k_{HT\ SLOW}$ (s^{-1})	$K_{d\ SLOW}$ (μM)	$k_{HT}/K_{d\ SLOW}$ (μM^{-1}·s^{-1})
WT	281 ± 14	15 ± 2	19 ± 2	14 ± 2	8 ± 3	1.8 ± 0.6
L7V	363 ± 26	23 ± 4	16 ± 2	18 ± 1	6 ± 1	3.0 ± 0.5
L7A	430 ± 31	28 ± 5	15 ± 2	13 ± 1	5 ± 1	2.5 ± 0.4
L10V	417 ± 59	24 ± 8	17 ± 2	16 ± 1	7 ± 1	2.3 ± 0.8
L10A	460 ± 30	32 ± 5	14 ± 2	13 ± 1	7 ± 1	1.8 ± 0.1
L30V	370 ± 42	21 ± 6	18 ± 2	10 ± 1	5 ± 1	2.1 ± 0.6

The influence of mutations on the NQO1 active site dynamics was further studied by evaluating the reductive half reaction with NADH and NADD at different temperatures.

We use only those mutants whose solubility extents are not significantly affected, namely L7A, L7V, L10A and L10V. At the low temperature, fast and slow reduction steps were slowed down when using NADD for all variants similarly than in the WT (Table S2), resulting in low KIEs (1.5–2.2). However, some mutants showed temperature dependences of KIEs in contrast to the WT protein (Figure S6).

The temperature-independent KIEs in WT NQO1 were interpreted (for the fast and slow steps) as transitions under the barrier and tunneling of both proton and deuterium, while KIEs decreasing with temperature may indicate only proton tunnelling [2,15,29]. Analysis of WT data in the context of the Arrhenius equation indicated tunneling ready states with movements of active site heavy atoms do not actively contributing to bring the donor and acceptor to the distance for efficient tunnelling and HT but increase that to achieve tunnel ready conformations (passive dynamics). Replacements at L7 and L10 modulated pre-exponential factors (A) and activation energies (E_a) in both steps (Figure S6 and Table S2). Larger differences were found for the fast steps in L7V and L7A, and particularly for the slow steps in L10V and L10A (with large changes increases in A values and in activation energies). These changes in activation parameters suggest: (i) Passive dynamics is the major contributor to achieve the proton tunnel ready configuration; (ii) Some reaction steps might require a larger structural reorganization to engage the coenzyme in the catalytic process as well as a little contribution of donor–acceptor distance sampling to achieve tunnel ready conformations. Thus, volume changes at L7 and L10 mildly alter the overall packing and general dynamics of NQO1 active sites, with catalytic enhancement achieved by promoting and optimizing vibrations in active sites that minimize DAD fluctuations [29]. Overall, these observations actually agree well with HDX analyses that showed changes in structural stability in the close environment of the isoalloxazine binding site of NQO1$_{holo}$ and NQO1$_{dic}$ (Section 3.6).

3.6. Effect of Cavity-Making Mutations on the Structural Stability of NQO1

To evaluate the effect of cavity-making mutations on the local stability in different ligation states (NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$), we have used hydrogen–deuterium exchange (HDX) followed by mass spectrometry [2,6]

3.6.1. The L7 Cavity-Making Mutants

We observed significant changes in the local stability of the NQO1$_{apo}$ state upon mutations L7V and L7A. In the L7V mutant, these effects seemed to spread to longer regions, but were generally mild ($\Delta\%D_{av}$ of about 10%; in yellow). In the case of L7A, the effects were stronger but more local (Figure 6). Binding of FAD and/or dicoumarol led to a dramatic decrease in these destabilizing effect. Therefore, mutations L7V and L7A seem to mostly target the stability of the NQO1$_{apo}$ state. We must also remember that the L7G mutant almost fully abolished expression of the protein (Figure 2), suggesting that conformational flexibility at this site can be critical for proper expression and/or folding of the protein. Changes in the local stability of the active site (Figure 6) might be associated with the effects observed in enzyme kinetics and dynamics along the HT process.
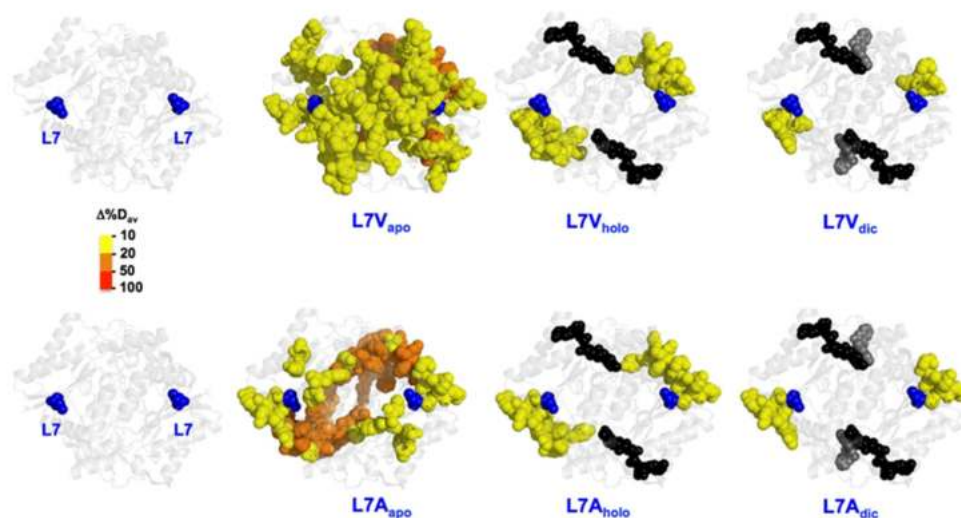
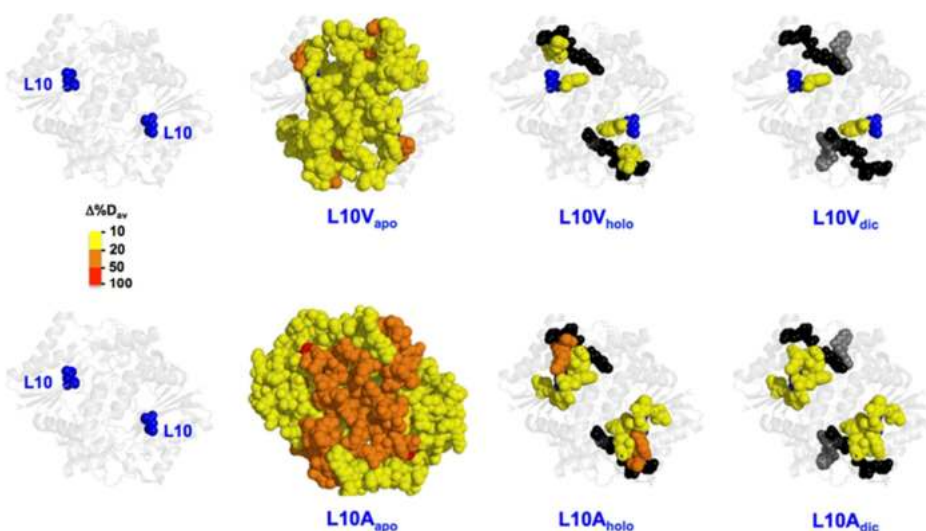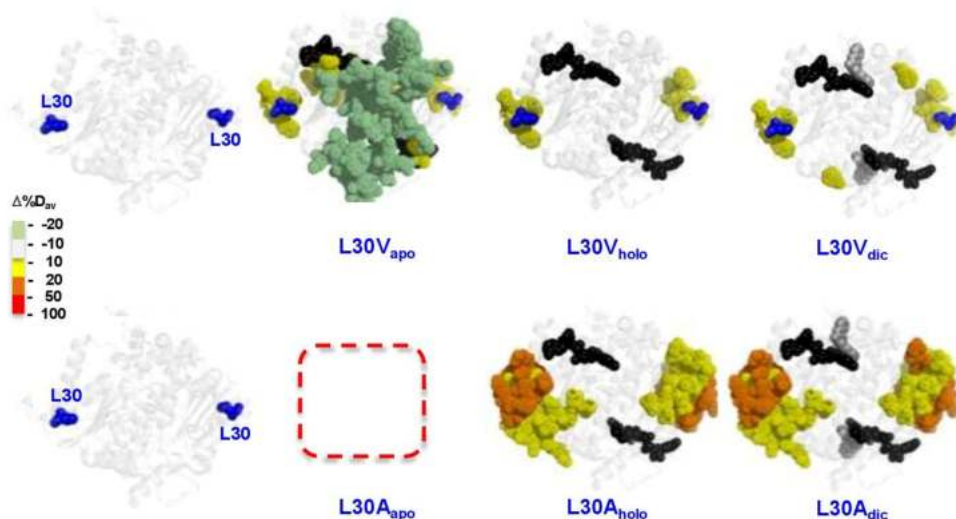**Figure 6. Representation of the changes in local stability (as Δ%D$_{av}$) due to mutations L7V and L7A vs. WT.** Results are shown as indicated in the color scale for different ligation states (NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$). FAD and Dic are shown as spheres (black and grey, respectively). The structural model used for display was PDB 2F1O [25].

3.6.2. The L10 Cavity-Making Mutants

As observed for L7 mutants (Section 3.6.1), the most remarkable changes in local stability caused by mutations at L10 were found in the NQO1$_{apo}$ state (Figure 7). We observed that the propagation of destabilizing effects due to mutations is substantially larger for the L10A than for L10V in the NQO1$_{apo}$ state (Figure 7). However, these effects were largely abolished in the NQO1$_{holo}$ and NQO1$_{dic}$ states, although some residual destabilization was still observed in the L10A mutant in both these states (Figure 7). This large destabilization due to the mutation L10A in the NQO1$_{apo}$ and NQO1$_{holo}$ states could explain the significant decrease for FAD binding affinity experimentally observed (Figure 4). We must also note that the L10G mutation severely reduced the solubility of the protein, making unfeasible its further characterization, and supporting that the combined effect of cavity making plus increased backbone flexibility has dramatic effects on NQO1 foldability. These results revealed a highly specific impact of mutations at L7 and L10 on the apo-state, in contrast to other mutations such as S82D and P187S that had severe impacts in different ligation states [2].

Figure 7. Representation of the changes in local stability (as Δ%D$_{av}$) due to mutations L10V and L10A vs. WT. Results are shown as indicated in the color scale for different ligation states (NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$). FAD and Dic are shown as spheres (black and grey, respectively). The structural model used for display was PDB 2F1O [25].

3.6.3. The L30 Cavity-Making Mutants

The results obtained by HDXMS for L30V and L30A were striking (the large loss in solubility and stability meant that probing the mutant L30G was not possible in any state) (Figure 2). L30V in the NQO1$_{apo}$ state showed mild destabilization around the mutation, but, surprisingly, other segments of the proteins were mildly stabilized. In the NQO1$_{holo}$ and NQO1$_{dic}$ states, only local and mild destabilization was observed (Figure 8). This could explain why this mutant has little effect on FAD binding, catalytic performance or stability of the protein (Figures 3–5, S2 and S3 and Table 1). Still, the mutation L30V reduced the levels of soluble and well folded protein (Figures 2 and S1).

Figure 8. Representation of the changes in local stability (as Δ%D$_{av}$) due to mutations L30V and L30A vs. WT. Results are shown as indicated in the color scale for different ligation states (NQO1$_{apo}$, NQO1$_{holo}$ and NQO1$_{dic}$). Note that light green indicates slight stabilization. FAD and Dic are shown as spheres (black and grey, respectively). The structural model used for display was PDB 2F1O [25].

The mutant L30A showed stronger destabilizing effects around the mutated site and these propagated to a longer distance in both NQO1$_{holo}$ and NQO1$_{dic}$ states (Figure 8). For this variant, we could not prepare samples of quality sufficient for HDXMS experiments in NQO1$_{apo}$ state possibly due to solubility/stability issues in this state.

3.7. Statistical Mechanical Calculations on the Effects of NQO1 Cavity-Making Mutants

To explore the role of cavity-making mutations on the conformational landscape of NQO1$_{apo}$, the folding free energy profile of the WT, as a function of the number of structured blocks, at 310 K is predicted employing the bWSME model [26]. The model accounts for an ensemble of more than 5 million microstates or conformational states (see Section 2.8 in Methods) whose probabilities are estimated employing a structure-based algorithm while accounting for van der Waals interactions, electrostatics and solvation free energy terms, apart from residue-level conformational entropy parameters [24]. The resulting free energy profile is indicative of a multistate-like folding behavior with several high free energy intermediates (Figure 9A). Three cavity-creating mutations—L7A, L10A and L30A—were introduced on the WT structure and fed into the model to predict the mutant free energy profiles. The mutant free energy profiles exhibit lower stability (Figure 9A)—note the lower free energy values for the mutants at ~10 structured blocks—as the mutation involves the replacement of a large aliphatic amino-acid (leucine) with a smaller amino acid (alanine). The folded ensemble, at ~70 structured blocks, appears to be relatively unperturbed on mutations. However, this could be a result of the fact that we are lumping together millions of microstates into a specific order parameter, i.e., the number of structured blocks.
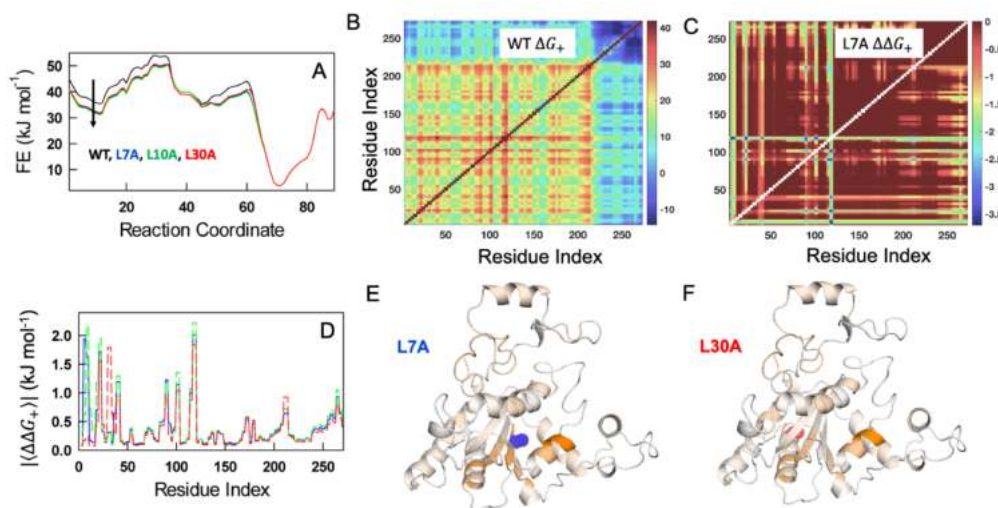
Figure 9. WSME model predictions on the effect of hydrophobic truncations. (**A**) Free energy profiles of WT NQO1 (black), L7A (blue), L10A (green), and L30A (red), as a function of the number of structured residues as the reaction coordinate. Note that the free energy of the unfolded state decreases (arrow), highlighting thermodynamic destabilization of the mutants due to the loss of van der Waals interactions upon truncation. (**B**) Positive coupling free energy matrix of the WT. The color bar shown to the right is in kJ mol^{-1}. (**C**) Difference in coupling free energy matrices between the WT and the L7A mutant. Color bar is in kJ·mol^{-1}. (**D**) Absolute values of the mean differences in positive coupling free energies between the WT and the mutants as a function of residue index. Since the differences are negative (panel (**C**), and as the WT residues are more strongly coupled), absolute values are employed for ease of visualization. (**E,F**) The values in panel (**D**) mapped onto the structure for L7A (panel (**E**)) and L30A (panel (**F**)) mutations. Dark orange corresponds to residues that are the most perturbed while white signals little or no change. The structural model used was PDB 2F1O [25].

The effect on the native ensemble can be more accurately gleaned by estimating the extent to which different residues are coupled to each other in the native ensemble. To this end, we constructed the matrix of positive coupling free energies (ΔG_+) that accounts for the sum of probabilities of states in which two residues simultaneously folded vs. those states in which they are decoupled (one residue is folded and the other is not) (Equation (2)). The positive coupling free energy matrix highlights that the majority of the residues are strongly coupled to each other (shades of red in Figure 9B) with the C-terminal region being weakly coupled (shades of blue in Figure 9B). The latter observation is consistent with HDX–MS experiments on WT NQO1$_{apo}$ [2,6]. This calculation is repeated for each of the mutants and a differential coupling matrix ($\Delta\Delta G_+ = \Delta G_{+,mut} - \Delta G_{+,WT}$) is generated to identify protein regions that are perturbed (shown for the L7A mutant in Figure 9C). It is observed that a large number of residues display lower coupling magnitudes in the native ensemble (negative values as the coupling is weaker; colors that are not red in Figure 9C). Plotting the absolute mean values of $\Delta\Delta G_+$ averaged across rows (or columns) as function of residue index provides a simpler view of the residue-wise altered coupling pattern (Figure 9D), which can also be visualized by mapping them on to the structure (Figure 9E,F). The average effects are small because we do not consider additional weakening of interactions in the second-shell around the mutated site [1]. However, this calculation suggests that the cavity-creating mutants have stronger effects in the NTD (residues 1–50 and 80–120), and that these uncoupling effects are dependent on the mutated site (Figure 9D–F). To summarize, the bWSME model predicts that mutations modulate the folding probability of multiple structural regions in NQO1 in a non-intuitive manner that is consistent with our experimental characterization carried out on these cavity-making mutants.

4. Conclusions

Human flavoproteins display multiple functional features, including oxido-reduction and antioxidant activities, interaction with small effectors and other biomacromolecules and transport to different subcellular and extracellular locations [2]. This multifunctionality is likely imprinted in their energy landscape in which different substates with different energetic and functional features can be populated. Mutations can affect this energy landscape by altering protein interaction networks and leading to allosteric effects due to long-range propagation of mutational effects [1,3,4]. Herein, we have perturbed the interaction network of the human flavoenzyme NQO1 at three buried positions (L7, L10 and L30) in the N-terminal domain by introducing cavity-making mutations (L→V and L→A), which may, additionally, enhance backbone flexibility (L→G). We showed that L→G mutations severely compromised protein folding and/or solubility, and L→V and L→A mutations affected protein stability, flavin binding, catalytic activity and functional cooperativity to different, and not always intuitive, extents. Using a combination of HDX–MS and statistical mechanical calculations, we also showed that these mutations affected the native state ensemble and folding landscape in a site-specific manner, mostly targeting the apo state of the enzyme at L7 and L10. Importantly, our work shows that cavity-making mutations affect the conformational states (e.g., apo-NQO1 and holo-NQO1) differently to naturally occurring and disease-associated mutations [2]. The fact that the pseudo-phosphorylating mutation S82D, the cancer-associated polymorphism P187S as well as natural and artificial mutations at different sites affect both the holo and apo states [2,19,28], whereas cavity making mostly targets the stability of the apo state nicely, shows the multiple pathways available for transmitting mutational effects to different functional features in different ligation states (i.e., allosteric communication). Our approach can be of general application to prove allosteric mechanisms at high resolution for this and other human flavoenzymes, which can in many cases be associated with disease upon mutation (please see the large genetic diversity found just for NQO1 in COSMIC at https://cancer.sanger.ac.uk/cosmic/search?q=NQO1 (accessed date: 20 May 2022) and

gnomAD at https://gnomad.broadinstitute.org/gene/ENSG00000181019?dataset=gnomad_r2_1 (accessed date: 20 May 2022).

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/article/10.3390/antiox11061110/s1. Figure S1: Purification of NQO1 proteins containing mutations at L7, L10 and L30; Figure S2: Thermal stability of NQO1 cavity-making mutants; Figure S3: Representative SDS-PAGE analysis for the proteolytic kinetics of NQO1 variants with thermolysin; Figure S4: Proteolysis kinetics of NQO1 variants with thermolysin; Figure S5: Reductive half reaction of FAD bound to NQO1 variants with NADH; Figure S6: Temperature dependence of kinetic parameters for the two hydride/deuteride transfer (HT/DT) processes from NADH to NQO1. Table S1: Observed rate constants (k_{obs}) for partial proteolysis of NQO1 variants with thermolysin. Table S2: Arrhenius parameters and KIEs for the HT/DT in the reduction of NQO1 variants by NADH/NADH.

Author Contributions: Conceptualization, A.L.P.; methodology, R.P.-M., P.M., M.M., A.N.N. and A.L.P.; software, P.M. and A.N.N.; validation, J.L.P.-G., D.S.L., E.A.-C. and P.V.; formal analysis, J.L.P.-G., D.S.L., E.A.-C., P.V., P.M., M.M., A.N.N. and A.L.P.; investigation, J.L.P.-G., D.S.L., E.A.-C., P.V., A.N.N. and A.L.P.; resources, E.S., P.M., M.M., A.N.N. and A.L.P.; data curation, P.M., M.M., A.N.N. and A.L.P.; writing—original draft preparation, P.M., M.M., A.N.N. and A.L.P.; writing—review and editing, J.L.P.-G., D.S.L., E.A.-C., P.V., R.P.-M., E.S., P.M., M.M., A.N.N. and A.L.P.; visualization, P.M., M.M., A.N.N. and A.L.P.; supervision, P.M., M.M., A.N.N. and A.L.P.; project administration, P.M., M.M., A.N.N. and A.L.P.; funding acquisition, E.S., P.M., M.M., A.N.N. and A.L.P. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are contained within the article and supplementary materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Naganathan, A.N. Modulation of Allosteric Coupling by Mutations: From Protein Dynamics and Packing to Altered Native Ensembles and Function. *Curr. Opin. Struct. Biol.* **2019**, *54*, 1–9.
2. Pacheco-Garcia, J.L.; Anoz-Carbonell, E.; Vankova, P.; Kannan, A.; Palomino-Morales, R.; Mesa-Torres, N.; Salido, E.; Man, P.; Medina, M.; Naganathan, A.N.; et al. Structural Basis of the Pleiotropic and Specific Phenotypic Consequences of Missense Mutations in the Multifunctional NAD(P)H: Quinone Oxidoreductase 1 and Their Pharmacological Rescue. *Redox Biol.* **2021**, *46*, 102112. https://doi.org/10.1016/j.redox.2021.102112.
3. Rajasekaran, N.; Suresh, S.; Gopi, S.; Raman, K.; Naganathan, A.N. A General Mechanism for the Propagation of Mutational Effects in Proteins. *Biochemistry* **2017**, *56*, 294–305. https://doi.org/10.1021/acs.biochem.6b00798.
4. Rajasekaran, N.; Sekhar, A.; Naganathan, A.N. A Universal Pattern in the Percolation and Dissipation of Protein Structural Perturbations. *J. Phys. Chem. Lett.* **2017**, *8*, 4779–4784. https://doi.org/10.1021/acs.jpclett.7b02021.
5. Beaver, S.K.; Mesa-Torres, N.; Pey, A.L.; Timson, D.J. NQO1: A Target for the Treatment of Cancer and Neurological Diseases, and a Model to Understand Loss of Function Disease Mechanisms. *Biochim. Biophys. Acta—Proteins Proteom.* **2019**, *1867*, 663–676.
6. Vankova, P.; Salido, E.; Timson, D.J.; Man, P.; Pey, A.L. A Dynamic Core in Human NQO1 Controls the Functional and Stability Effects of Ligand Binding and Their Communication across the Enzyme Dimer. *Biomolecules* **2019**, *9*, 728. https://doi.org/10.3390/biom9110728.

7.   Megarity, C.F.; Abdel-Aal Bettley, H.; Caraher, M.C.; Scott, K.A.; Whitehead, R.C.; Jowitt, T.A.; Gutierrez, A.; Bryce, R.A.; Nolan, K.A.; Stratford, I.J.; et al. Negative Cooperativity in NAD(P)H Quinone Oxidoreductase 1 (NQO1). *ChemBioChem* **2019**, *20*, 2841–2849. https://doi.org/10.1002/cbic.201900313.

8.   Pey, A.L.; Megarity, C.F.; Timson, D.J. NAD(P)H Quinone Oxidoreductase (NQO1): An Enzyme Which Needs Just Enough Mobility, in Just the Right Places. *Biosci. Rep.* **2019**, *39*, BSR20180459.

9.   Ross, D.; Siegel, D. The Diverse Functionality of NQO1 and Its Roles in Redox Control. *Redox Biol.* **2021**, *41*, 101950.

10.   Li, R.; Bianchet, M.A.; Talalayt, P.; Amzel, L.M. The Three-Dimensional Structure of NAD(P)H:Quinone Reductase, a Flavoprotein Involved in Cancer Chemoprotection and Chemotherapy: Mechanism of the Two-Electron. Reduction (x-Ray Diffraction/Flavin). *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8846–8850.

11.   Lienhart, W.D.; Gudipati, V.; Uhl, M.K.; Binter, A.; Pulido, S.A.; Saf, R.; Zangger, K.; Gruber, K.; Macheroux, P. Collapse of the Native Structure Caused by a Single Amino Acid Exchange in Human NAD(P)H: Quinone Oxidoreductase. *FEBS J.* **2014**, *281*, 4691–4704. https://doi.org/10.1111/febs.12975.

12.   Faig, M.; Bianchet, M.A.; Talalay, P.; Chen, S.; Winski, S.; Ross, D.; Amzel, L.M. Structures of Recombinant Human and Mouse NAD(P)H: Quinone Oxidoreductases: Species Comparison and Structural Changes with Substrate Binding and Release. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3177–3182.

13.   Medina-Carmona, E.; Neira, J.L.; Salido, E.; Fuchs, J.E.; Palomino-Morales, R.; Timson, D.J.; Pey, A.L. Site-to-Site Interdomain Communication May Mediate Different Loss-of-Function Mechanisms in a Cancer-Associated NQO1 Polymorphism. *Sci. Rep.* **2017**, *7*, 44532. https://doi.org/10.1038/srep44532.

14.   Encarnación, M.C.; Palomino-Morales, R.J.; Fuchs, J.E.; Esperanza, P.G.; Noel, M.T.; Salido, E.; Timson, D.J.; Pey, A.L. Conformational Dynamics Is Key to Understanding Loss-of-Function of NQO1 Cancer-Associated Polymorphisms and Its Correction by Pharmacological Ligands. *Sci. Rep.* **2016**, *6*, 20331. https://doi.org/10.1038/srep20331.

15.   Anoz-Carbonell, E.; Timson, D.J.; Pey, A.L.; Medina, M. The Catalytic Cycle of the Antioxidant and Cancer-Associated Human NQO1 Enzyme: Hydride Transfer, Conformational Dynamics and Functional Cooperativity. *Antioxidants* **2020**, *9*, 772. https://doi.org/10.3390/antiox9090772.

16.   Xue, M.; Wakamoto, T.; Kejlberg, C.; Yoshimura, Y.; Nielsen, T.A.; Risør, M.W.; Sanggaard, K.W.; Kitahara, R.; Mulder, F.A.A. How Internal Cavities Destabilize a Protein. *Proc. Natl. Acad. Sci USA* **2019**, *116*, 21031–21036. https://doi.org/10.1073/pnas.1911181116.

17.   Medina-Carmona, E.; Fuchs, J.E.; Gavira, J.A.; Mesa-Torres, N.; Neira, J.L.; Salido, E.; Palomino-Morales, R.; Burgos, M.; Timson, D.J.; Pey, A.L. Enhanced Vulnerability of Human Proteins towards Disease-Associated Inactivation through Divergent Evolution. *Hum. Mol. Genet.* **2017**, *26*, 3531–3544. https://doi.org/10.1093/hmg/ddx238.

18.   Pey, A.L.; Megarity, C.F.; Timson, D.J. FAD Binding Overcomes Defects in Activity and Stability Displayed by Cancer-Associated Variants of Human NQO1. *Biochim. Biophys. Acta—Mol. Basis Dis.* **2014**, *1842*, 2163–2173. https://doi.org/10.1016/j.bbadis.2014.08.011.

19.   Pacheco-García, J.L.; Cano-Muñoz, M.; Sánchez-Ramos, I.; Salido, E.; Pey, A.L. Naturally-Occurring Rare Mutations Cause Mild to Catastrophic Effects in the Multifunctional and Cancer-Associated NQO1 Protein. *J. Pers. Med.* **2020**, *10*, 207. https://doi.org/10.3390/jpm10040207.

20.   Sánchez-Azqueta, A.; Catalano-Dupuy, D.L.; López-Rivero, A.; Tondo, M.L.; Orellano, E.G.; Ceccarelli, E.A.; Medina, M. Dynamics of the Active Site Architecture in Plant-Type Ferredoxin-NADP + Reductases Catalytic Complexes. *Biochim. Biophys. Acta—Bioenerg.* **2014**, *1837*, 1730–1738. https://doi.org/10.1016/j.bbabio.2014.06.003.

21.   Trcka, F.; Durech, M.; Man, P.; Hernychova, L.; Muller, P.; Vojtesek, B. The Assembly and Intermolecular Properties of the Hsp70-Tomm34-Hsp90 Molecular Chaperone Complex. *J. Biol. Chem.* **2014**, *289*, 9887–9901. https://doi.org/10.1074/jbc.M113.526046.

22.   Zhang, Z.; Smith, D.L. Determination of Amide Hydrogen Exchange by Mass Spectrometry: A New Tool for Protein Structure Elucidation. *Protein Sci.* **1993**, *2*, 522–531. https://doi.org/10.1002/pro.5560020404.

23.   Naganathan, A.N.; Dani, R.; Gopi, S.; Aranganathan, A.; Narayan, A. Folding Intermediates, Heterogeneous Native Ensembles and Protein Function. *J. Mol. Biol.* **2021**, *433*, 167325. https://doi.org/10.1016/j.jmb.2021.167325.

24.   Gopi, S.; Aranganathan, A.; Naganathan, A.N. Thermodynamics and Folding Landscapes of Large Proteins from a Statistical Mechanical Model. *Curr. Res. Struct. Biol.* **2019**, *1*, 6–12. https://doi.org/10.1016/j.crstbi.2019.10.002.

25.   Asher, G.; Dym, O.; Tsvetkov, P.; Adler, J.; Shaul, Y. The Crystal Structure of NAD(P)H Quinone Oxidoreductase 1 in Complex with Its Potent Inhibitor Dicoumarol. *Biochemistry* **2006**, *45*, 6372–6378. https://doi.org/10.1021/bi0600087.

26.   Naganathan, A.N.; Kannan, A. A Hierarchy of Coupling Free Energies Underlie the Thermodynamic and Functional Architecture of Protein Structures. *Curr. Res. Struct. Biol.* **2021**, *3*, 257–267. https://doi.org/10.1016/j.crstbi.2021.09.003.

27.   DeLano, W.L. *The PyMOL Molecular Graphics System*; Schrodinger LLC: New York, NY, USA, 2002.

28.   Pey, A.L. Biophysical and Functional Perturbation Analyses at Cancer-Associated P187 and K240 Sites of the Multifunctional NADP(H): Quinone Oxidoreductase 1. *Int. J. Biol. Macromol.* **2018**, *118*, 1912–1923. https://doi.org/10.1016/j.ijbiomac.2018.07.051.

29.   Nagel, Z.D.; Klinman, J.P. Update 1 of: Tunneling and Dynamics in Enzymatic Hydride Transfer. *Chem. Rev.* **2010**, *110*, PR41–PR67. https://doi.org/10.1021/cr1001035.

Supplementary Materials: Allosteric Communication in the Multifunctional and Redox NQO1 Protein Studied by Cavity-Making Mutations

Juan Luis Pacheco-Garcia [1,*,†], Dmitry S. Loginov [2,†], Ernesto Anoz-Carbonell [3,†], Pavla Vankova [4,5], Rogelio Palomino-Morales [6], Eduardo Salido [7], Petr Man [2], Milagros Medina [3], Athi N. Naganathan [8] and Angel L. Pey [9,*]

[1] Departamento de Química Física, Universidad de Granada, Av. Fuentenueva s/n, 18071 Granada, Spain
[2] Institute of Microbiology—BioCeV, Academy of Sciences of the Czech Republic, Prumyslova 595, 252 50 Vestec, Czech Republic; dmitry.loginov@biomed.cas.cz (D.S.L.); pman@biomed.cas.cz (P.M.)
[3] Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias, Instituto de Biocomputación y Física de Sistemas Complejos (BIFI) (GBsC-CSIC Joint Unit), Universidad de Zaragoza, 50009 Zaragoza, Spain; eanoz@unizar.es (E.A.-C.); mmedina@unizar.es (M.M.)
[4] Institute of Biotechnology—BioCeV, Academy of Sciences of the Czech Republic, Prumyslova 595, 252 50 Vestec, Czech Republic.; pavla.vankova@ibt.cas.cz or pavla.vankova@biomed.cas.cz
[5] Department of Biochemistry, Faculty of Science, Charles University, Hlavova 2030/8, 128 43 Prague, Czech Republic
[6] Departamento de Bioquímica y Biología Molecular I, Facultad de Ciencias y Centro de Investigaciones Biomédicas (CIBM), Universidad de Granada, 18016 Granada, Spain; rpm@ugr.es
[7] Center for Rare Diseases (CIBERER), Hospital Universitario de Canarias, Universidad de la Laguna, 38320 Tenerife, Spain; edsalido@gmail.com
[8] Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras (IITM), Chennai 600036, India; athi@ijtm.ac.in
[9] Departamento de Química Física, Unidad de Excelencia en Química Aplicada a Biomedicina y Medioambiente e Instituto de Biotecnología, Universidad de Granada, Av. Fuentenueva s/n, 18071 Granada, Spain
* Correspondence: juanlupacheco@correo.ugr.es (J.L.P.-G.); angelpey@ugr.es (A.L.P.); Tel.: +34-958243173 (A.L.P.)
† These authors contributed equally to this work.
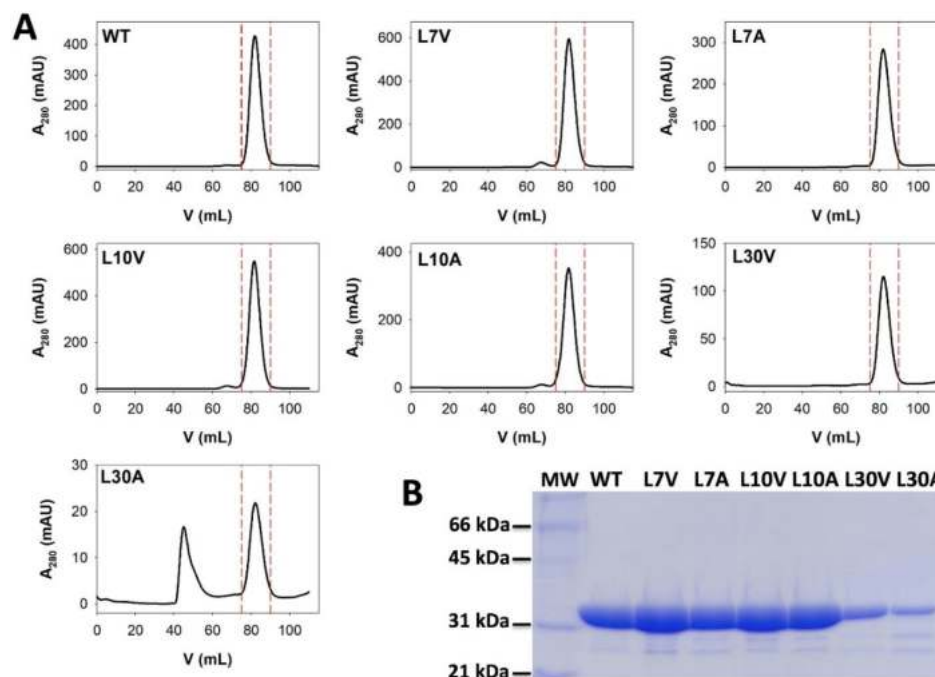
Figure S1. Purification of NQO1 proteins containing mutations at L7, L10 and L30. A) SEC profiles of NQO1 proteins previously isolated by IMAC. Each experiment corresponds to the chromatographic separation of the protein isolated by IMAC from 2.4-4.8 L of bacterial culture.

Chromatography was carried out in a HiLoad® 16/600 Superdex® 200 prep grade column (GE Healthcare) using 20 mM HEPES-KOH   200 mM NaCl at pH 7.4 as mobile phase. Vertical dashed lines correspond to the fractions (volume 75-90 mL) collected and containing NQO1 dimeric fraction. B) SDS-PAGE analyses (12 % acrylamide) of NQO1 dimeric fractions obtained from profiles shown in panel A.
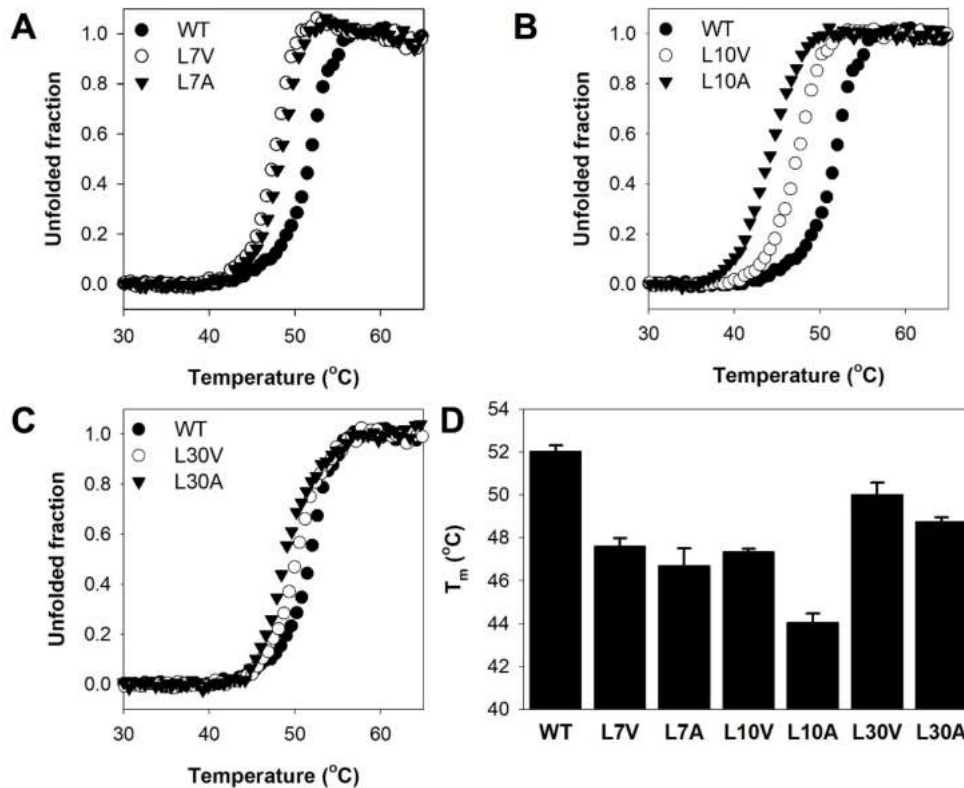
Figure S2. Thermal stability of NQO1 cavity-making mutants.   Panels A-C show representative experimental thermal denaturation experiments. Panel D displays the average ± s.d. from at least four different replicates. The experiments were carried out using 2 μM of proteins (in monomer) and 100 μM FAD.
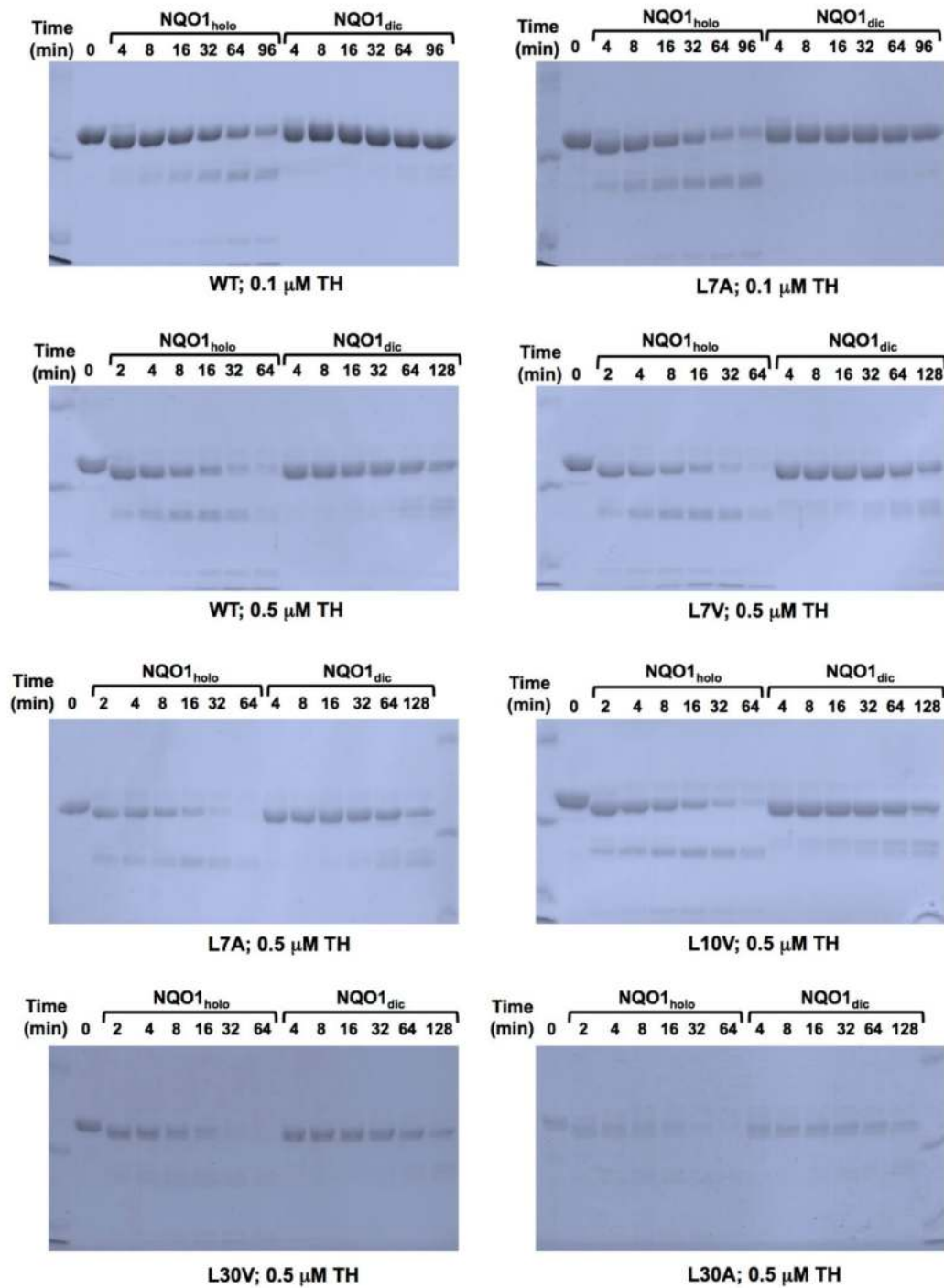
Figure S3. Representative SDS-PAGE analysis for the proteolytic kinetics of NQO1 variants with thermolysin. NQO1 samples were prepared at 10 μM NQO1 protein concentration in the presence of 100 μM FAD (NQO1holo) or 100 μM FAD + 100 μM Dic (NQO1dic). Thermolysin concentration is indicated for each experiment.
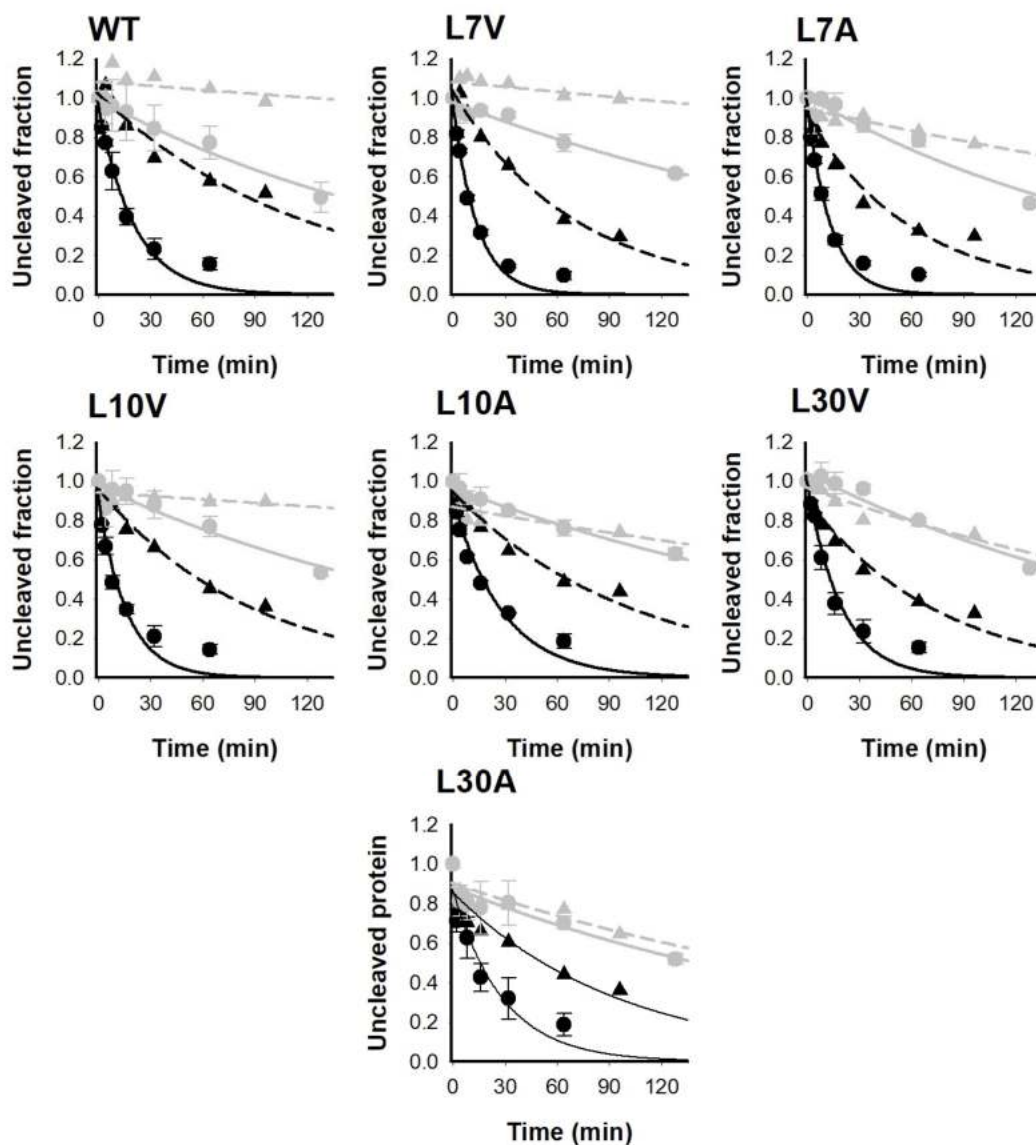
Figure S4. Proteolysis kinetics of NQO1 variants with thermolysin. NQO1holo is shown as black symbols and NQO1dic as grey symbols. Circles indicate a 0.5 μM thermolysin concentration (mean±s.d. from three experiments) and triangles show a 0.1 μM thermolysin concentration (from a single experiment). Lines are fittings using an exponential function.
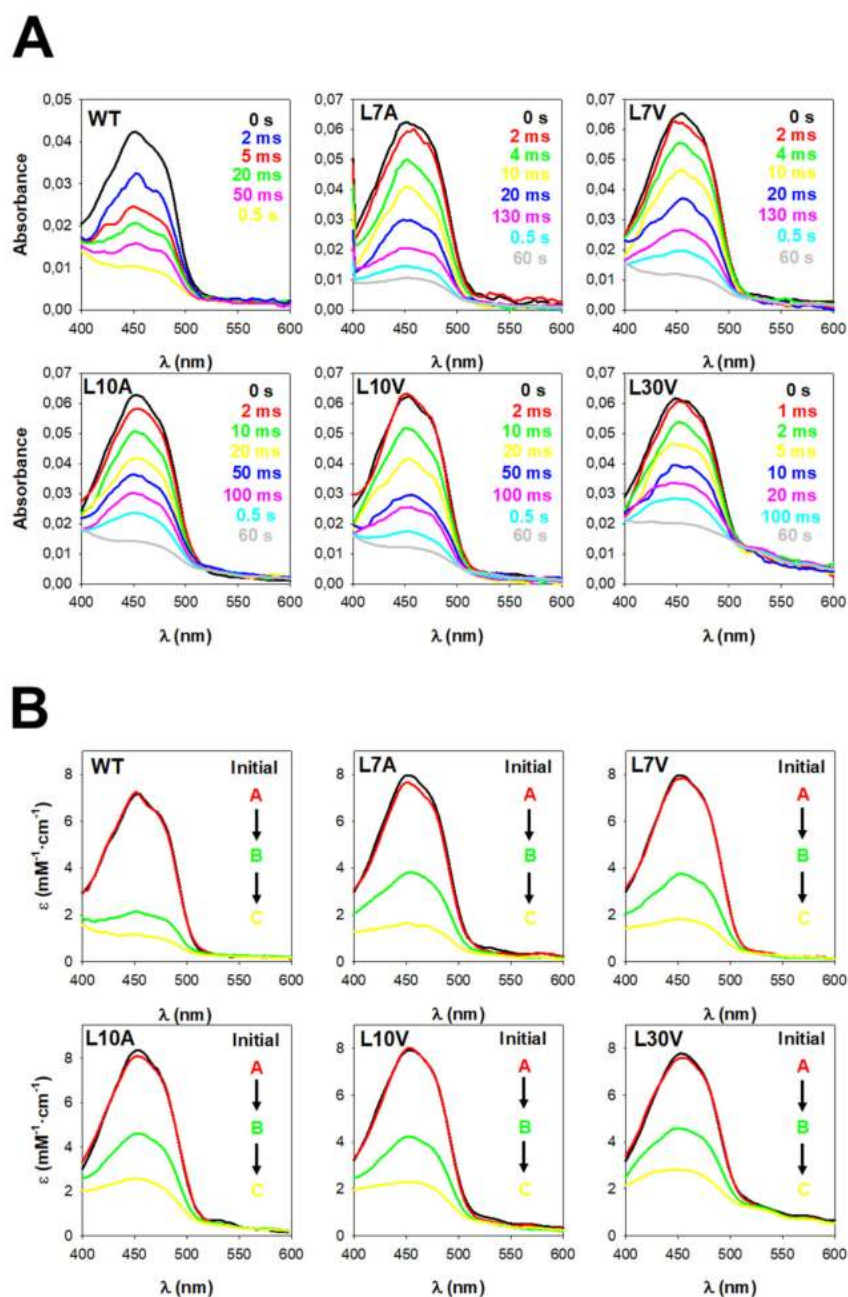
Figure S5. Reductive half-reaction of FAD bound to NQO1 variants with NADH. A) Spectral evolution in a 0-60 s timescale after mixing holo-NQO1 (7.5 μM) with NADH (7.5 μM) in 20 mM HEPES-KOH, pH 7.4, at 6 °C. Different colored lines correspond to the spectra at different reaction times. B) Spectral deconvolution of different species observed during the reaction when fitting to a three-state model. Different colored lines correspond to the different species along the reaction.
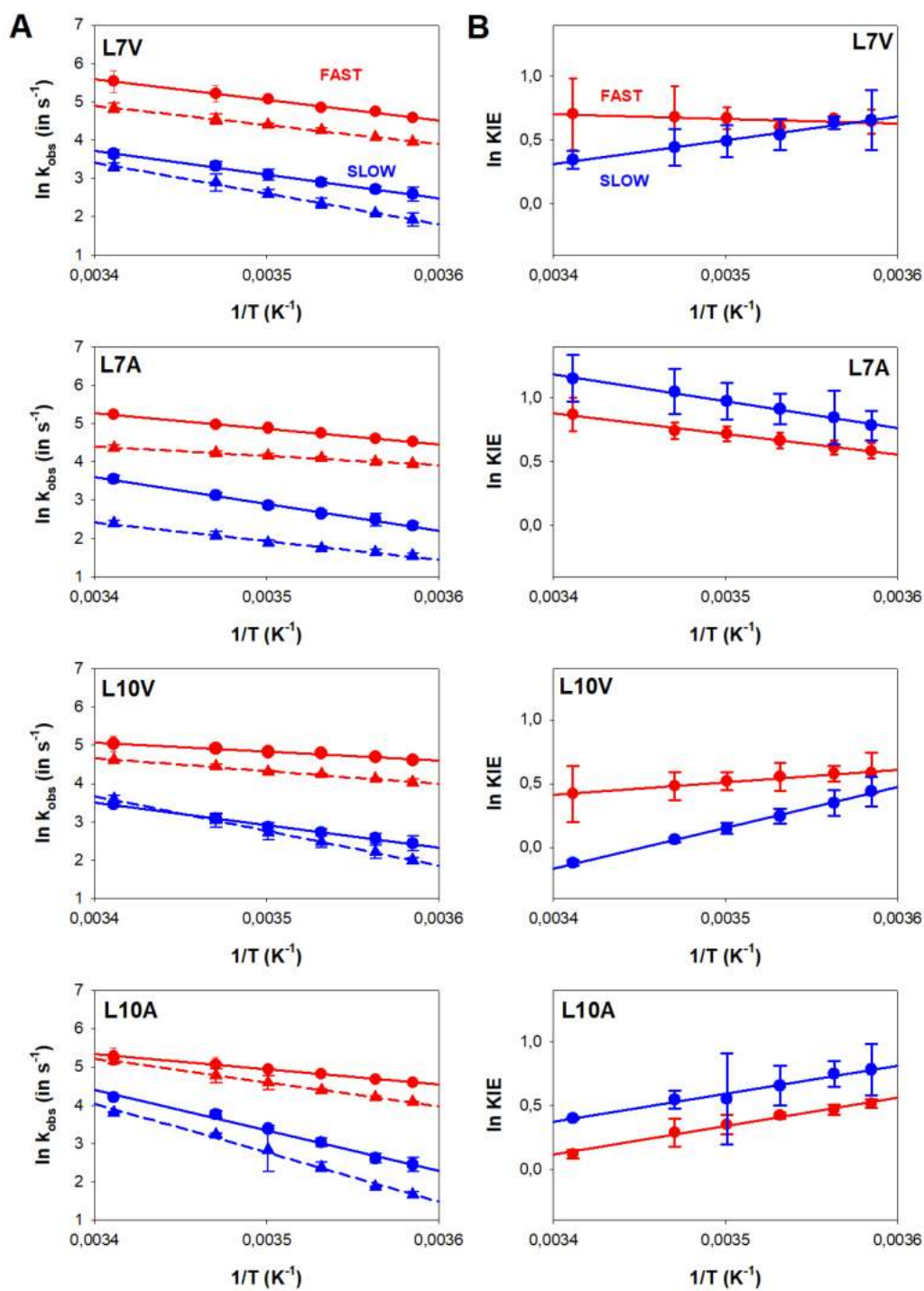
Figure S6. Temperature dependence of kinetic parameters for the two hydride/deuteride transfer (HT/DT) processes from NADH to NQO1. (A) Arrhenius plots of kinetic constants. Data for HT and DT are shown as circles and triangles, respectively. (B) Temperature dependence of the kinetic isotope effects (KIEs).

Table S1. Observed rate constants (k_{obs}) for partial proteolysis of NQO1 variants with thermolysin.

Variant	Species	Thermolysin concentration (µM)	k_{obs} (min^{-1})
WT	Holo	0.1	$8.4 \pm 1.2 \cdot 10^{-3}$
	Holo	0.5	$4.8 \pm 0.6 \cdot 10^{-2}$
	Holo+Dic	0.1	$6.4 \pm 8.0 \ 10^{-4}$
	Holo+Dic	0.5	$5.0 \pm 0.4 \cdot 10^{-3}$
L7V	Holo	0.1	$1.4 \pm 0.1 \cdot 10^{-2}$
	Holo	0.5	$7.2 \pm 0.8 \cdot 10^{-2}$
	Holo+Dic	0.1	$8.1 \pm 4.8 \cdot 10^{-4}$
	Holo+Dic	0.5	$3.5 \pm 0.4 \cdot 10^{-3}$
L7A	Holo	0.1	$1.6 \pm 0.3 \cdot 10^{-2}$
	Holo	0.5	$7.4 \pm 0.9 \cdot 10^{-2}$
	Holo+Dic	0.1	$2.1 \pm 0.5 \cdot 10^{-3}$
	Holo+Dic	0.5	$5.2 \pm 0.8 \cdot 10^{-3}$
L10V	Holo	0.1	$1.1 \pm 0.1 \cdot 10^{-2}$
	Holo	0.5	$6.1 \pm 1.2 \cdot 10^{-2}$
	Holo+Dic	0.1	$6.5 \pm 3.9 \cdot 10^{-4}$
	Holo+Dic	0.5	$4.3 \pm 0.7 \cdot 10^{-3}$
L10A	Holo	0.1	$9.6 \pm 1.1 \cdot 10^{-3}$
	Holo	0.5	$3.5 \pm 0.6 \cdot 10^{-2}$
	Holo+Dic	0.1	$1.8 \pm 1.1 \cdot 10^{-3}$
	Holo+Dic	0.5	$3.5 \pm 0.3 \cdot 10^{-3}$
L30V	Holo	0.1	$1.3 \pm 0.2 \cdot 10^{-2}$
	Holo	0.5	$5.1 \pm 0.7 \cdot 10^{-2}$
	Holo+Dic	0.1	$3.3 \pm 0.5 \cdot 10^{-3}$
	Holo+Dic	0.5	$4.4 \pm 0.5 \cdot 10^{-3}$
L30A	Holo	0.1	$1.0 \pm 0.2 \cdot 10^{-2}$
	Holo	0.5	$7.4 \pm 1.0 \cdot 10^{-2}$

Holo+Dic	0.1	$3.3 \pm 0.9 \cdot 10^{-3}$
Holo+Dic	0.5	$4.1 \pm 1.0 \cdot 10^{-3}$

Table S2. Arrhenius parameters and KIEs for the HT/DT in the reduction of NQO1 variants by NADH/NADD. All values correspond to data obtained with equimolecular concentrations of the reactants in the stopped-flow equipment. (n > 3, mean ± s.d.). Analysis were performed as previously described (Anoz-Carbonell et al., 2020)..

Step	HT			DT			KIE [a]	ΔE_{aDT-HT} (kcal·mol^{-1})	A_{HT}/A_{DT}
	k_{obsHT} [a] (s^{-1})	E_{aHT} (kcal·mol^{-1})	A_{HT} (s^{-1})	k_{obsDT} [a] (s^{-1})	E_{aDT} (kcal·mol^{-1})	A_{DT} (s^{-1})			
FAST									
WT[b]	78±1	6.1±0.2	$5.3 \pm 1.2 \cdot 10^{6}$	44±2	6.3±0.4	$4.1 \pm 1.1 \cdot 10^{6}$	1.8±0.1	0.2±0.1	1.3±0.6
L7V	98±11	10.7±0.4	$2.4 \pm 0.1 \cdot 10^{10}$	52±2	10.0±0.2	$3.6 \pm 0.1 \cdot 10^{9}$	1.9±0.3	-0.7±0.1	6.7±0.5
L7A	92±8	8.0±0.2	$2.0 \pm 0.1 \cdot 10^{8}$	51±1	4.8±0.1	$3.3 \pm 0.1 \cdot 10^{5}$	1.8±0.2	-3.2±0.2	$6.1 \pm 0.5 \cdot 10^{2}$
L10V	100±6	4.7±0.3	$5.4 \pm 0.2 \cdot 10^{5}$	56±5	6.7 ±0.2	$9.8 \pm 0.1 \cdot 10^{6}$	1.8±0.3	2±0.2	$5.5 \pm 0.2 \cdot 10^{2}$
L10A	99±5	8.0±0.2	$1.6 \pm 0.1 \cdot 10^{8}$	59±1	12.4 ±0.2	$2.9 \pm 0.1 \cdot 10^{11}$	1.7±0.1	4.4±0.2	$5.5 \pm 0.5 \cdot 10^{4}$
SLOW									
WT[b]	8.9±0.9	10.9±0.5	$3.4 \pm 0.9 \cdot 10^{9}$	5.3±0.2	9.8±0.5	$2.6 \pm 0.6 \cdot 10^{8}$	1.8±0.3	-1.1±0.2	13±6
L7V	13.2±2.4	12.3±0.4	$5.9 \pm 0.2 \cdot 10^{10}$	6.9±1.2	16.0±0.4	$2.2 \pm 0.1 \cdot 10^{13}$	1.9±0.6	3.7±0.2	$2.6 \pm 0.2 \cdot 10^{3}$
L7A	10.3±0.8	14.0±0.5	$8.0 \pm 0.3 \cdot 10^{11}$	4.7±0.3	9.5±0.5	$2.0 \pm 0.1 \cdot 10^{8}$	2.2±0.3	-4.5±0.4	$4.0 \pm 0.3 \cdot 10^{3}$
L10V	11.3±2.2	11.7±0.5	$1.6 \pm 0.1 \cdot 10^{10}$	7.3±0.5	18.1±0.4	$9.6 \pm 0.1 \cdot 10^{14}$	1.5±0.4	6.4±0.4	$1.7 \pm 0.1 \cdot 10^{2}$
L10A	11.6±2.2	21.0±1.1	$3.2 \pm 0.2 \cdot 10^{17}$	5.3±0.4	25.4±0.1	$4.3 \pm 0.2 \cdot 10^{20}$	2.2±0.6	4.4±0.3	$7.4 \pm 0.8 \cdot 10^{4}$

[a] Experiments carried out at 6ºC. [b] From (Anoz-Carbonell et al., 2020).

Supplementary references

1. Anoz-Carbonell, E., Timson, D. J., Pey, A. L., & Medina, M. (2020a). The catalytic cycle of the antioxidant and cancer-associated human NQO1 enzyme: Hydride transfer, conformational dynamics and functional cooperativity. Antioxidants, 9(9), 1–22. https://doi.org/10.3390/antiox9090772

ANEXO II

Effect of naturally-occurring mutations on the stability and function of cancer-associated NQO1: Comparison of experiments and computation

Effect of naturally-occurring mutations on the stability and function of cancer-associated NQO1: Comparison of experiments and computation

Juan Luis Pacheco-Garcia[1†], Matteo Cagiada[2†],
Kelly Tienne-Matos[1], Eduardo Salido[3], Kresten Lindorff-Larsen[2‡]
and Angel L. Pey[4*‡]

[1]Departamento de Química-Física, Universidad de Granada, Granada, Spain, [2]Department of Biology, Linderstrøm-Lang Centre for Protein Science, University of Copenhagen, Copenhagen, Denmark, [3]Center for Rare Diseases (CIBERER), Hospital Universitario de Canarias, Universidad de la Laguna, La Laguna, Tenerife Tenerife, Spain, [4]Departamento de Química Física, Unidad de Excelencia en Química Aplicada a Biomedicina y Medioambiente e Instituto de Biotecnología, Universidad de Granada, Granada, Spain

Recent advances in DNA sequencing technologies are revealing a large individual variability of the human genome. Our capacity to establish genotype-phenotype correlations in such large-scale is, however, limited. This task is particularly challenging due to the multifunctional nature of many proteins. Here we describe an extensive analysis of the stability and function of naturally-occurring variants (found in the COSMIC and gnomAD databases) of the cancer-associated human NAD(P)H:quinone oxidoreductase 1 (NQO1). First, we performed *in silico* saturation mutagenesis studies (>5,000 substitutions) aimed to identify regions in NQO1 important for stability and function. We then experimentally characterized twenty-two naturally-occurring variants in terms of protein levels during bacterial expression, solubility, thermal stability, and coenzyme binding. These studies showed a good overall correlation between experimental analysis and

computational predictions; also the magnitude of the effects of the substitutions are similarly distributed in variants from the COSMIC and gnomAD databases. Outliers in these experimental-computational genotype-phenotype correlations remain, and we discuss these on the grounds and limitations of our approaches. Our work represents a further step to characterize the mutational landscape of NQO1 in the human genome and may help to improve high-throughput *in silico* tools for genotype-phenotype correlations in this multifunctional protein associated with disease.

1 Introduction

Advances in technologies for whole-genome or exome sequencing and high-throughput functional assays have increased our knowledge on the consequences of the genetic variability in humans and the relationship to disease (McInnes et al., 2021; Arnedo-Pac et al., 2022; Høie et al., 2022; Katsonis et al., 2022). However, our capacity to predict the pathogenicity of single amino acid variants is still limited, with some approaches providing good overall results but failing to predict correlation for some individual mutations or phenotypes (Katsonis et al., 2022).

Current approaches for correlating genotype and phenotype can broadly be classified into two classes. First, experimental approaches are based on the characterization of one or several functional features (for example enzymatic function and regulation, protein-protein interactions, transport to different intracellular or extracellular locations, protein turnover, ligand binding) (Xu et al., 2017; Abildgaard et al., 2019; Pacheco-Garcia et al., 2021; Høie et al., 2022). In the case of high-throughput experimental approaches typically only one or two aspects of protein function are analysed (for example protein abundance or ability to rescue a growth phenotype) (Cagiada et al., 2021). Second, the use of structure- or sequence-based methods to predict pathogenicity are becoming increasingly robust (Abildgaard et al., 2019; Arnedo-Pac et al., 2022; Høie et al., 2022). Although experiments may be implemented in a high-throughput fashion, it has until now been limited to a relatively small set of proteins and assays (Arnedo-Pac et al., 2022; Høie et al., 2022). Thus, while computational approaches also have limitations, they may be appealing due to their potential application on a proteomic scale (Arnedo-Pac et al., 2022; Høie et al., 2022).

In this work, we apply both types of approaches to increase our understanding of the correlation between genotype and phenotype for missense variants of the human NAD(P)H: quinone oxidoreductase 1 (NQO1) protein. NQO1 is associated with several diseases including cancer, Alzheimer´s and Parkinson's disease (Beaver et al., 2019; Luo et al., 2019).

NQO1 is a multifunctional protein, displaying both enzymatic and non-enzymatic functions. As an enzyme, it catalyzes the FAD-dependent (two-electron) reduction of a large set of quinone substrates, with functions including redox maintenance of vitamins, detoxification of xenobiotics and activation of cancer pro-drugs (Ross and Siegel, 2018; Beaver et al., 2019; Anoz-Carbonell et al., 2020; Salido et al., 2022). Among non-enzymatic functions, NQO1 may interact with proteins and RNA, controlling their stability and function (Beaver et al., 2019; Asher et al., 2005; Ben-Nissan and Sharon, 2014; di Francesco et al., 2016; Oh et al., 2016). Many of these functions are associated with the catalytic competence and FAD binding, such as protein-protein interactions, intracellular stability and association with microtubules (Asher et al., 2005; Martínez-Limón et al., 2016; Martínez-Limón et al., 2020; Siegel et al., 2021). The native form of NQO1 is dimeric, containing two different domains: an N-terminal domain (NTD, residues 1–225), that tightly binds one FAD molecule/domain required for catalysis and contains most of the monomer-monomer interface (MMI), whereas the C-terminal domain (CTD, residues 225–274) complete the active site and the MMI (Li et al., 1995; Faig et al., 2000; Lienhart et al., 2014; Medina-Carmona et al., 2017a; Pacheco-Garcia et al., 2021).

We have recently shown that ligand binding and variant effects on stability propagate to long distances in the native state, affecting different functional features in a counterintuitive fashion (Medina-Carmona et al., 2016; Pey, 2018; Medina-Carmona et al., 2019a; Vankova et al., 2019; Pacheco-García et al., 2020; Pacheco-Garcia et al., 2021; Pacheco-Garcia et al., 2022a). Therefore, NQO1 represents a challenging and biomedically relevant system to compare the performance of computational and experimental methods to explain and to predict genotype-phenotype in a large-scale for a multi-functional protein. Here, we use computational tools to probe 5,187 variants of NQO1 that includes a set of clinically relevant missense variants which we then experimentally characterized. In this set, thirteen variants come from large-scale human sequencing data (gnomAD) and nine from the catalogue of somatic mutations in human cancer lines (COSMIC)

TABLE 1 Set of NQO1 variants experimentally characterized in this work.

Mutation	Database	% ASA[a]	Variant class	Residue class
G3S	gnomAD	6.0 ± 4.7	WT-like	WT-like
G3D	COSMIC	6.0 ± 4.7	WT-like	WT-like
L7P	COSMIC	0.3 ± 0.2	Total-loss	Total-loss
L7R	gnomAD	0.3 ± 0.2	Total-loss	Total-loss
V9I	gnomAD	0.0 ± 0.0	WT-like	Total-loss
T16M	gnomAD	43 ± 14	Stable but inactive	WT-like
Y20N	gnomAD	21 ± 5	WT-like	WT-like
A29T	COSMIC	2.2 ± 0.5	WT-like	Unstable but active
K32N	gnomAD	79 ± 11	WT-like	WT-like
G34V	gnomAD	54 ± 11	Total-loss	Stable but Inactive
E36K	gnomAD	64 ± 3	WT-like	WT-like
S40L	gnomAD	0.0 ± 0.0	WT-like	Total-loss
D41G	gnomAD	14 ± 2	Total-loss	Stable but inactive
D41Y	COSMIC	14 ± 2	Stable but inactive	Stable but inactive
M45L	COSMIC	28 ± 5	WT-like	WT-like
M45I	COSMIC	28 ± 5	WT-like	WT-like
I51V	gnomAD	14 ± 1	WT-like	Stable but inactive
W106R	gnomAD	10 ± 1	Total-loss	Total-loss
W106C	COSMIC	10 ± 1	Total-loss	Total-loss
F107C	gnomAD	7.6 ± 0.5	Unstable but active	Unstable but active
M155I	COSMIC	13 ± 3	WT-like	WT-like
H162N	COSMIC	6.6 ± 0.5	WT-like	WT-like

The table indicates whether the variants are found in the COSMIC/gnomAD databases as well as the solvent exposure (as % ASA) determined using a crystallographic model of WT NQO1 [PDB code 2F1O (Asher et al., 2006)], the software Getarea and the computational classification at variant and residue level using a combination of predictions of thermodynamic stability change upon mutation and evolutionary conservation.
[a]Using GetArea (http://curie.utmb.edu/getarea.html) and the structure with PDB code 2F1O (Asher et al., 2006). Data are the average ±s.d. from eight monomers.

(Table 1). As of ninth of January 2022, none of these variants were found in both databases. Whether variants found in COSMIC or gnomAD databases are associated with disease (e.g. predisposition to cancer development) is unknown. The set of variants we studied experimentally comprises very different amino acid side chain characteristics and display different levels of solvent exposure (Table 1).

2 Materials and methods

2.1 Experimental methods

2.1.1 Protein expression and purification

Mutations were introduced by site-directed mutagenesis in the wild-type (WT) NQO1 cDNA cloned into the pET-15b vector (pET-15b-NQO1) by GenScript (Leiden, Netherlands). Mutated codons were optimized for expression in *E. coli* and mutagenesis was confirmed by sequencing the entire cDNA. The plasmids were transformed in *E. coli* BL21 (DE3) cells (Agilent Technologies, Santa Clara, CA, United States) for protein expression.

To determine the amount of soluble NQO1 at 37°C, 5 ml of LB medium containing 0.1 mg mL^{-1} ampicillin (purchased from Canvax Biotech, Córdoba, Spain) was inoculated with transformed cells and grown for 16 h at 37°C. 0.5 ml of these cultures were diluted into 10 ml of LB containing 0.1 mg mL^{-1} ampicillin (LBA) and grown at 37°C for 3 h. After that, cultures were induced with 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG, Canvax Biotech) at 37 °C for 4 h. Cells were harvested by centrifugation at 2,900 *g* in a bench centrifuge at 4°C and frozen at −80°C for 16 h. Pellets were resuspended in binding buffer (20 mM Na-phosphate, 300 mM NaCl, 50 mM imidazole, pH 7.4) with 1 mM phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich, Madrid, Spain) and sonicated in an ice bath. These *total extracts* were centrifugated (24,000 *g*, 30 min, 4°C in a bench centrifuge) to obtain the *soluble extracts*. The amount of NQO1 in total and soluble extracts was determined by Western-blotting providing the S/T (soluble/total) ratio for each variant. Samples were denatured using Laemmli's buffer, resolved in polyacrylamide gel electrophoresis in the presence of sodium dodecylsulphate (SDS-PAGE, 12% acrylamide) gels and transferred to PVDF membranes (GE Healthcare, Chicago, IL, United States) using

standard procedures. Immunoblotting was carried out using primary monoclonal antibody against NQO1 (sc-393736, Santa Cruz Biotechnology, Dallas, TX, United States) at 1: 500 dilution and, as secondary antibody, an anti-mouse IgGκ BP-HRP (sc-516102, Santa Cruz Biotechnology) at 1: 2000 dilution was used. Samples were visualized using luminol-based enhanced chemiluminescence (from BioRad Laboratories, Hercules, CA, United States), scanned and analysed using Image Lab (from BioRad Laboratories).

For large-scale purifications, a preculture (100 ml) was prepared from a single clone for each variant and grown for 16 h at 37°C in LBA and diluted into 2.4–4.8 L of LBA. After 3 h at 37°C, NQO1 expression was induced by the addition of 0.5 mM IPTG for 6 h at 25°C. Cells were harvested by centrifugation at 8,000 g and frozen overnight at −80 °C. NQO1 proteins were purified using immobilized nickel affinity chromatography columns (IMAC, GE Healthcare) as described (Anoz-Carbonell et al., 2020). Isolated dimeric fractions of NQO1 variants were exchanged to HEPES-KOH buffer 50 mM pH 7.4 using PD-10 columns (GE Healthcare). The UV–visible spectra of purified NQO1 proteins were measured in a Cary spectrophotometer (Agilent Technologies, Waldbronn, Germany) and used to quantify NQO1 concentration and the content of FAD as described in (Anoz-Carbonell et al., 2020). Apo-proteins were obtained as described in (Vankova et al., 2019). Briefly, holo-proteins were incubated with 2 M urea and 2 M KBr in HEPES-KOH 50 mM pH 7.4 in the presence of 2 mM β-mercaptoethanol and 1 mM PMSF and loaded into IMAC columns, washed with 2 M urea and 2 M KBr in HEPES-KOH 50 mM pH 7.4, 2 mM β-mercaptoethanol, then with HEPES-KOH 50 mM pH 7.4, 2 mM β-mercaptoethanol eluted with 20 mM Na-Phosphate 300 mM NaCl 500 mM imidazole pH 7.4 and finally exchanged to HEPES-KOH 50 mM pH 7.4. These apo-proteins contained less than 2% bound FAD based on UV-visible spectra. Samples were stored at −80°C upon flash freezing in liquid N_2. Protein purity and integrity was checked by SDS-PAGE.

2.1.2 *In vitro* characterization of NQO1 variants

Thermal denaturation of NQO1 proteins, as holo-proteins (2 μM in monomer +100 μM FAD) was monitored by following changes in tryptophan emission fluorescence in HEPES-KOH 50 mM at pH 7.4 as described in (Medina-Carmona et al., 2017b). T_m values were reported as mean ± s.d. of four independent measurements.

Fluorescence titrations were carried out at 25°C using 1 cm × 0.3 cm path-length cuvettes in a Cary Eclipse spectrofluorimeter (Agilent Technologies, Waldbronn, Germany). Experiments were performed in 20 mM K-phosphate, pH 7.4, essentially as described in (Pacheco-García et al., 2020). Briefly, apo-NQO1 (0.25 μM in subunit) was mixed with 0–2 μM FAD in K-phosphate 20 mM pH 7.4. Samples were incubated at 25°C in the dark for at least 10 min before measurements. Spectra were acquired in the 340–360 nm range upon excitation at 280 nm (slits 5 nm), and

spectra were averaged over 10 scans registered at a scan rate of 200 nm min^{-1}. FAD binding fluorescence intensities at 350 nm were fitted using a single and identical type of binding sites as described in (Pacheco-García et al., 2020). Variant effects on the FAD binding free energy ($\Delta\Delta G_{FAD}$) were calculated as:

$$\Delta\Delta G_{FAD} = R \cdot T \cdot \ln\frac{K_{d(mut)}}{K_{d(WT)}} \tag{1}$$

Where R is the ideal gas constant (1.987 cal mol·K^{-1}), T is the experimental temperature (298.15 K), and $K_{d(mut)}$ and $K_{d(WT)}$ are the FAD binding dissociation constant of the mutant and WT protein variants, respectively. A positive value of $\Delta\Delta G_{FAD}$ indicates that the mutation reduces the affinity for FAD.

For proteolysis studies, NQO1 samples (10 μM in monomer) were prepared in HEPES-KOH 50 mM at pH 7.4 in the presence of 100 μM FAD (NQO1$_{holo}$) and incubated at 25°C for 5 min. The proteolysis reaction was initiated upon addition of 0.02–1.2 μM thermolysin (from *Geobacillus stearothermophilus*, Sigma-Aldrich) and a final concentration of 10 mM CaCl$_2$. Samples were incubated at 25°C and aliquots were collected over time and the reaction quenched by addition of EDTA pH 8 (final concentration of 20 mM) and Laemmli´s buffer (2x). Controls (time 0) were prepared likewise but without thermolysin. Samples were resolved by SDS-PAGE under reducing conditions in gels containing 12% acrylamide. Gels were stained with Coomassie blue G-250. Densitometry was carried out using ImageJ. Data were analyzed using an exponential function to provide the apparent rate constant (k_{obs}). From the linear dependence of k_{obs} vs. [thermolysin], we obtained the second-order rate constant k_{prot}. Linearity in these plot indicate that proteolysis occurs under a EX2 mechanism, thus reflecting the thermodynamic stability of the thermolysin cleavage site (Ser72-Val73) between non-cleavable and cleavable states (Park and Marqusee, 2004). These k_{prot} values were used to determine mutational effects on the local stability of thermolysin cleave site ($\Delta\Delta G_{PROT}$) using Eq. 2:

$$\Delta\Delta G_{PROT} = R \cdot T \cdot ln\frac{k_{prot(mut)}}{k_{prot\ (WT)}} \tag{2}$$

Where R is the ideal gas constant (1.987 cal mol·K^{-1}), T is the experimental temperature (298.15 K), and $k_{prot(mut)}$ and $k_{prot(WT)}$ are the second-order proteolysis rate constants of the mutant and WT protein variants, respectively. A positive value of $\Delta\Delta G_{PROT}$ indicates that the mutation thermodynamically destabilizes the thermolysin cleavage site.

2.2 Computational analyses

2.2.1 Evolutionary conservation analysis

We used GEMME (Laine et al., 2019) to evaluate evolutionary distances from the WT NQO1 sequence (Uniprot ID: P15559 — isoform 1). We used HHblits (Remmert et al., 2011; Steinegger et al., 2019) to generate a multiple sequence alignment

(MSA) using UniClust30 as sequence database and an E-value threshold of 10^{-10}. The resulting MSA contained 1,602 sequences and was refined using two additional filters: first, all the columns that were not present in the WT NQO1 sequence were removed; second, all the sequences with more than the 50% of gaps were removed. Application of these two filters yielded 1,414 sequences in the MSA. The GEMME package was run using default parameters. For each position, a median score was evaluated using all the available substitutions.

2.2.2 Thermodynamic stability predictions

Changes in thermodynamic stability ($\Delta\Delta G$) were calculated using the crystal structure (Faig et al., 2000) (PDB 1D4A) and the Rosetta package (GitHub SHA1 *c7009b3115c22daa9efe2805d9d1ebba08426a54*) with the Cartesian $\Delta\Delta G$ protocol (Park et al., 2016; Frenz et al., 2020). The values obtained from Rosetta in internal Energy Unit were divided by 2.9 to bring them on to a scale corresponding to kcal·mol^{-1} (Park et al., 2016). Median scores were evaluated for each position using all the available substitutions.

We used DSSP (Kabsch and Sander, 1983) to calculate the solvent accessible surface area (SASA) when identifying interface residues in NQO1. Interface residues were defined as those residues for a difference larger than 0.2 was detected between SASA calculations based on the dimer and monomer structure.
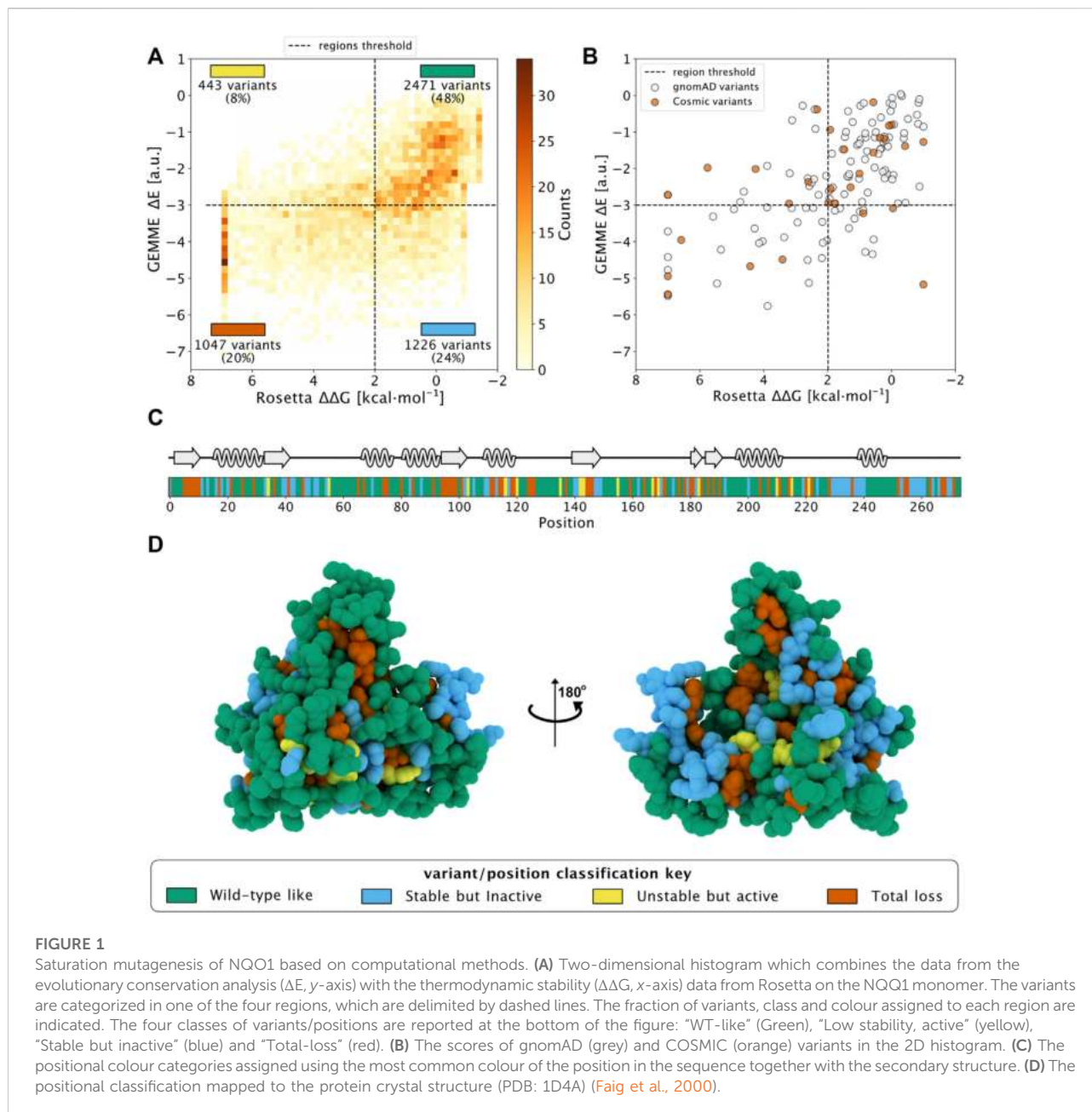
3 Results

3.1 Saturation mutagenesis by computational methods

We first used the predictive ability of evolutionary conservation analysis combined with thermodynamic stability calculations to classify all possible variants (i.e. saturation mutagenesis) in NQO1 based on their effects on the protein function and stability (Cagiada et al., 2021). For evolutionary conservation studies, we used GEMME (Laine et al., 2019) which provides a score (ΔE) for all possible single amino acid change variants of NQO1 (Supplementary Figure S1A). ΔE represents the evolutionary distance of a variant from the WT NQO1 sequence, and ΔE has been shown to be a useful predictor of the deleterious effects on function and stability of the given substitution. We used Rosetta (Park et al., 2016) to predict variant effects on thermodynamic stability ($\Delta\Delta G$) using a crystal structure of NQO1 (Faig et al., 2000) as input (Supplementary Figure S1B) and subsequently calculated the median $\Delta\Delta G$ for all variants at each position. We performed $\Delta\Delta G$ evaluations using both the monomeric and dimeric structure of NQO1 to separate effects on overall stability and effects on dimerization. Specifically, we calculated $\Delta\Delta G$ from the monomer (Supplementary Figures S1B,S2A) to predict the

change in thermodynamic stability relative to wild type of each variant. We also performed similar $\Delta\Delta G$ calculations using the dimer structure as input (Supplementary Figures S1C,S2B), introducing each missense variant in both chains (i.e., treating this as a homodimer) and used the resulting values to compare with experiments. Based on these two calculations, we also evaluated the $\Delta\Delta G$ of dimerization as the difference between the two Rosetta runs (Supplementary Figures S2C,D) to highlight which residues are involved in stabilizing the dimer and thus also those variants that might weaken dimer formation. Then, we evaluated the difference in the SASA between the dimeric and monomeric residues of NQO1 (Supplementary Figure S2C) and we classified 33 of them as interface residues. We found that for 20 of these interface residues the median $\Delta\Delta G$ of dimerization was >1 kcal mol^{-1}. Of these 20, 15 were stable upon mutation in the monomeric form (median $\Delta\Delta G$ <2 kcal mol^{-1}) and a subset of seven display a median $\Delta\Delta G$ of dimerization >2 kcal mol^{-1}.

Then, we combined the evolutionary conservation scores and stability calculations based on the monomeric protein for the 5,187 variants of NQO1 and plotted the results in a two-dimensional histogram (Figure 1A). We used cutoff values of 2 kcal mol^{-1} for Rosetta $\Delta\Delta G$ values and -3 for GEMME ΔE scores as thresholds for all the variants in order to separate them based on their effects (Luo et al., 2019). To ease analyses and interpretations, we associated each of the four defined regions with a color (Cagiada et al., 2021). 'WT-like' variants represent 48% of the available NQO1 variants (shown in green). 20% of NQO1 substitutions show high $\Delta\Delta G$ and high evolutionary distances and are classified as "Total-loss." These variants have substitutions that are unlikely in the evolutionary analysis (ΔE < −3) and lead to decreased stability ($\Delta\Delta G$ >2 kcal mol^{-1}); they thus likely compromise protein function *via* loss of protein stability (colored in red). Variants with high negative ΔE and low $\Delta\Delta G$ belong to the "Stable but inactive" class (colored in blue). This class contains 24% of the variants and represent those for which the evolutionary and stability analysis suggests that the protein function has been compromised, but not for stability reasons. Lastly, the remaining 8% of the variants show low stability and low evolutionary distance, and were classified as "Unstable but active" (colored in yellow). Having predicted the effects of all possible missense variants, we performed a similar classification of amino acid positions, assigning the most common variant class to each position (Figures 1C,D) and found 48% of the total positions classified as "WT-like," 25% as "-Total-loss," 22% as "Stable but inactive" and 5% as "Unstable but active."

In addition, we used the data from the dimer analysis to evaluate the number of residues involved in the stabilization of the dimer form. We found that 14 residues at the interface changed their classification to "Total-loss" if $\Delta\Delta G$ was evaluated using the dimer structure. Of these 14, 9 were classified as "Stable but inactive," while 5 were classified as
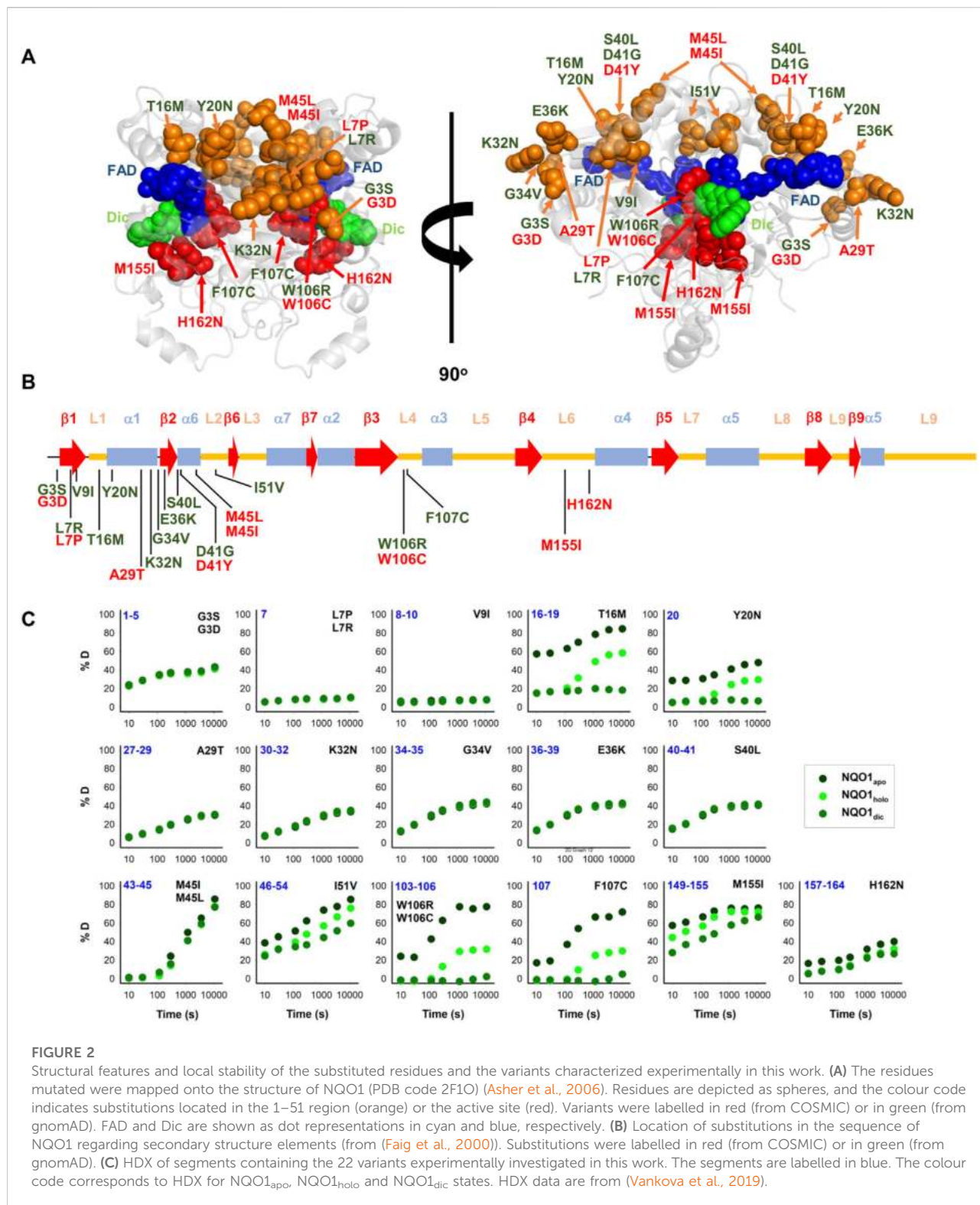
FIGURE 1
Saturation mutagenesis of NQO1 based on computational methods. **(A)** Two-dimensional histogram which combines the data from the evolutionary conservation analysis (ΔE, y-axis) with the thermodynamic stability (ΔΔG, x-axis) data from Rosetta on the NQO1 monomer. The variants are categorized in one of the four regions, which are delimited by dashed lines. The fraction of variants, class and colour assigned to each region are indicated. The four classes of variants/positions are reported at the bottom of the figure: "WT-like" (Green), "Low stability, active" (yellow), "Stable but inactive" (blue) and "Total-loss" (red). **(B)** The scores of gnomAD (grey) and COSMIC (orange) variants in the 2D histogram. **(C)** The positional colour categories assigned using the most common colour of the position in the sequence together with the secondary structure. **(D)** The positional classification mapped to the protein crystal structure (PDB: 1D4A) (Faig et al., 2000).

WT-like using monomeric ΔΔG data. Thus, many residues at the interface appear to be conserved during evolution to preserve the stability of the dimeric form of NQO1.

Having analyzed all possible missense variants, we next looked at the results for a subset of variants that have been found in the human population. Specifically, we looked at variants that are found in the COSMIC (COSMIC v.92; https://cancer.sanger.ac.uk/cosmic) or gnomAD (gnomAD v.2. 1.1; https://gnomad.broadinstitute.org/) databases, and did not find clear differences between these two sets (Figure 1B). In particular, we found variants in both sets that would be predicted

as functional and others for which stability and/or conservation analyses predict loss-of-function (LoF). This result is in line with the notion that both databases may contain both potentially pathogenic as well as benign variants.

3.2 Selection of NQO1 variants to be experimentally characterized

After studying the NQO1 variants computationally, we next examined a set of the variants using a series of different

FIGURE 2
Structural features and local stability of the substituted residues and the variants characterized experimentally in this work. **(A)** The residues mutated were mapped onto the structure of NQO1 (PDB code 2F1O) (Asher et al., 2006). Residues are depicted as spheres, and the colour code indicates substitutions located in the 1–51 region (orange) or the active site (red). Variants were labelled in red (from COSMIC) or in green (from gnomAD). FAD and Dic are shown as dot representations in cyan and blue, respectively. **(B)** Location of substitutions in the sequence of NQO1 regarding secondary structure elements (from (Faig et al., 2000)). Substitutions were labelled in red (from COSMIC) or in green (from gnomAD). **(C)** HDX of segments containing the 22 variants experimentally investigated in this work. The segments are labelled in blue. The colour code corresponds to HDX for $NQO1_{apo}$, $NQO1_{holo}$ and $NQO1_{dic}$ states. HDX data are from (Vankova et al., 2019).

experiments. In this study, we have thus extended our previous work on 8 naturally-occurring variants in NQO1 (Pacheco-García et al., 2020) to a set of 22 variants (Table 1; Figure 2).
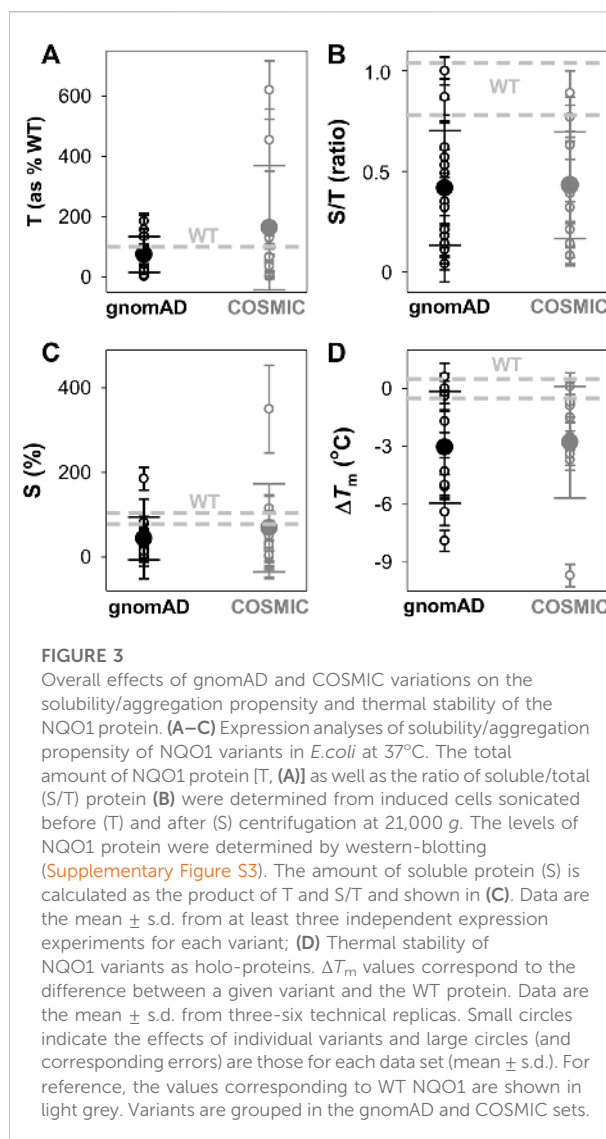
Overall, this set included thirteen variants found in the gnomAD database and nine variants found in the COSMIC database. Seventeen of these variants clustered in the N-terminal part of

the protein (residues 1–51), whereas five were located in the segment comprising residues 106–162 (in close proximity to the active site). Nine variants affected residues buried inside the protein structure (with less than 10% of SASA), whereas the rest are at positions that are partially or highly solvent-exposed (Table 1). The chemical nature of the changes introduced by the substitutions is also quite diverse, and the substitutions are located in different elements of secondary structure (Figure 2B). Based on our computational analysis the 22 variants represent well the heterogeneous scale of effects on NQO1 function and stability. Indeed, of the 22 variants selected 14 are classified by the computational models as "WT-like," 4 as "Total-loss" and 4 as "stable but inactive."

Results from a recent hydrogen/deuterium exchange (HDX) study on WT NQO1 (Vankova et al., 2019) enables us to evaluate the local stability of the protein segments in which these residues are found as well as the effect of FAD and dicoumarol binding (two ligands of functional and stability relevance) (Figure 2C). The L7P, L7R and V9I substitutions are located in regions with high stability that do not change upon binding of FAD or dicoumarol (Dic; a competitive inhibitor of NADH). Variants G3S, G3D, A29T, K32N, G34V, E36K, S40L and H162N are located in regions with intermediate HDX stability and whose local stability is hardly sensitive to ligand binding. Nevertheless, these positions could still, in principle, affect protein folding, stability, or solubility and indirectly affect the binding of the substrates. M45L and M45I are found in a region with low stability and not responsive to ligand binding. Y20N is located in a region with intermediate stability and where HDX shows a response to ligand binding; the remaining six variants (T16M, I51V, W106R, W106C, F107C and M155I) are in regions with low stability and also their local stability respond directly to ligand binding. This last group of variants may thus have a greater potential to affect enzymatic activity [preventing either the formation of the "stable" holo-protein, a precatalytic state, or the formation of a catalytically relevant state, the Dic state, with FAD and the inhibitor Dic bound (Anoz-Carbonell et al., 2020)]. Although this suggestion is simple, it must be noted that regions of either high or low local stability may play roles in enzyme functional and allosteric response, and that local effects can propagate far from the perturbed site (due to ligand binding or amino acid substitutions) (Luque and Freire, 2000; Luque et al., 2002; Naganathan, 2019). Therefore, we next performed an experimental characterization of variant effects on protein stability and function and compared the results with our computational analyses.

3.3 Variant effects on the expression levels, solubility and stability of NQO1

We first experimentally analysed the effect of these 22 NQO1 variants on the expression levels and solubility of

FIGURE 3
Overall effects of gnomAD and COSMIC variations on the solubility/aggregation propensity and thermal stability of the NQO1 protein. (A–C) Expression analyses of solubility/aggregation propensity of NQO1 variants in *E.coli* at 37°C. The total amount of NQO1 protein [T, (A)] as well as the ratio of soluble/total (S/T) protein (B) were determined from induced cells sonicated before (T) and after (S) centrifugation at 21,000 *g*. The levels of NQO1 protein were determined by western-blotting (Supplementary Figure S3). The amount of soluble protein (S) is calculated as the product of T and S/T and shown in (C). Data are the mean ± s.d. from at least three independent expression experiments for each variant; (D) Thermal stability of NQO1 variants as holo-proteins. ΔT_m values correspond to the difference between a given variant and the WT protein. Data are the mean ± s.d. from three-six technical replicas. Small circles indicate the effects of individual variants and large circles (and corresponding errors) are those for each data set (mean ± s.d.). For reference, the values corresponding to WT NQO1 are shown in light grey. Variants are grouped in the gnomAD and COSMIC sets.

the protein (upon expression in *E.coli*) as well as their effects on thermal stability (Supplementary Table S1; Figures 3, 4).

The analysis of the total (T) expression of the variants vs. WT NQO1 at 37°C (Figures 3A,B; Supplementary Figure S3; Supplementary Table S1) revealed that some variants (G3S, G3D, L7P, and V9I) showed higher total expression levels, in agreement with our previous report (Pacheco-García et al., 2020). This is likely the result of codon optimization used in the mutagenesis. Most of the variants showed relatively high expression levels, ranging from 25% to full WT levels, indicating that these variants mildly to moderately reduced total expression levels. The L7R, G34V, S40L, D41G, and D41Y variants showed extremely low expression levels, thus preventing further biophysical characterization. Interestingly, overall, no substantial differences were observed between the average effect of the gnomAD and COSMIC sets of variants.
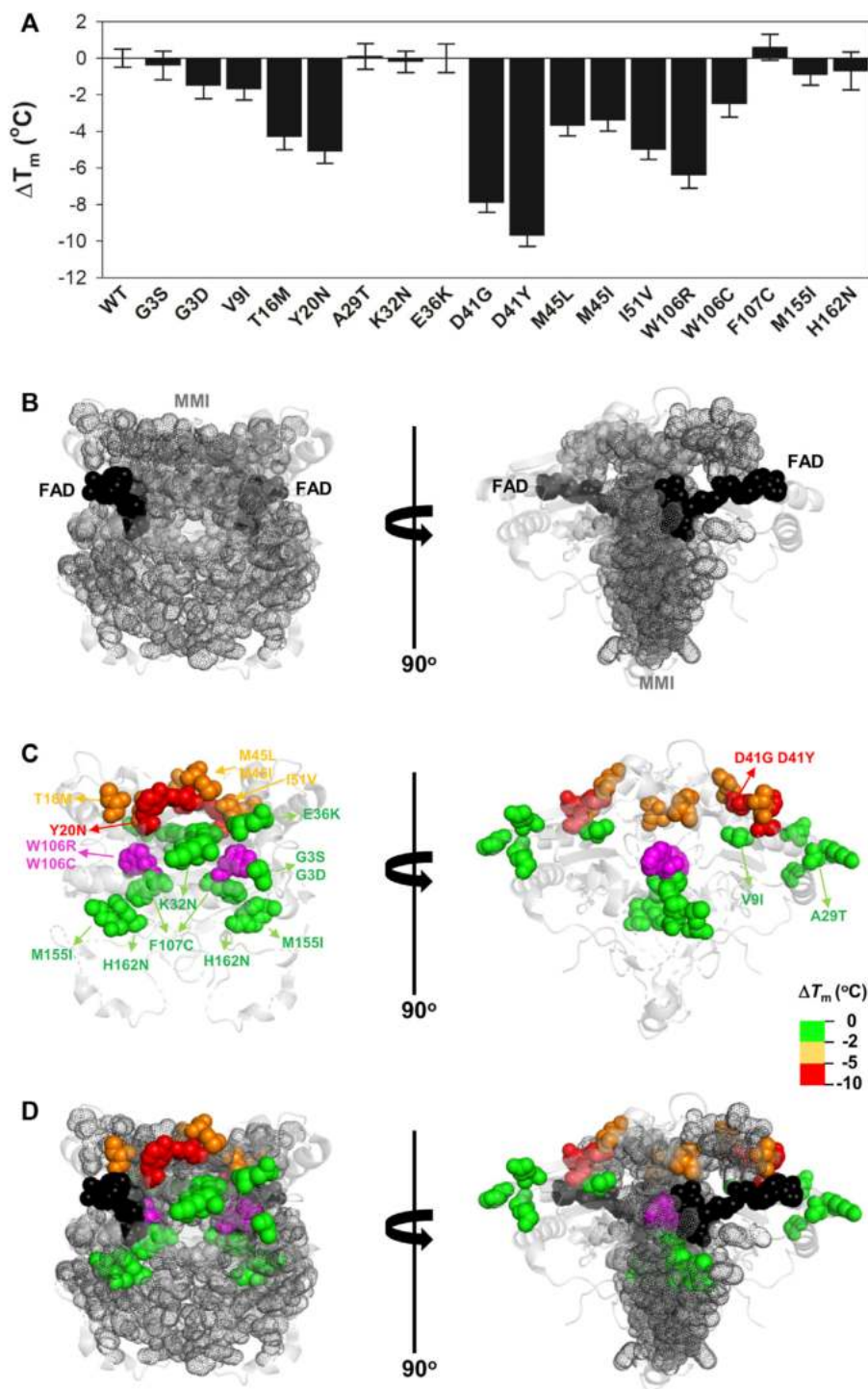
FIGURE 4
Variant effects on thermal stability related to their location near the MMI or the bound FAD. **(A)** Experimental ΔT_m values for individual variants; **(B–D)** Structural location of the FAD (black spheres) and the MMI (grey dots) **(B)** and mutated residues (colour scale according to their destabilizing effect) **(C)**. **(D)** shows an overlay of **(B,C)**. Note that two views rotated 90° are shown. The structural model used for display was PDB code 2F1O (Asher et al., 2006). The residue W106 is highlighted in magenta due to the widely different effects of the W106R/W106C substitutions.

We then determined the fraction of NQO1 existing as soluble protein (S/T ratios; Figure 3B; Supplementary Table S1). WT NQO1 showed a ratio of ~0.9 (Supplementary Table S1; Figure 3B). Again, although some variants showed much lower S/T ratios than WT NQO1, most of them showed values between 0.2 and the WT ratio. Only five variants showed lower S/T ratios than 0.15 (L7P, L7R, S40L, F107C and M155I). Expression under milder conditions (25°C) did not allow for purification of enough protein of the L7P, L7R, S40L, and G34V variants for further biophysical characterization.

We used the product of total expression levels and S/T ratios (i.e. the S values) as a proxy to evaluate the overall effect of amino acid substitutions on NQO1 solubility/aggregation propensity when expressed at 37°C (Supplementary Table S1; Figure 3C). Nine substitutions reduced the solubility below 20% of the WT protein (L7R, G34V, S40L, D41G, D41Y, M45I, W106R, F107C, and M155I).

Next, we determined the thermal stability of the remaining eighteen variants as holo-proteins (i.e. those that were expressed well as soluble proteins and were stable during purification) (Figures 3D, 4, Supplementary Table S1). Nine variants showed a thermal stability close that of the WT protein ($\Delta T_m \leq 2$°C; variants G3S, G3D, V9I, A29T, K32N, E36K, F107C, M155I, and H162N), whereas five variants decreased thermal stability by 2–5°C (variants T16M, M45L, M45I, I51V, and W106C) and four decreased the stability by 5–10°C (Y20N, D41G, D41Y, and W106R) (Supplementary Table S1). Most of the variants that destabilized the holo-protein by more than 2°C are found in the MMI or close to the FAD bound (Figure 4). Here, we note that the reported T_m and ΔT_m values are "apparent" values that cannot be regarded as reporting effects on thermodynamic stability since thermal unfolding is irreversible and kinetically-controlled (Pey et al., 2004). The W106R and W106C variants show different effects, highlighting the importance of both the location and the chemical nature of the change. The effects of the two mutations at W106 were intriguing. Both substitutions are non-conservative changes at a residue in the active site with low solvent exposure and a low structural stability with strong ligand-dependent responsiveness based on HDX studies (Vankova et al., 2019). However, their effects are very different, with W106R causing a much larger decrease in stability than W106C.

To end this section, the observed effects pinpoint that some variants in both the COSMIC and gnomAD databases decrease solubility and thermal stability of NQO1, and overall the two groups do so to similar extents (Figure 3); this observation was also seen in the computational predictions (Figure 1).

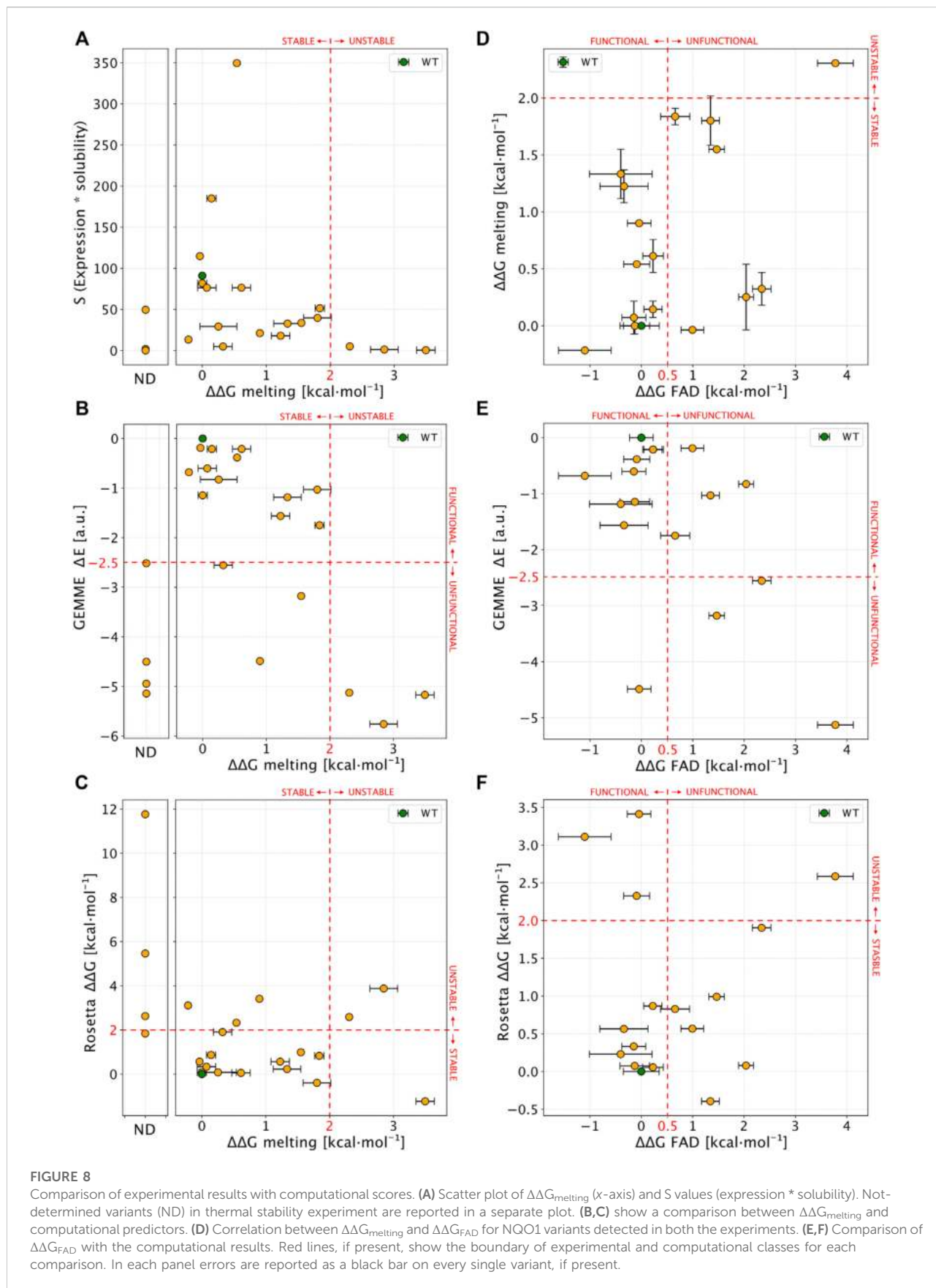## 3.4 FAD binding and the stability of the FAD binding site

Sixteen out of the eighteen variants that yielded good levels of soluble proteins were prepared as apo-proteins to determine their

FIGURE 5
Overall effects of gnomAD and COSMIC mutations on the FAD binding affinity and the local stability of the FAD binding site (FBS). **(A)** Changes in apparent FAD binding free energies ($\Delta\Delta G_{FAD}$) calculated from the difference of binding affinity between WT and a given variant from titrations. Errors are those from linear propagation; **(B)** Local stability of the TCS (next to the FBS) from proteolysis kinetics. Changes in local stability ($\Delta\Delta G_{PROT}$) calculated from the difference of the second-order rate constant between WT and a given variant. Errors are those from linear propagation; For reference, the values corresponding to WT NQO1 are shown in light grey. Variants are grouped in the gnomAD and COSMIC sets.

affinity for FAD by titrations monitored by tryptophan fluorescence (Figures 5A, 6; Supplementary Figure S4; Supplementary Table S2). The D41Y and D41G variants were too unstable to obtain apo-proteins in sufficient amounts and quality.

The G3S, G3D, V9I, K32N, E36K, M45L, M45I, W106C, and F107C variants showed less than a 0.5 kcal mol^{-1} increase in FAD binding free energy (corresponding to less than a 2.5-fold increase in K_d). The Y20N and A29T variants showed a moderate decrease in binding affinity (between 3 and 5-fold higher K_d; thus, a change in FAD binding free energy of 0.5–1.0 kcal mol^{-1}). Five variants (T16M, I51V, W106R, M155I and H162N) markedly decreased the binding affinity for FAD (10–500-fold increase in K_d; between 1 and 3.7 kcal mol^{-1} decrease in binding free energy). Inspection of a structural model of NQO1 allows us to rationalize the effect of these substitutions due to their proximity to the bound FAD, with some exceptions. For instance, the W106R, W106C and F107C substitutions are in proximity to the FAD molecule, but have widely different effects (from ca. 500-fold lower affinity in W106R, to WT-like affinity for W106C and even *higher* affinity than WT for F107C). These results show that NQO1 responds to natural variations very differently even at the same site (i.e. two highly non-conservative variants at the site W106).

Mutational effects on FAD binding affinity ($\Delta\Delta G_{FAD}$) and proteolysis rates ($\Delta\Delta G_{PROT}$) with thermolysin have previously been shown to correlate well (Medina-Carmona et al., 2017a; Medina-Carmona et al., 2019b; Pacheco-García et al., 2020). The site for initial cleavage by thermolysin (TCS) is generally between Ser72–Val73, close to the FAD binding site (Medina-Carmona et al., 2016). All the variants investigated showed proteolysis
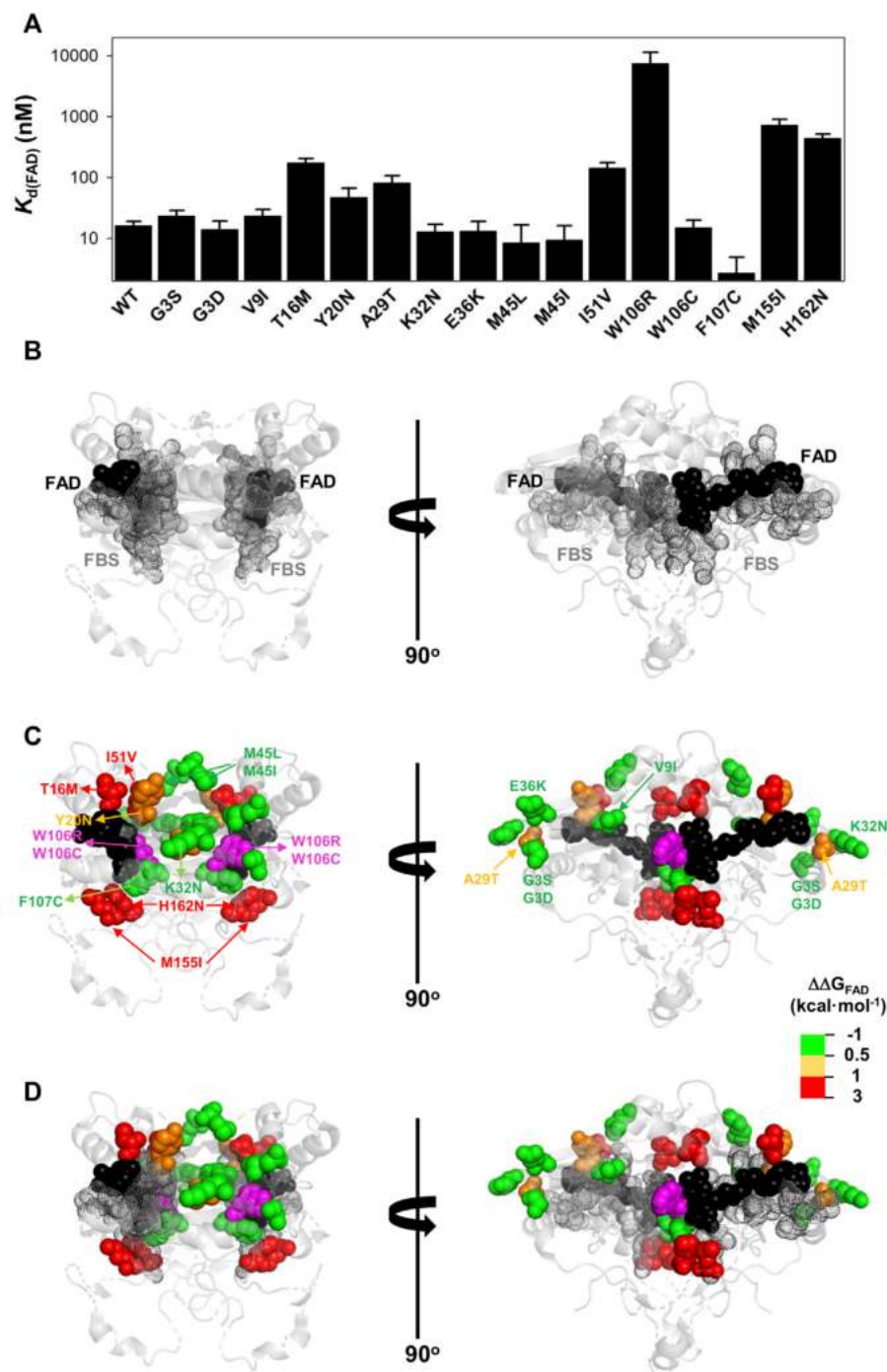
FIGURE 6
Variant effects on the FAD binding affinity. **(A)** FAD binding affinity of each variant was determined by titrations of apo-proteins. At least two independent experiments were carried out for each variant and fitted using a single-type-of-independent binding sites to obtain K_d values (note the logarithmic scale of the y-axis). Errors are those fittings. These K_d values were used to calculate the binding free energy difference ($\Delta\Delta G_{FAD}$) between a given variant and the WT protein (note that a positive value indicates lower affinity). **(B–D)** Structural location of the FAD (black spheres) and the FBS (grey dots) **(B)** and mutated residues (colour scale according to their destabilizing effect on FAD binding) **(C)**. The residue W106 is highlighted in magenta due to the widely different effects of the W106R/W106C substitutions. **(D)** shows an overlay of **(B,C)**. Note that two views rotated 90° are shown. The structural model used for display was PDB code 2F1O (Asher et al., 2006).
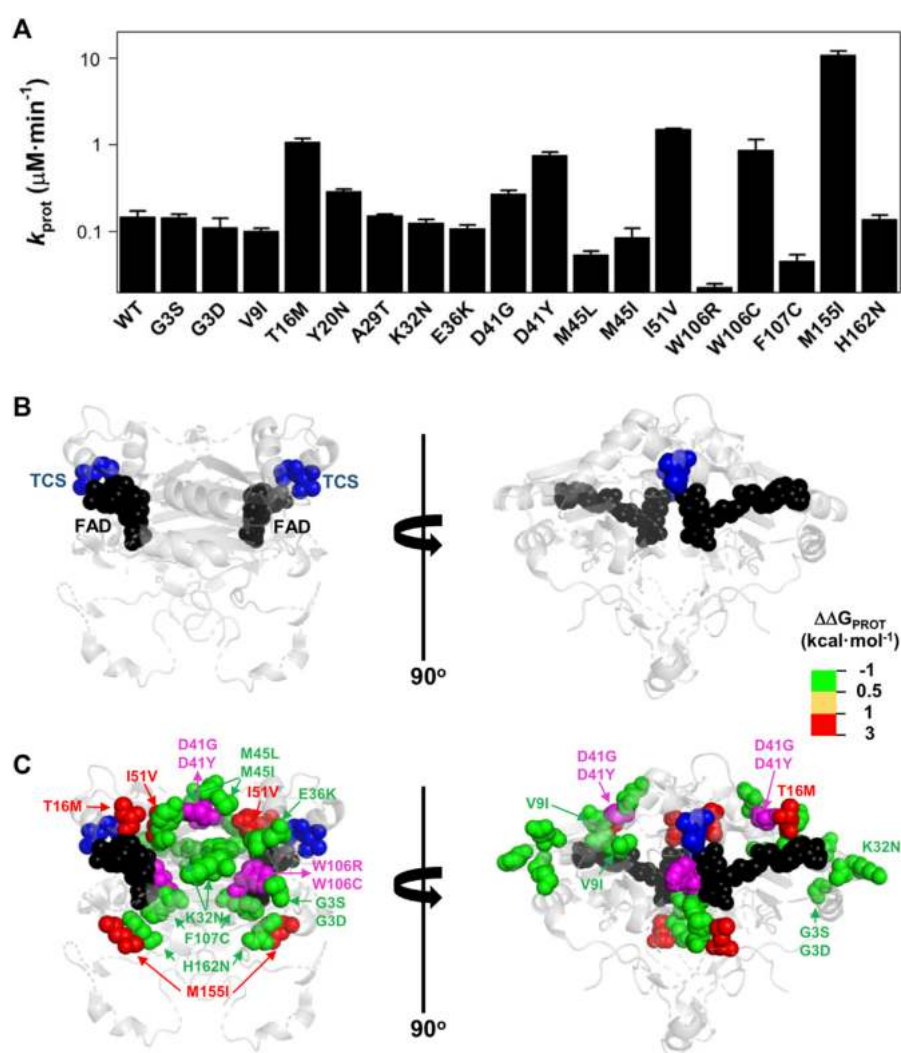
FIGURE 7
Variant effects on the local stability of the FBS from proteolysis. **(A)** Second-order rate constants for proteolysis of NQO1 variants (Supplementary Figure S5C). Errors are those fittings. These k_{PROT} values were used to calculate the TCS local energy free difference ($\Delta\Delta G_{PROT}$) between a given variant and the WT protein (note that a positive value indicates lower affinity). **(B,C)** Structural location of the FAD (black spheres) and the TCS (blue spheres) **(B)**. In **(C)**, mutated residues (colour scale according to their destabilizing effect on FAD binding) **(C)** are overlayed with TCS and FAD. The residues D41 and W106 are highlighted in magenta due to the widely different effects of the D41G/D41Y and W106R/W106C substitutions. Note that two views rotated 90° are shown. The structural model used for display was PDB code 2F1O (Asher et al., 2006).

patterns that were similar to that of WT NQO1 (Supplementary Figure S5). The previously observed correlation between $\Delta\Delta G_{FAD}$ and $\Delta\Delta G_{PROT}$ holds for the 16 variants for which both FAD binding affinity and protease sensitivity can be compared (Supplementary Figure S6). The T16M, I51V, W106C, and M155I mutations destabilized locally the TCS by 1–2.5 kcal mol^{-1} and the residues affected by these substitutions are in general close to the TCS (with the exception of M155I, the most destabilizing mutation) (Figure 7). Again, the results for the W106R/C variants were very different: both affected the local stability by ∼ 1 kcal mol^{-1}, but with opposite signs (Supplementary Table S2).

Overall, the negative impact on FAD binding affinity and the stability of the FAD binding site in the holo-state was similar between variants from COSMIC and gnomAD sets (Figure 5).

3.5 Comparison of experimental analyses and computational predictions

We then proceeded to compare the experimental data to each other and to computational predictions. To ease comparison between the calculated $\Delta\Delta G$ values and thermal melting measurements, we first estimated $\Delta\Delta G_{melting}$ from the ΔT_m

FIGURE 8
Comparison of experimental results with computational scores. **(A)** Scatter plot of $\Delta\Delta G_{melting}$ (*x*-axis) and S values (expression * solubility). Not-determined variants (ND) in thermal stability experiment are reported in a separate plot. **(B,C)** show a comparison between $\Delta\Delta G_{melting}$ and computational predictors. **(D)** Correlation between $\Delta\Delta G_{melting}$ and $\Delta\Delta G_{FAD}$ for NQO1 variants detected in both the experiments. **(E,F)** Comparison of $\Delta\Delta G_{FAD}$ with the computational results. Red lines, if present, show the boundary of experimental and computational classes for each comparison. In each panel errors are reported as a black bar on every single variant, if present.

values using an empirical relationship (Watson et al., 2018), again noting that these are not strictly experimental thermodynamic values as unfolding is not reversible either by temperature (Pey et al., 2004) or chemical denaturants (Figure S7). We first compared the experimental values of $\Delta\Delta G_{melting}$ to the levels of soluble protein (S values, Figure 8A). We found that unstable variants (those with $\Delta\Delta G_{melting} > 2$ kcal mol^{-1} or not amenable for purification, not-determined or ND) mostly showed level of S close to zero (<5) except for L7P. Stable variants (here defined as $\Delta\Delta G_{melting} < 2$ kcal mol^{-1}) instead showed a wide range of S values (76 ± 85%; mean ± s.d.).

We next compared $\Delta\Delta G_{melting}$ with the computational scores (Figures 8B,C). Overall, we found a good agreement for most of the unstable and Not-Determined (ND) variants, which showed $\Delta\Delta G$ >2.0 kcal mol^{-1} and evolutionary distance, $\Delta E < -3$ indicating predicted loss of stability and function. The only exception was D41Y which displayed a stabilizing behaviour in Rosetta $\Delta\Delta G$ predictions. Experimentally stable variants ($\Delta\Delta G_{melting} < 2$ kcal mol^{-1}) displayed low evolutionary distances ($\Delta E > -2.5$ kcal mol^{-1}) except for W106C and T16M. This observation for T16M supports the notion that detrimental effects on protein function may not be connected to thermodynamic destabilization for this variant (Pacheco-García et al., 2020).

We then compared $\Delta\Delta G_{FAD}$ with the other experimental and computational observables (Figures 8D–F). For most (15 out of 16) of the variants where $\Delta\Delta G_{FAD}$ could be measured, the $\Delta\Delta G_{melting}$ was <2 kcal mol^{-1}, indicating stable variants which was also confirmed by Rosetta $\Delta\Delta G$ calculations (13 out of 16 substitutions). Seven of the fifteen variants showed a $\Delta\Delta G_{FAD} > 0.5$ kcal mol^{-1} indicating loss of function. Of these, three variants were captured by evolutionary conservation analysis ($\Delta E < -2.5$ kcal mol^{-1}).

To summarize, for 14 out of 22 tested variants (G3S L7P L7D V9I T16M Y20N K32N G34V E36K D41G M45L M45I W106R, M155I) the predictions from the computational protocols match the experimental results, in terms of variant effects on protein stability and function.

While the results are overall encouraging, there remains differences between computation and experiments for the effects of some mutations (eight out of twenty-two; G3D, A29T, S40L, D41Y, I51V, W106C, F107C, and H162N). For five of these (G3D, S40L, D41Y, W106C, and F107C) it appears that there is a difference between the stability prediction by Rosetta and experiments (noting again that the latter are not equilibrium measurements). For the partially exposed S40L and D41Y the reason for the discrepancy is perhaps related to specific interactions made by these two residues whose effects are not captured by the Rosetta calculations. Both W106C and F107C involve substituting aromatic residues with a cysteine, suggesting problems with evaluating such substitutions. In other two cases (G3D, A29T, and H162N) the GEMME scores did not capture properly the variant effects, possibly because some specific

interactions in human NQO1 may not be present in other homologs of NQO1 and thus, not captured by the evolutionary analysis. Lastly, for I51V the behaviour is opposite from both computational predictions.

4 Discussion

With advances in sequencing technologies, we are uncovering the large genetic variability in the *human genome*. To exploit the availability of this information at the clinical level, we must be able to establish genotype-phenotype correlations accurately and at a large-scale. Although detailed characterization of mutational effects is obviously useful, it is difficult to perform this at such scale (many genes, many variants). However, we may use experimental characterization on a more modest scale to test the performance of current predictive tools in the hope that we can improve them. In this work, we have carried out such an exercise for the human NQO1 protein. The rationale for selecting this system is three-fold: 1) human NQO1 is a multifunctional protein and mutational effects may affect these functions through complex mechanisms (Pacheco-Garcia et al., 2022a; Pacheco-Garcia et al., 2022b). Therefore, contrasting experimental characterization and computational predictions can be challenging for current predictive tools and may help to improve them; 2) Altered NQO1 functionality is associated with increased risk of developing cancer and neurological disorders (Salido et al., 2022). Indeed, the presence of a highly deleterious polymorphic variant in NQO1 is associated with increased cancer risk and affects multiple protein functions through allosteric effects (Lajin and Alachkar, 2013; Pacheco-Garcia et al., 2022b); 3) Over a hundred of missense variants in human NQO1 have been found in human population (i.e. the gnomAD database) or in cancer cell lines as somatic mutations (i.e. the COSMIC database). However, the impact of these mutations on NQO1 multifunctionality and their potential role in cancer development are largely unexplored.

Theoretical advances and new methodologies in the fields of sequence evolution and structure predictions allow us to perform large-scale *in silico* mutagenesis studies on target proteins. Although the current state-of-art algorithms are often [but not always (Frazer et al., 2021)] less accurate at predicting pathogenicity compared to detailed experimental testing, they provide a fast and effective way to predict LoF and sometimes to generate mechanistic hypotheses regarding which properties a variant might affect (Stein et al., 2019; Cagiada et al., 2021).

Here, we first performed *in silico* saturation mutagenesis of WT NQO1, predicting the changes in thermodynamic stability ($\Delta\Delta G$) and evolutionary conservation (ΔE) for 5,187 variants. We combined the two scores to perform a global analysis on how the NQO1 function may be perturbed. Approximately 44% of variants are predicted to cause LoF, with 45% of these drastically

affecting the protein stability. This analysis enabled us to obtain an overview on the possible biologically relevant positions and variants. Indeed, although we know that the ability of computational tools to assign biological functions and predict overall pathogenicity is rapidly increasing, we are still at a point where computational methods may not predict LoF perfectly, and often do not shed much light on the mechanisms of action. This might in particular hold for proteins like NQO1 where multiple biological functions are present, and where some of which may differ between orthologues.

We then used the information provided from the *in silico* saturation mutagenesis to select 22 naturally-occurring variants with a diverse range of predicted effects on protein stability and function to be experimentally tested. We selected nine mutations found in COSMIC and thirteen from gnomAD. Of these variants, 36% severely affected protein foldability and solubility (upon expression in *E.coli*) or reduced conformational stability (at least a $5°C$ decrease in T_m). A quarter of the variants had severely affected FAD binding (a 5-fold decrease in affinity, i.e. a 1 kcal mol^{-1} of binding free energy penalization). For 64% of the variants, experimental characterization and computational predictions agreed in the variant effects on protein stability and function, whereas the remaining 36% of the mutations might be explained by limitations known for the tools used in the prediction process. Although, at this point, this level of agreement is reasonable, it also pinpoints the necessity of improving these predictive tools.

COSMIC mutations are in general somatic (actually, 86% of the COSMIC mutations of NQO1 are labelled as *confirmed* in this database; accessed by 17 August 2022) and likely come from samples that underwent many mutational events in different genes. Thus, the identification of a mutation in the COSMIC database does not imply that this mutation is a driver mutation [here we may define a driver mutation as a mutation with the ability to drive tumorigenesis and confer selective advantages in a tumor cell and a somatic tissue (Martínez-Jiménez et al., 2020)]. Mutations in the gnomAD database belong to heterogeneous groups (many different sequencing projects, some of them case-control studies), and likely reflect genetic variability in the *germline* and in general presence or absence in gnomAD is not sufficient to assign a label as pathogenic or benign. When we examine the NQO1 variants investigated in this work found in the gnomAD database (v.2.1), allele frequencies are overall comparable in *control* vs. *all* samples (Supplementary Table S3). This suggests that there is no strong bias towards *case* samples, and thus the allele frequencies in gnomAD may represent well their presence in a *healthy* population. The presence of these mutations in the germline may predispose somatic cells towards a new mutational event in the WT allele [as occurs in familial cancer cases (Martínez-Jiménez et al., 2020)], thus largely decreasing the NQO1 activity and function.

Our combined experimental and computational analyses provide information on the potential LoF character and the mechanisms by which the variants may exert their effects (protein stability and/or function). Due to its role in the antioxidant defense and stabilization of oncosuppressor proteins, it is likely that NQO1 play a role in cancer development. Homozygous NQO1 knock-out mice revealed cancer-associated phenotypic traits when exposed to chemical or radiological insults (Radjendirane et al., 1998; Long et al., 2000; Iskander et al., 2005; Iskander et al., 2008). Thus, the presence of LoF variants in NQO1 and increased cancer risk may resemble a recessive inheritance (Lajin and Alachkar, 2013). The p.P187S polymorphism (with an allele frequency of ~0.25, Supplementary Table S3) dramatically decreases the intracellular stability of NQO1 thus preventing its interaction with oncosuppressors, reducing enzyme activity and affecting almost the entire structure of NQO1 (Pacheco-Garcia et al., 2021). Noteworthy, it only associates with cancer in homozygotes (Lajin and Alachkar, 2013). Due to the low frequency of most gnomAD NQO1 variants, their presence would be rare even in compound heterozygotes. In fact, 98% of the homozygous samples containing NQO1 missense variations correspond to homozygotes for P187S. However, an additional (*somatic*) mutational event in a WT/P187S genetic background (about 25% of human population) may substantially enhance the LOF phenotype in this common genetic background.

To conclude, we present a test of predictive tools against the experimental characterization of large set of naturally-occurring mutations on NQO1 stability and function. Further steps will be taken to provide a wider perspective on the multifunctionality of NQO1 (i.e. intracellular degradation and stability, high-resolution structural stability in different ligation states, enzyme function and cooperativity, interaction with protein partners, allosteric communication of mutational effects) and the relationships between the genetic diversity of NQO1 in human population and its link with individual propensity towards disease development.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

AP conceived the project and supervised the experimental work; JP-G and KT-M carried out expression, purification and characterization of proteins; MC carried out computational analysis; ES contributed to selection of variants and contributed with reagents; KL-L supervised the computational work; KL-L and AP received funding; M.C, KL-L and AP wrote the original draft; All authors contributed to and approved the final version of the manuscript.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.1063620/full#supplementary-material

References

Abildgaard, A. B., Stein, A., Nielsen, S. V., Schultz-Knudsen, K., Papaleo, E., Shrikhande, A., et al. (2019). Computational and cellular studies reveal structural destabilization and degradation of MLH1 variants in Lynch syndrome. *Elife* 8, e49138. doi:10.7554/eLife.49138

Anoz-Carbonell, E., Timson, D. J., Pey, A. L., and Medina, M. (2020). The catalytic cycle of the antioxidant and cancer-associated human NQO1 enzyme: Hydride transfer, conformational dynamics and functional cooperativity. *Antioxidants* 9, 772. doi:10.3390/antiox9090772

Arnedo-Pac, C., Lopez-Bigas, N., and Muiños, F. (2022). Predicting disease variants using biodiversity and machine learning. *Nat. Biotechnol.* 40, 27–28. doi:10.1038/s41587-021-01187-w

Asher, G., Dym, O., Tsvetkov, P., Adler, J., and Shaul, Y. (2006). The crystal structure of NAD(P)H quinone oxidoreductase 1 in complex with its potent inhibitor dicoumarol. *Biochemistry* 45, 6372–6378. doi:10.1021/bi0600087

Asher, G., Tsvetkov, P., Kahana, C., and Shaul, Y. (2005). A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73. *Genes Dev.* 19, 316–321. doi:10.1101/gad.319905

Beaver, S. K., Mesa-Torres, N., Pey, A. L., and Timson, D. J. (2019). NQO1: A target for the treatment of cancer and neurological diseases, and a model to understand loss of function disease mechanisms. *Biochim. Biophys. Acta. Proteins Proteom.* 1867, 663–676. doi:10.1016/j.bbapap.2019.05.002

Ben-Nissan, G., and Sharon, M. (2014). Regulating the 20S proteasome ubiquitin-independent degradation pathway. *Biomolecules* 4, 862–884. doi:10.3390/biom4030862

Cagiada, M., Johansson, K. E., Valanciute, A., Nielsen, S. V., Hartmann-Petersen, R., Yang, J. J., et al. (2021). Understanding the origins of loss of protein function by analyzing the effects of thousands of variants on activity and abundance. *Mol. Biol. Evol.* 38, 3235–3246. doi:10.1093/molbev/msab095

di Francesco, A., di Germanio, C., Panda, A. C., Huynh, P., Peaden, R., Navas-Enamorado, I., et al. (2016). Novel RNA-binding activity of NQO1 promotes SERPINA1 mRNA translation. *Free Radic. Biol. Med.* 99, 225–233. doi:10.1016/j.freeradbiomed.2016.08.005

Faig, M., Bianchet, M. A., Talalay, P., Chen, S., Winski, S., Ross, D., et al. (2000). Structures of recombinant human and mouse NAD(P)H:quinone oxidoreductases: Species comparison and structural changes with substrate binding and release. *Proc. Natl. Acad. Sci. U. S. A.* 97, 3177–3182. doi:10.1073/pnas.050585797

Frazer, J., Notin, P., Dias, M., Gomez, A., Min, J. K., Brock, K., et al. (2021). Disease variant prediction with deep generative models of evolutionary data. *Nature* 599, 91–95. doi:10.1038/s41586-021-04043-8

Frenz, B., Lewis, S. M., King, I., DiMaio, F., Park, H., and Song, Y. (2020). Prediction of protein mutational free energy: Benchmark and sampling improvements increase classification accuracy. *Front. Bioeng. Biotechnol.* 8, 558247. doi:10.3389/fbioe.2020.558247

Høie, M. H., Cagiada, M., Beck Frederiksen, A. H., Stein, A., and Lindorff-Larsen, K. (2022). Predicting and interpreting large-scale mutagenesis data using analyses of protein stability and conservation. *Cell Rep.* 38, 110207. doi:10.1016/j.celrep.2021.110207

Iskander, K., Barrios, R. J., and Jaiswal, A. K. (2008). Disruption of NAD(P)H: quinone oxidoreductase 1 gene in mice leads to radiation-induced myeloproliferative disease. *Cancer Res.* 68, 7915–7922. doi:10.1158/0008-5472.CAN-08-0766

Iskander, K., Gaikwad, A., Paquet, M., Long, D. J., Brayton, C., Barrios, R., et al. (2005). Lower induction of p53 and decreased apoptosis in NQO1-null mice lead to increased sensitivity to chemical-induced skin carcinogenesis. *Cancer Res.* 65, 2054–2058. doi:10.1158/0008-5472.CAN-04-3157

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637. doi:10.1002/bip.360221211

Katsonis, P., Wilhelm, K., Williams, A., and Lichtarge, O. (2022). Genome interpretation using *in silico* predictors of variant impact. *Hum. Genet.* 141, 1549–1577. doi:10.1007/s00439-022-02457-6

Laine, E., Karami, Y., and Carbone, A. (2019). Gemme: A simple and fast global epistatic model predicting mutational effects. *Mol. Biol. Evol.* 36, 2604–2619. doi:10.1093/molbev/msz179

Lajin, B., and Alachkar, A. (2013). The NQO1 polymorphism C609T (Pro187Ser) and cancer susceptibility: A comprehensive meta-analysis. *Br. J. Cancer* 109, 1325–1337. doi:10.1038/bjc.2013.357

Li, R., Bianchet, M. A., Talalay, P., and Amzel, L. M. (1995). The three-dimensional structure of NAD(P)H:quinone reductase, a flavoprotein involved in cancer chemoprotection and chemotherapy: Mechanism of the two-electron reduction. *Proc. Natl. Acad. Sci. U. S. A.* 92, 8846–8850. doi:10.1073/pnas.92.19.8846

Lienhart, W. D., Gudipati, V., Uhl, M. K., Binter, A., Pulido, S. A., Saf, R., et al. (2014). Collapse of the native structure caused by a single amino acid exchange in human NAD(P)H:Quinone oxidoreductase. *FEBS J.* 281, 4691–4704. doi:10.1111/febs.12975

Long, D. J., Waikel, R. L., Wang, X. J., Perlaky, L., Roop, D. R., and Jaiswal, A. K. (2000). NAD(P)H:quinone oxidoreductase 1 deficiency increases susceptibility to benzo(a)pyrene-induced mouse skin carcinogenesis. *Cancer Res.* 60, 5913–5915.

Luo, S., Kang, S. S., Wang, Z. H., Liu, X., Day, J. X., Wu, Z., et al. (2019). Akt phosphorylates NQO1 and triggers its degradation, abolishing its antioxidative activities in Parkinson's disease. *J. Neurosci.* 39, 7291–7305. doi:10.1523/JNEUROSCI.0625-19.2019

Luque, I., and Freire, E. (2000). Structural stability of binding sites: Consequences for binding affinity and allosteric effects. *Proteins* 4, 63–71. doi:10.1002/1097-0134(2000)41:4+<63::aid-prot60>3.3.co;2-y

Luque, I., Leavitt, S. A., and Freire, E. (2002). The linkage between protein folding and functional cooperativity: Two sides of the same coin? *Annu. Rev. Biophys. Biomol. Struct.* 31, 235–256. doi:10.1146/annurev.biophys.31.082901.134215

Martínez-Jiménez, F., Muiños, F., Sentís, I., Deu-Pons, J., Reyes-Salazar, I., Arnedo-Pac, C., et al. (2020). A compendium of mutational cancer driver genes. *Nat. Rev. Cancer* 20, 555–572. doi:10.1038/s41568-020-0290-x

Martínez-Limón, A., Alriquet, M., Lang, W. H., Calloni, G., Wittig, I., and Vabulas, R. M. (2016). Recognition of enzymes lacking bound cofactor by Protein quality control. *Proc. Natl. Acad. Sci. U. S. A.* 113, 12156–12161. doi:10.1073/pnas.1611994113

Martínez-Limón, A., Calloni, G., Ernst, R., and Vabulas, R. M. (2020). Flavin dependency undermines proteome stability, lipid metabolism and cellular proliferation during vitamin B2 deficiency. *Cell Death Dis.* 11, 725. doi:10.1038/s41419-020-02929-5

McInnes, G., Sharo, A. G., Koleske, M. L., Brown, J. E. H., Norstad, M., Adhikari, A. N., et al. (2021). Opportunities and challenges for the computational interpretation of rare variation in clinically important genes. *Am. J. Hum. Genet.* 108, 535–548. doi:10.1016/j.ajhg.2021.03.003

Medina-Carmona, E., Betancor-Fernández, I., Santos, J., Mesa-Torres, N., Grottelli, S., Batlle, C., et al. (2019). Insight into the specificity and severity of pathogenic mechanisms associated with missense mutations through experimental and structural perturbation analyses. *Hum. Mol. Genet.* 28, 1–15. doi:10.1093/hmg/ddy323

Medina-Carmona, E., Fuchs, J. E., Gavira, J. A., Mesa-Torres, N., Neira, J. L., Salido, E., et al. (2017). Enhanced vulnerability of human proteins towards disease-associated inactivation through divergent evolution. *Hum. Mol. Genet.* 26, 3531–3544. doi:10.1093/hmg/ddx238

Medina-Carmona, E., Neira, J. L., Salido, E., Fuchs, J. E., Palomino-Morales, R., Timson, D. J., et al. (2017). Site-to-site interdomain communication may mediate different loss-of-function mechanisms in a cancer-associated NQO1 polymorphism. *Sci. Rep.* 7, 44532. doi:10.1038/srep44532

Medina-Carmona, E., Palomino-Morales, R. J., Fuchs, J. E., Padín-Gonzalez, E., Mesa-Torres, N., Salido, E., et al. (2016). Erratum: Conformational dynamics is key to understanding loss-of-function of NQO1 cancer-associated polymorphisms and its correction by pharmacological ligands. *Sci. Rep.* 6 (1), 21939. doi:10.1038/srep21939

Medina-Carmona, E., Rizzuti, B., Martín-Escolano, R., Pacheco-García, J. L., Mesa-Torres, N., Neira, J. L., et al. (2019). Phosphorylation compromises FAD binding and intracellular stability of wild-type and cancer-associated NQO1: Insights into flavo-proteome stability. *Int. J. Biol. Macromol.* 125, 1275–1288. doi:10.1016/j.ijbiomac.2018.09.108

Naganathan, A. N. (2019). Modulation of allosteric coupling by mutations: From protein dynamics and packing to altered native ensembles and function. *Curr. Opin. Struct. Biol.* 54, 1–9. doi:10.1016/j.sbi.2018.09.004

Oh, E. T., Kim, J. W., Kim, J. M., Kim, S. J., Lee, J. S., Hong, S. S., et al. (2016). NQO1 inhibits proteasome-mediated degradation of HIF-1α. *Nat. Commun.* 7, 13593. doi:10.1038/ncomms13593

Pacheco-Garcia, J. L., Anoz-Carbonell, E., Loginov, D. S., Vankova, P., Salido, E., Man, P., et al. (2022). Different phenotypic outcome due to site-specific phosphorylation in the cancer-associated NQO1 enzyme studied by phosphomimetic mutations. *Arch. Biochem. Biophys.* 729, 109392. doi:10.1016/j.abb.2022.109392

Pacheco-Garcia, J. L., Anoz-Carbonell, E., Vankova, P., Kannan, A., Palomino-Morales, R., Mesa-Torres, N., et al. (2021). Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue. *Redox Biol.* 46, 102112. doi:10.1016/j.redox.2021.102112

Pacheco-García, J. L., Cano-Muñoz, M., Sánchez-Ramos, I., Salido, E., and Pey, A. L. (2020). Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein. *J. Pers. Med.* 10, E207–E231. doi:10.3390/jpm10040207

Pacheco-Garcia, J. L., Loginov, D. S., Anoz-Carbonell, E., Vankova, P., Palomino-Morales, R., Salido, E., et al. (2022). Allosteric communication in the multifunctional and redox NQO1 protein studied by cavity-making mutations. *Antioxidants* 11, 1110. doi:10.3390/antiox11061110

Park, C., and Marqusee, S. (2004). Probing the high energy states in proteins by proteolysis. *J. Mol. Biol.* 343, 1467–1476. doi:10.1016/j.jmb.2004.08.085

Park, H., Bradley, P., Greisen, P., Liu, Y., Mulligan, V. K., Kim, D. E., et al. (2016). Simultaneous optimization of biomolecular energy functions on features from small molecules and macromolecules. *J. Chem. Theory Comput.* 12, 6201–6212. doi:10.1021/acs.jctc.6b00819

Pey, A. L. (2018). Biophysical and functional perturbation analyses at cancer-associated P187 and K240 sites of the multifunctional NADP(H):quinone oxidoreductase 1. *Int. J. Biol. Macromol.* 118, 1912–1923. doi:10.1016/j.ijbiomac.2018.07.051

Pey, A. L., Megarity, C. F., and Timson, D. J. (2004). FAD binding overcomes defects in activity and stability displayed by cancer-associated variants of human NQO1. *Biochim. Biophys. Acta* 1842, 2163–2173. doi:10.1016/j.bbadis.2014.08.011

Radjendirane, V., Joseph, P., Lee, Y. H., Kimura, S., Klein-Szanto, A. J., Gonzalez, F. J., et al. (1998). Disruption of the DT diaphorase (NQO1) gene in mice leads to increased menadione toxicity. *J. Biol. Chem.* 273, 7382–7389. doi:10.1074/jbc.273.13.7382

Remmert, M., Biegert, A., Hauser, A., and Söding, J. (2011). HHblits: Lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods* 9, 173–175. doi:10.1038/nmeth.1818

Ross, D., and Siegel, D. (2018). NQO1 in protection against oxidative stress. *Curr. Opin. Toxicol.* 7, 67–72. doi:10.1016/j.cotox.2017.10.005

Salido, E., Timson, D. J., Betancor-Fernández, I., Palomino-Morales, R., Anoz-Carbonell, E., Pacheco-García, J. L., et al. (2022). Targeting HIF-1α function in cancer through the chaperone action of NQO1: Implications of genetic diversity of NQO1. *J. Pers. Med.* 12, 747. doi:10.3390/jpm12050747

Siegel, D., Bersie, S., Harris, P., di Francesco, A., Armstrong, M., Reisdorph, N., et al. (2021). A redox-mediated conformational change in NQO1 controls binding to microtubules and α-tubulin acetylation. *Redox Biol.* 39, 101840. doi:10.1016/j.redox.2020.101840

Stein, A., Fowler, D. M., Hartmann-Petersen, R., and Lindorff-Larsen, K. (2019). Biophysical and mechanistic models for disease-causing protein variants. *Trends biochem. Sci.* 44, 575–588. doi:10.1016/j.tibs.2019.01.003

Steinegger, M., Meier, M., Mirdita, M., Vöhringer, H., Haunsberger, S. J., and Söding, J. (2019). HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinforma.* 20, 473. doi:10.1186/s12859-019-3019-7

Vankova, P., Salido, E., Timson, D. J., Man, P., and Pey, A. L. (2019). A dynamic core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer. *Biomolecules* 9, 728. doi:10.3390/biom9110728

Watson, M. D., Monroe, J., and Raleigh, D. P. (2018). Size-dependent relationships between protein stability and thermal unfolding temperature have important implications for analysis of protein energetics and high-throughput assays of protein-ligand interactions. *J. Phys. Chem. B* 122, 5278–5285. doi:10.1021/acs.jpcb.7b05684

Xu, Q., Tang, Q., Katsonis, P., Lichtarge, O., Jones, D., Bovo, S., et al. (2017). Benchmarking predictions of allostery in liver pyruvate kinase in CAGI4. *Hum. Mutat.* 38, 1123–1131. doi:10.1002/humu.23222

Supplementary Table 1. Variant effects on the solubility and thermal stability of NQO1. Expression analyses of NQO1 were carried out in *E.coli* at 37°C and the total expression levels (T), and the fraction of soluble protein (S/T ratio) determined by western-blot (mean ± s.d. from three independent experiments). The % of soluble protein (S) was thus calculated as the T x S/T product. T_m values were determined upon purification of holo-NQO1 proteins by thermal denaturation (mean±s.d. from at least three replicates). N.Det. not determined due to very low expression levels as soluble protein.

Variant	T (% vs.WT)	S/T ratio	S (%)	T_m (°C)
WT	100	0.91±0.13	91±13	51.8±0.5
G3S	185±26	1.00±0.07	185±27	51.4±0.6*
G3D	454±103	0.77±0.10	350±103	50.3±0.5*
L7P	620±97	0.08±0.04	50±10	N.Det.
L7R	9±10	0.11±0.03	~ 1	N.Det.
V9I	156±48	0.49±0.37	76±60	50.1±0.3*
T16M	96±10	0.35±0.12	34±16	47.5±0.5*
Y20N	97±10	0.53±0.08	51±13	46.7±0.4*
A29T	129±26	0.89±0.11	115±28	51.9±0.5*
K32N	134±7	0.52±0.17	76±18	51.6±0.3
G34V	6±1	0.32±0.11	~ 2	N.Det.
E36K	94±11	0.87±0.09	82±14	51.8±0.6
S40L	2±1	0.04±0.03	> 1	N.Det.
D41G	7±2	0.18±0.23	~ 1	43.9±0.2
D41Y	2±1	0.21±0.14	~ 1	42.1±0.3
M45L	52±41	0.63±0.07	33±42	48.1±0.2
M45I	46±29	0.39±0.26	18±39	48.4±0.3
I51V	64±21	0.62±0.13	40±25	46.8±0.2
W106R	24±16	0.21±0.07	~ 5	45.4±0.5
W106C	66±45	0.32±0.08	21±45	49.3±0.5
F107C	95±64	0.14±0.10	13±65	52.4±0.5
M155I	35±33	0.14±0.11	~ 5	50.9±0.3
H162N	65±71	0.45±0.38	29±80	51.1±0.9

* From Pacheco-García JL, Cano-Muñoz M, Sánchez-Ramos I, Salido E, Pey AL. (2020). Naturally-Occurring Rare Mutations Cause Mild to Catastrophic Effects in the Multifunctional and Cancer-Associated NQO1 Protein. J. Pers. Med. 10, 207. doi: 10.3390/jpm10040207.

Supplementary Table 2. Variant effects on FAD binding. The apparent dissociation constants ($K_{d\ FAD}$) were determined by titration of apo-NQO1 variants with FAD. Data are best-fit parameters from at least two independent titrations. The local stability of the TCS (close to the FAD binding site) was determined by proteolysis with thermolysin (k_{prot}). k_{prot} is the second-order rate constant for proteolysis obtained from the linear dependence of the apparent first-order rate constant on protease concentration.

| Variant	$K_{d\ FAD}$ (nM)	k_{prot} (μM·min^{-1})
WT	16.1±2.7	0.149±0.024
G3S	23.2±5.4	0.146±0.013
G3D	13.9±5.2	0.111±0.032
L7P	N.Det.	N.Det.
L7R	N.Det.	N.Det.
V9I	23.3±6.5	0.101±0.009
T16M	174±31	1.075±0.115
Y20N	46.8±20.0	0.292±0.017
A29T	80.8±25.7	0.153±0.005
K32N	12.7±4.3	0.126±0.013
G34V	N.Det.	N.Det.
E36K	13.1±5.7	0.109±0.011
S40L	N.Det.	N.Det.
D41G	N.Det.	0.271±0.029
D41Y	N.Det.	0.754±0.072
M45L	8.4±8.2	0.054±0.006
M45I	9.3±6.9	0.086±0.025
I51V	143±32	1.521±0.032
W106R	7400±4000*	0.023±0.002
W106C	15±5*	0.864±0.285
F107C	2.7±2.2*	0.046±0.008
M155I	722±173*	10.75±1.26
H162N	441±74*	0.138±0.018

* Unpublished work.

Supplementary Table 3. Allelic frequency of NQO1 mutations found in the gnomAD database and experimentally analysed in this work. Frequencies are reported from the gnomAD v.2.1. using all or control samples. For sake of comparison, data on the common P187S polymorphism is also added.

| Variation	Allelic frequency	
	All (141.456) samples	**Control (60.146) samples**
G3S	$1.21 \cdot 10^{-5}$	$1.86 \cdot 10^{-5}$
L7R	$7.80 \cdot 10^{-5}$	$5.49 \cdot 10^{-5}$
V9I	$3.90 \cdot 10^{-5}$	$4.16 \cdot 10^{-5}$
T16M	$2.83 \cdot 10^{-5}$	$4.64 \cdot 10^{-4}$
Y20N	$2.12 \cdot 10^{-5}$	$1.83 \cdot 10^{-5}$
K32N	$7.96 \cdot 10^{-5}$	$4.57 \cdot 10^{-5}$
G34V	$3.98 \cdot 10^{-6}$	$9.14 \cdot 10^{-6}$
E36K	$6.37 \cdot 10^{-5}$	$3.33 \cdot 10^{-5}$
S40L	$1.99 \cdot 10^{-5}$	$9.14 \cdot 10^{-6}$
D41G	$3.98 \cdot 10^{-6}$	$9.14 \cdot 10^{-6}$
I51V	$6.37 \cdot 10^{-5}$	$3.33 \cdot 10^{-5}$
W106R	$2.79 \cdot 10^{-5}$	$4.57 \cdot 10^{-5}$
F107C	$1.77 \cdot 10^{-5}$	0
P187S	$2.47 \cdot 10^{-1}$	$2.55 \cdot 10^{-1}$
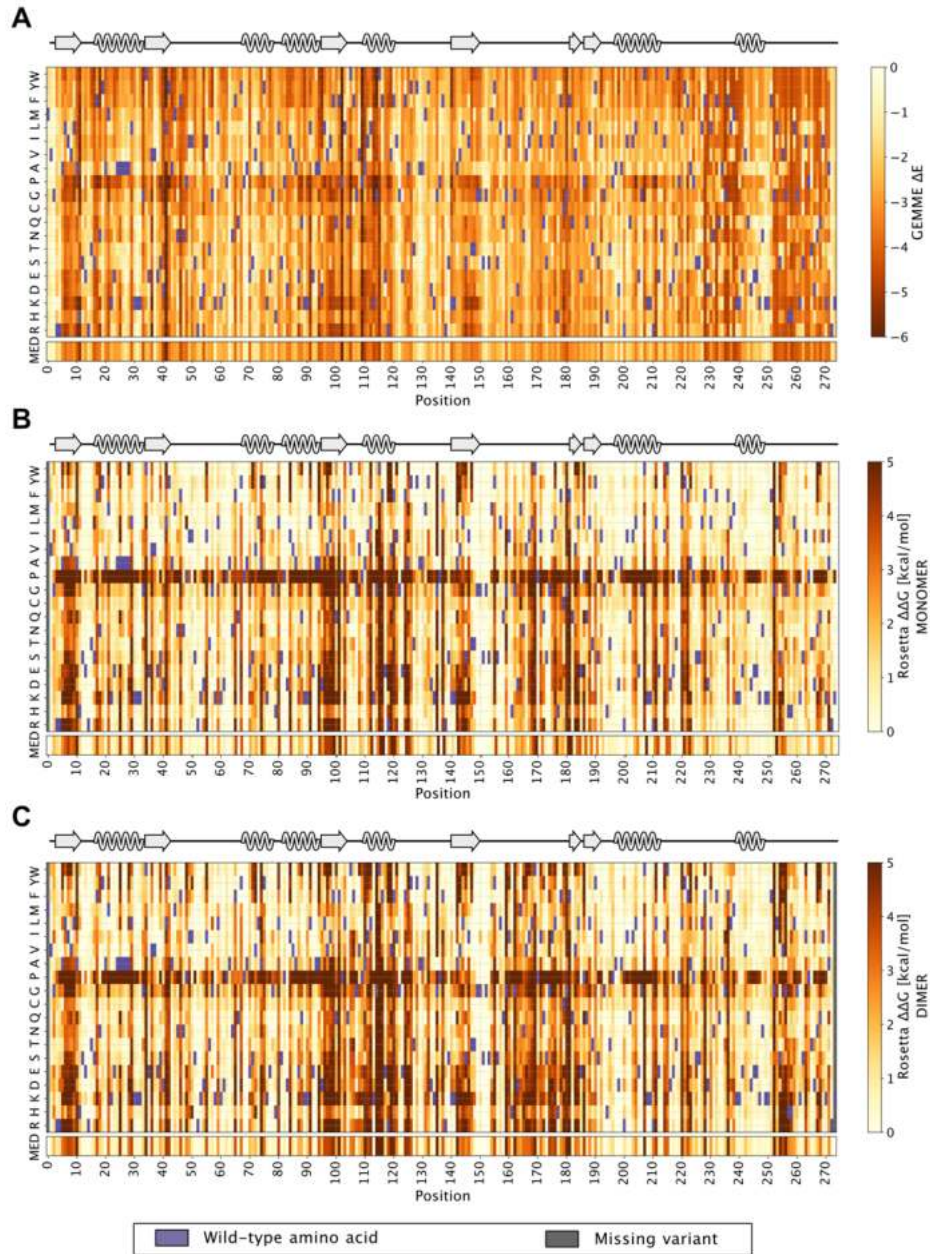
Supplementary Figure 1. Computational prediction maps of thermodynamic stability and evolutionary conservation of NQO1. Panel A shows the results of evolutionary conservation analysis using GEMME on NQO1 sequence. Variants evolutionary close to the WT have score close to zero, while variants with detrimental effects on the protein have high negative scores (red shadows). WT amino acid is indicated with a purple box and the median score is shown for each residue. Panel B and C show the thermodynamic stability ($\Delta\Delta G$) maps of NQO1 evaluated using Rosetta using the monomer (B) and the dimer (C) structure. Variants with a stability similar to the wild-type have scores close to zero, while variants with detrimental effects on the protein's stability have large positive scores (coloured with shadows of red). Wild-type amino acids are indicated with purple boxed, positions (and therefore variants) which are missed in the PDB structure used are coloured in grey and the median score is shown for each position.
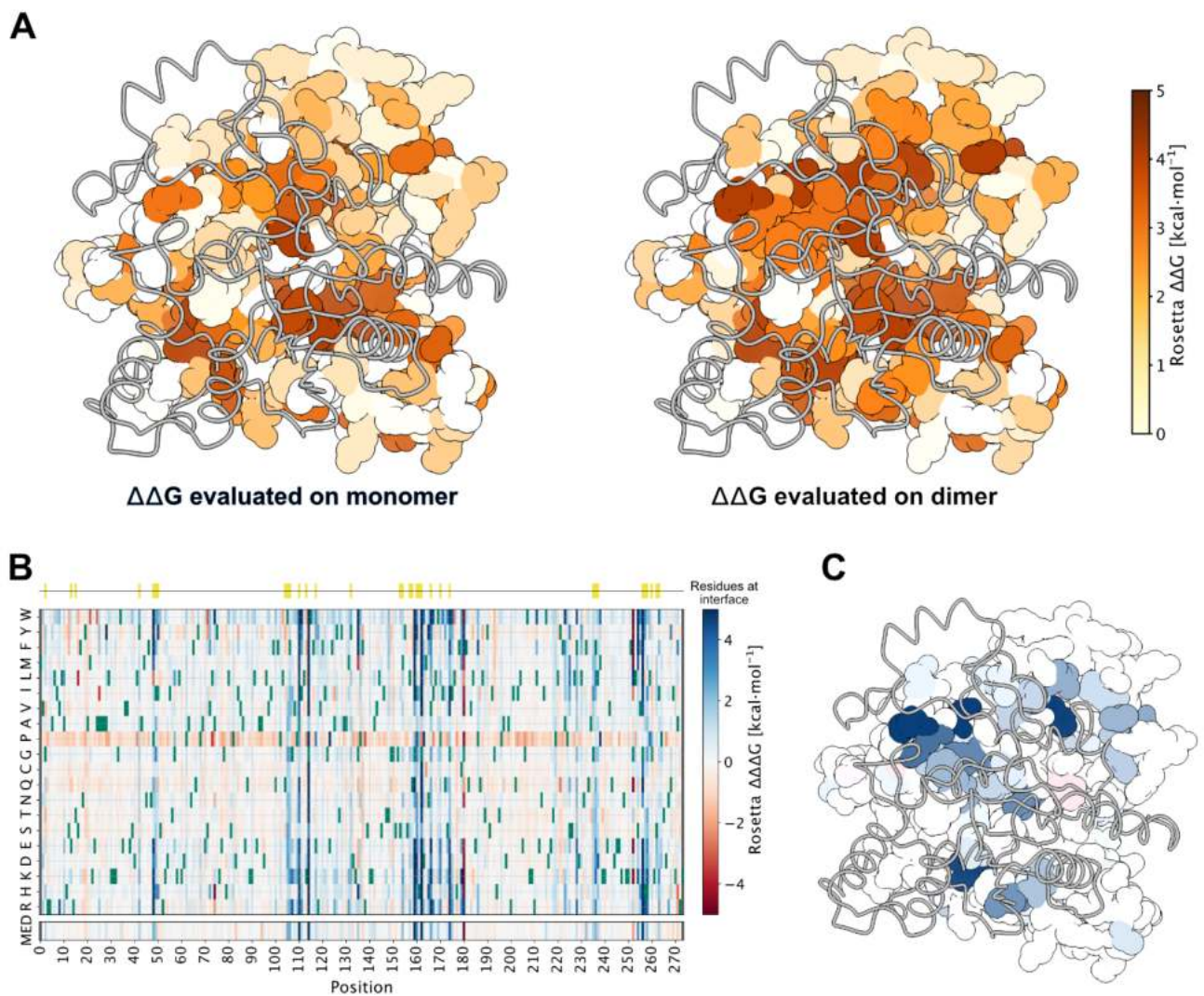
Supplementary Figure 2. Thermodynamic stability differences (ΔΔG) between the monomer and dimer based on computational methods. Panel A shows the median ΔΔG value for each residue on the crystal structure of NQO1, the figure on the left show median ΔΔGs evaluated keeping only the monomer structure of the NQO1 while the figure on the right report median ΔΔG evaluation using homozygous mutation on the dimer. Positions with neutral effect are coloured in white, while detrimental positions are coloured in shadows of red. Panel B shows the heatmap with the Δ(ΔΔG) between the two different ΔΔG evaluations on the monomer and dimer. Variants with a higher stability in dimer evaluation of ΔΔG are coloured in blue while variants with a higher stability in the monomer ΔΔG are coloured in red. WT residues are reported in green. On the top of the heatmap, interface residues are reported with a yellow marker. Panel C shows the median positional Δ(ΔΔG) difference between the two evaluations mapped to the crystal structure of NQO1.
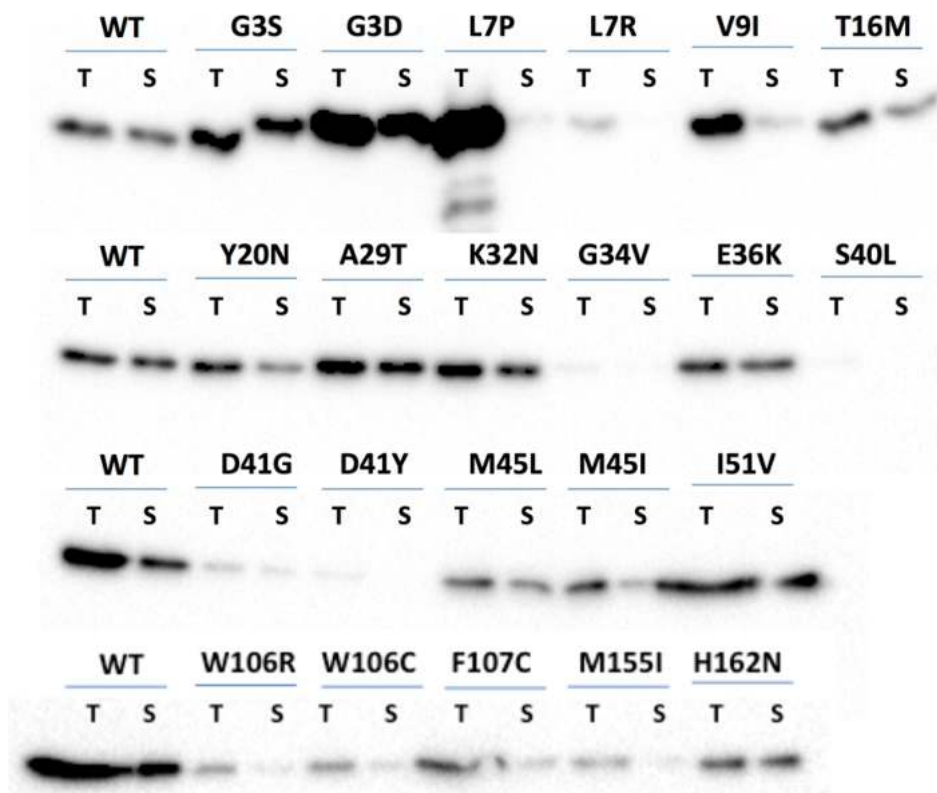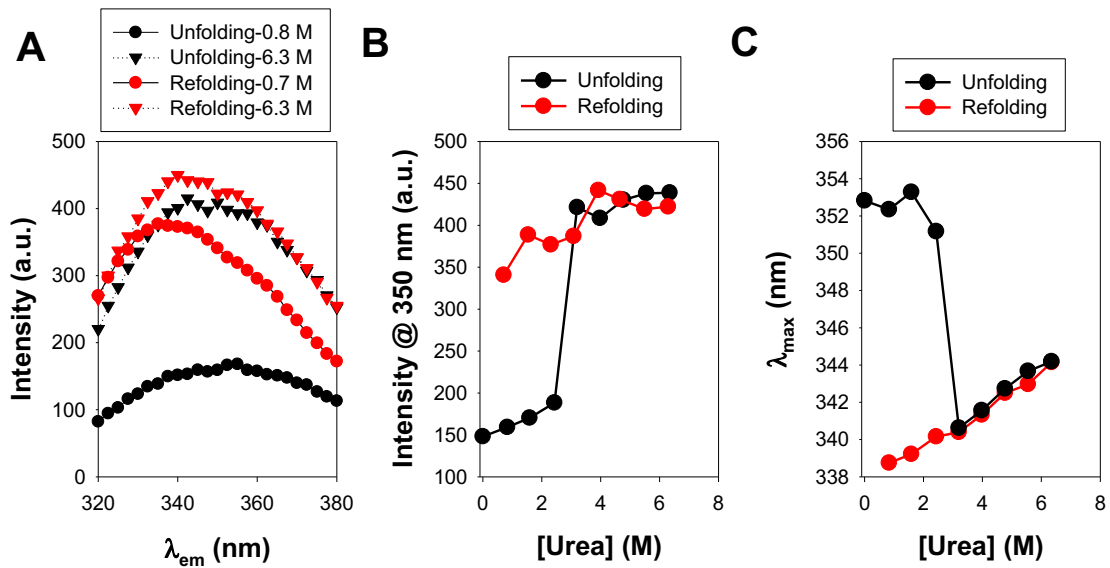
Supplementary Figure 3. Representative Western-blot analyses of total (T) and soluble (S) expression levels for NQO1 variants in *E.coli*. Experimental details can be found in the main text
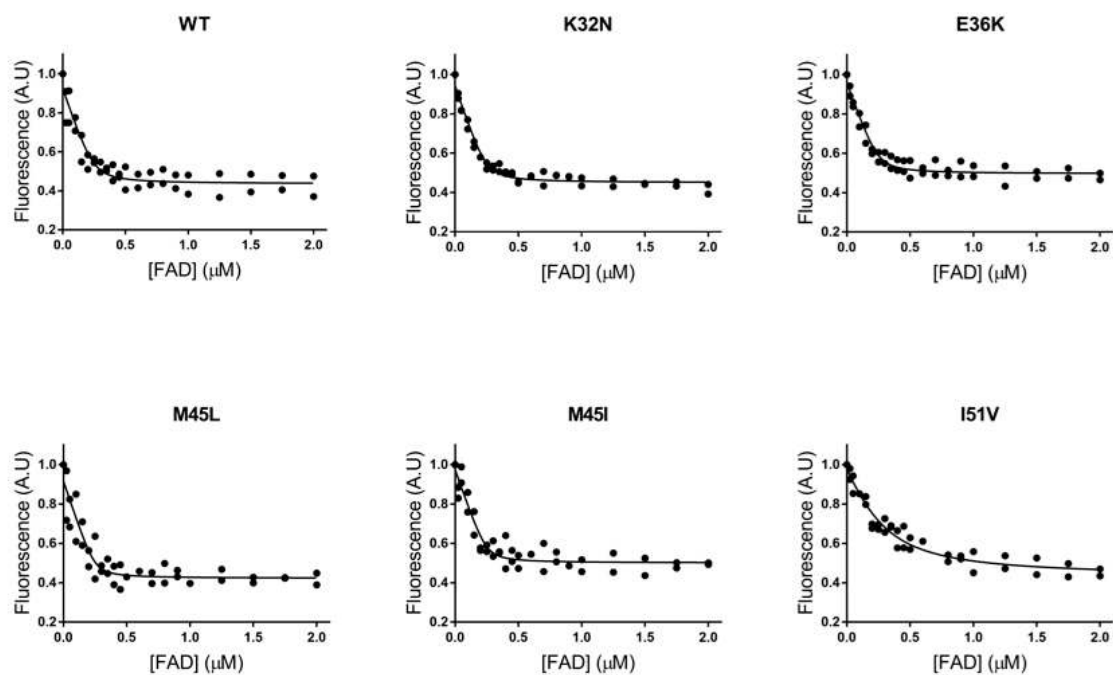
Supplementary Figure 4. Titrations of apo-NQO1 proteins with FAD. Experiments were replicated twice and all data were fitted to single-site binding model. In all the cases, titrations were carried out by fluorescence measurements. Experimental details can be found in the main text.
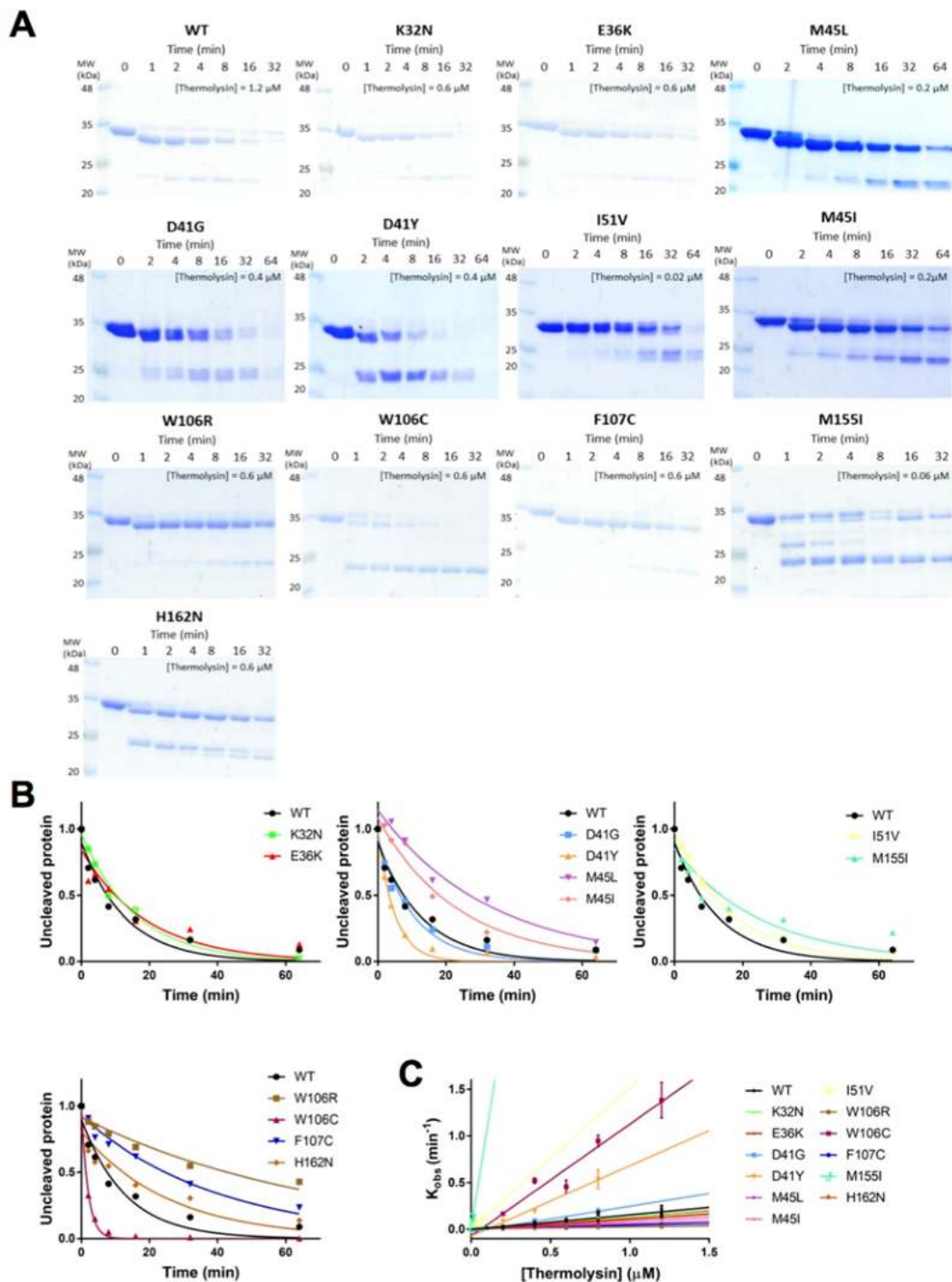
Supplementary Figure 5. Partial proteolysis of NQO1 variants by thermolysin. Panel A shows representative kinetic experiments with the protease and NQO1 variants. The concentration of protease used is indicated in each case. Panel B shows the corresponding densitometric analysis of native protein over time used to obtained the observed first-order rate constant using 0.4 μM thermolysin. Panel C shows the linear dependence of the observed rate constants that can be interpreted as effects on the thermodynamic stability of the TCS. Experimental details can be found in the main text.

Supplementary Figure 6. Correlation between changes in free energy for limited proteolysis ($\Delta\Delta G_{PROT}$) at the NTD and binding affinity for FAD ($\Delta\Delta G_{FAD}$). The plot includes the best linear correlation between both parameters. W106R was excluded as a clear outlier.

Supplementary Figure 7. Urea induced denaturation of holo-NQO1 WT is not reversible. A) Fluorescence spectra of hNQO1 in unfolding (black symbols) and refolding (red symbols) experiments. B) the fluorescence intensities at 350 nm in unfolding (black symbols) and refolding (red symbols) experiments. C) λ_{max} of fluorescence emission spectra in unfolding (black symbols) and refolding (red symbols) experiments. For denaturation experiments, 2 µM holo-NQO1 protein was incubated with 2 µM FAD in HEPES-KOH 50 mM, pH 7.4 and 2 mM β-mercaptoethanol in the presence of 0-7.2 M urea for 16 h at 25 ℃ in the darkness. Denaturation was followed by intrinsic fluorescence (excitation, 280 nm; emission, 320-380 nm; slits 5 nm) at 25 ℃. using a Cary Eclipse spectrofluorimeter (Agilent Technologies, Waldbronn, Germany) and a 3 x 3 mm quartz cuvette. Four scans were measured for each sample and averaged. Blanks in the absence of protein were acquired and subtracted. For renaturation experiments, a sample of 20 µM holo-NQO1 protein was incubated with 20 µM FAD in HEPES-KOH 50 mM, pH 7.4 and 2 mM β-mercaptoethanol in the presence of 7.2 M urea for 16 h at 25 ℃ in the darkness. Then, protein samples were diluted 10-fold to different urea solutions in the range 0.8-6.4 M concentration in HEPES-KOH 50 mM, pH 7.4 and 2 mM β-mercaptoethanol and allowed to "renature" for 6 h at 25 ℃ before fluorescence spectra were acquired at the same conditions than those described in denaturation experiments. In all cases, urea concentration was determined by refractive index measurements.

REFERENCIAS

1.   Chaudhuri, T. K. & Paul, S. Protein-misfolding diseases and chaperone-based therapeutic approaches. *FEBS J.* **273**, 1331–1349 (2006).

2.   Anfinsen, C. B. Principles that govern protein folding. *Science (80-.).* **181**, 223–230 (1973).

3.   Herczenik, E. & Gebbink, M. F. B. G. Molecular and cellular aspects of protein misfolding and disease. *FASEB J.* **22**, 2115–2133 (2008).

4.   Van Der Lee, R. *et al.* Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **114**, 6589–6631 (2014).

5.   Dill, K. A., Ozkan, S. B., Shell, M. S. & Weikl, T. R. The protein folding problem. *Annu. Rev. Biophys.* **37**, 289–316 (2008).

6.   Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

7.   Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science (80-.).* **373**, 871–876 (2021).

8.   Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins Struct. Funct. Bioinforma.* **21**, 167–195 (1995).

9.   Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry* **29**, 7133–7155 (1990).

10.   Gershenson, A. & Gierasch, L. M. Protein folding in the cell: Challenges and progress. *Curr. Opin. Struct. Biol.* **21**, 32–41 (2011).

11.   Ellis, R. J. & Minton, A. P. Protein aggregation in crowded environments. *Biol. Chem.* **387**, 485–497 (2006).

12.   Fedyukina, D. V. & Cavagnero, S. Protein folding at the exit tunnel. *Annu. Rev. Biophys.* **40**, 337–359 (2011).

13.   Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M. & Ulrich Hartl, F. *Molecular chaperone functions in protein folding and proteostasis. Annual*

Review of Biochemistry vol. 82 (2013).

14. Coux, O., Tanaka, K. & Goldberg, A. L. Structure and functions of the 20S and 26S proteasomes. *Annu. Rev. Biochem.* **65**, 801–847 (1996).

15. Chen, B., Retzlaff, M., Roos, T. & Frydman, J. Cellular strategies of protein quality control. *Cold Spring Harb. Perspect. Biol.* **3**, 1–14 (2011).

16. Ellis, R. J. O. H. N. The general concept of molecular chaperones. 257–261 (1993).

17. Ben-Zvi, A., De Los Rios, P., Dietler, G. & Goloubinoff, P. Active solubilization and refolding of stable protein aggregates by cooperative unfolding action of individual Hsp70 chaperones. *J. Biol. Chem.* **279**, 37298–37303 (2004).

18. McClellan, A. J., Scott, M. D. & Frydman, J. Folding and quality control of the VHL tumor suppressor proceed through distinct chaperone pathways. *Cell* **121**, 739–748 (2005).

19. Kaushik, S. & Cuervo, A. M. Chaperone-mediated autophagy: A unique way to enter the lysosome world. *Trends Cell Biol.* **22**, 407–417 (2012).

20. Albanèse, V., Yam, A. Y. W., Baughman, J., Parnot, C. & Frydman, J. Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells. *Cell* **124**, 75–88 (2006).

21. Haslbeck, M., Franzmann, T., Weinfurtner, D. & Buchner, J. Some like it hot: The structure and function of small heat-shock proteins. *Nat. Struct. Mol. Biol.* **12**, 842–846 (2005).

22. Young, J. C., Barral, J. M. & Hartl, F. U. More than folding: Localized functions of cytosolic chaperones. *Trends Biochem. Sci.* **28**, 541–547 (2003).

23. Nath, D. & Shadan, S. The ubiquitin system. *Nature* **458**, 421 (2009).

24. Erales, J. & Coffino, P. Ubiquitin-independent proteasomal degradation. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 216–221 (2014).

25. Tanaka, K. The proteasome: Overview of structure and functions. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* **85**, 12–36 (2009).

26. Shendure, J. & Akey, J. M. The origins, determinants, and consequences of

human mutations. *Science (80-.).* **349**, 1478–1483 (2015).

27.   Brookes, A. J. The essence of SNPs. *Gene* **234**, 177–186 (1999).

28.   Stein, A., Fowler, D. M., Hartmann-petersen, R. & Lindorff-larsen, K. Biophysical and Mechanistic Models for Disease-Causing Protein Variants. *Trends Biochem. Sci.* 1–14 (2019) doi:10.1016/j.tibs.2019.01.003.

29.   Zhou, Y., Hall, C. K. & Karplus, M. The calorimetric criterion for a two-state process revisited. *Protein Sci.* **8**, 1064–1074 (1999).

30.   King, K. J. L. and M. C. The Development of Transition-State Theory. *J. Phys. Chem.* **87**, 2657–2664 (1983).

31.   Sanchez-Ruiz, J. M. Protein kinetic stability. *Biophys. Chem.* **148**, 1–15 (2010).

32.   Gerasimavicius, L., Livesey, B. J. & Marsh, J. A. Loss-of-function, gain-of-function and dominant-negative mutations have profoundly different effects on protein structure. *Nat. Commun.* **13**, 1–15 (2022).

33.   Guo, Y. *et al.* A Rare KIF1A Missense Mutation Enhances Synaptic Function and Increases Seizure Activity. *Front. Genet.* **11**, 1–11 (2020).

34.   Meyer, K. *et al.* Mutations in Disordered Regions Can Cause Disease by Creating Dileucine Motifs. *Cell* **175**, 239-253.e17 (2018).

35.   Goehler, H. *et al.* A protein interaction network links GIT1, an enhancer of Huntingtin aggregation, to Huntington's disease (Molecular Cell (September 24, 2004) 15 (853-865) PII: S1097-2765(04)00545-3 and DOI: 10.1016/j.molcel.2004.09. 016). *Mol. Cell* **19**, 287 (2005).

36.   Yue, P., Li, Z. & Moult, J. Loss of protein structure stability as a major causative factor in monogenic disease. *J. Mol. Biol.* **353**, 459–473 (2005).

37.   Correia, A. R., Pastore, C., Adinolfi, S., Pastore, A. & Gomes, C. M. Dynamics, stability and iron-binding activity of frataxin clinical mutants. *FEBS J.* **275**, 3680–3690 (2008).

38.   Amorós, I. *et al.* Functional effects of a missense mutation in HERG associated with type 2 long QT syndrome. *Hear. Rhythm* **8**, 463–470 (2011).

39.   Tamura, H. *et al.* Liddle disease caused by a missense mutation of β subunit of

the epithelial sodium channel gene. *J. Clin. Invest.* **97**, 1780–1784 (1996).

40. Pauling, L., Itano, H. A., Singer, S. J. & Wells, I. C. 'Sickle cell anemia, a molecular disease'. *Science (80-.).* **110**, 1488–1490 (1949).

41. Ingram, V. M. Gene mutations in human haemoglobin: Difference between normal and sickle cell haemoglobin. *Nature* **180**, 326–328 (1957).

42. Pey, A. L., Serrano, L. & Martinez, A. Predicted Effects of Missense Mutations on Native-State Stability Account for Phenotypic Outcome in Phenylketonuria , a Paradigm of Misfolding Diseases. **81**, 1006–1024 (2007).

43. Mesa-Torres, N. *et al.* The Role of Protein Denaturation Energetics and Molecular Chaperones in the Aggregation and Mistargeting of Mutants Causing Primary Hyperoxaluria Type I. *PLoS One* **8**, (2013).

44. Coelho, A. I. *et al.* Functional and structural impact of the most prevalent missense mutations in classic galactosemia. *Mol. Genet. Genomic Med.* **2**, 484–496 (2014).

45. Quaresima, B. *et al.* Missense mutations of BRCA1 gene affect the binding with p53 both in vitro and in vivo. *Oncol. Rep.* **16**, 811–815 (2006).

46. Kuzmanov, U. & Emili, A. Protein-protein interaction networks: probing disease mechanisms using model systems. 1–12 (2013).

47. Yarwood, R., Hellicar, J., Woodman, P. G. & Lowe, M. Membrane trafficking in health and disease. *DMM Dis. Model. Mech.* **13**, (2020).

48. Duan, G. & Walther, D. The Roles of Post-translational Modifications in the Context of Protein Interaction Networks. *PLoS Comput. Biol.* **11**, 1–23 (2015).

49. Nurk, S. *et al.* The complete sequence of a human genome. *Science (80-.).* **376**, 44–53 (2022).

50. Albertin, W. *et al.* Linking post-translational modifications and variation of phenotypic traits. *Mol. Cell. Proteomics* **12**, 720–735 (2013).

51. Jin, H. & Zangar, R. C. Protein Modifications as Potential Biomarkers in Breast Cancer. *Biomark. Insights* **4**, 191–200 (2009).

52. Gong, C. X., Liu, F., Grundke-Iqbal, I. & Iqbal, K. Post-translational

modifications of tau protein in Alzheimer's disease. *J. Neural Transm.* **112**, 813–838 (2005).

53.   Morino, K. *et al.* Reduced mitochondrial density and increased IRS-1 serine phosphorylation in muscle of insulin-resistant offspring of type 2 diabetic parents. *J. Clin. Invest.* **115**, 3587–3593 (2005).

54.   Apweiler, R. *et al.* Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **40**, 71–75 (2012).

55.   Hunter, T. Why nature chose phosphate to modify proteins. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2513–2516 (2012).

56.   Ardito, F., Giuliani, M., Perrone, D., Troiano, G. & Muzio, L. Lo. The crucial role of protein phosphorylation in cell signalingand its use as targeted therapy (Review). *Int. J. Mol. Med.* **40**, 271–280 (2017).

57.   Cieśla, J., Fraczyk, T. & Rode, W. Phosphorylation of basic amino acid residues in proteins: Important but easily missed. *Acta Biochim. Pol.* **58**, 137–148 (2011).

58.   Iakoucheva, L. M. *et al.* The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049 (2004).

59.   Henriques, J. & Lindorff-Larsen, K. Protein Dynamics Enables Phosphorylation of Buried Residues in Cdk2/Cyclin-A-Bound p27. *Biophys. J.* **119**, 2010–2018 (2020).

60.   Li, P., Martins, I. R. S., Amarasinghe, G. K. & Rosen, M. K. Internal dynamics control activation and activity of the autoinhibited Vav DH domain. *Nat. Struct. Mol. Biol.* **15**, 613–618 (2008).

61.   Pérez-Mejías, G. *et al.* Exploring protein phosphorylation by combining computational approaches and biochemical methods. *Comput. Struct. Biotechnol. J.* **18**, 1852–1863 (2020).

62.   Engholm-Keller, K. & Larsen, M. R. Technologies and challenges in large-scale phosphoproteomics. *Proteomics* **13**, 910–931 (2013).

63.   Paradela, A. & Albar, J. P. Advances in the analysis of protein phosphorylation. *J. Proteome Res.* **7**, 1809–1818 (2008).

64. Tariq, A. *et al.* Potentiating Hsp104 activity via phosphomimetic mutations in the middle domain. *FEMS Yeast Res.* **18**, 1–14 (2018).

65. Maciejewski, P. M., Peterson, F. C., Anderson, P. J. & Brooks, C. L. Mutation of serine 90 to glutamic acid mimics phosphorylation of bovine prolactin. *J. Biol. Chem.* **270**, 27661–27665 (1995).

66. Stateva, S. R. *et al.* Characterization of phospho-(tyrosine)-mimetic calmodulin mutants. *PLoS One* **10**, 1–20 (2015).

67. Zisch, A. H. *et al.* Replacing two conserved tyrosines of the EphB2 receptor with glutamic acid prevents binding of SH2 domains without abrogating kinase activity and biological responses. *Oncogene* **19**, 177–187 (2000).

68. Fang, K. Y., Lieblich, S. A. & Tirrell, D. A. Incorporation of non-canonical amino acids into proteins by global reassignment of sense codons. *Methods Mol. Biol.* **1798**, 173–186 (2018).

69. Léger, J., Kempf, M., Lee, G. & Brandt, R. Conversion of serine to aspartate imitates phosphorylation-induced changes in the structure and function of microtubule-associated protein tau. *J. Biol. Chem.* **272**, 8441–8446 (1997).

70. Rothman, D. M. *et al.* Caged phosphoproteins. *J. Am. Chem. Soc.* **127**, 846–847 (2005).

71. Rogerson, D. T. *et al.* Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol.* **11**, 496–503 (2015).

72. Guerra-Castellano, A., Díaz-Moreno, I., Velázquez-Campoy, A., De La Rosa, M. A. & Díaz-Quintana, A. Structural and functional characterization of phosphomimetic mutants of cytochrome c at threonine 28 and serine 47. *Biochim. Biophys. Acta - Bioenerg.* **1857**, 387–395 (2016).

73. Levy, R., Gregory, E., Borcherds, W. & Daughdrill, G. p53 phosphomimetics preserve transient secondary structure but reduce binding to Mdm2 and MdmX. *Biomolecules* **9**, 1–14 (2019).

74. Zhang, M. S. *et al.* Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nat. Methods* **14**, 729–736 (2017).

75. Fan, C., Ip, K. & Söll, D. Expanding the genetic code of Escherichia coli with phosphotyrosine. *FEBS Lett.* **590**, 3040–3047 (2016).

76. Luo, X. *et al.* Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. *Nat. Chem. Biol.* **13**, 845–849 (2017).

77. Kawahata, N. *et al.* A novel phosphotyrosine mimetic 4′-carboxymethyloxy-3′-phosphonophenylalanine (Cpp): Exploitation in the design of nonpeptide inhibitors of pp60Src SH2 domain. *Bioorganic Med. Chem. Lett.* **11**, 2319–2323 (2001).

78. Choo-Kang, L. R. & Zeitlin, P. L. Induction of HSP70 promotes ΔF508 CFTR trafficking. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **281**, 58–68 (2001).

79. McDermott, C. J. Clinical trials in amyotrophic lateral sclerosis. *Curr. Opin. Neurol.* **32**, 758–763 (2019).

80. Liguori, L. *et al.* Pharmacological chaperones: A therapeutic approach for diseases caused by destabilizing missense mutations. *Int. J. Mol. Sci.* **21**, (2020).

81. Diamant, S., Eliahu, N., Rosenthal, D. & Goloubinoff, P. Chemical Chaperones Regulate Molecular Chaperones in Vitro and in Cells under Combined Salt and Heat Stresses. *J. Biol. Chem.* **276**, 39586–39591 (2001).

82. Pradeep, L. & Udgaonkar, J. B. Osmolytes induce structure in an early intermediate on the folding pathway of barstar. *J. Biol. Chem.* **279**, 40303–40313 (2004).

83. Tsukimura, T. *et al.* Molecular mechanism for stabilization of a mutant α-galactosidase A involving M51I amino acid substitution by imino sugars. *Mol. Genet. Metab.* **103**, 26–32 (2011).

84. Boyd, R. E. *et al.* Pharmacological chaperones as therapeutics for lysosomal storage diseases. *J. Med. Chem.* **56**, 2705–2725 (2013).

85. Benito, J. M., García Fernández, J. M. & Mellet, C. O. Pharmacological chaperone therapy for Gaucher disease: A patent review. *Expert Opin. Ther. Pat.* **21**, 885–903 (2011).

86. Maegawa, G. H. B. *et al.* Identification and characterization of ambroxol as an

enzyme enhancement agent for Gaucher disease. *J. Biol. Chem.* **284**, 23502–23516 (2009).

87. Santos-sierra, S. *et al.* Novel pharmacological chaperones that correct phenylketonuria in mice. *Hum. Mol. Genet.* **21**, 1877–1887 (2012).

88. Carlile, G. W. *et al.* A novel triple combination of pharmacological chaperones improves F508del-CFTR correction. *Sci. Rep.* **8**, 1–16 (2018).

89. Kopytova, A. E. *et al.* Potential Binding Sites of Pharmacological Chaperone NCGC00241607 on Mutant β-Glucocerebrosidase and Its Efficacy on Patient-Derived Cell Cultures in Gaucher and Parkinson's Disease. *Int. J. Mol. Sci.* **24**, 1–17 (2023).

90. Martínez-limón, A., Calloni, G., Ernst, R. & Vabulas, R. M. Flavin dependency undermines proteome stability , lipid metabolism and cellular proliferation during vitamin B2 de fi ciency. *Cell Death Dis.* (2020) doi:10.1038/s41419-020-02929-5.

91. Martínez-Limón, A. *et al.* Recognition of enzymes lacking bound cofactor by protein quality control. *Proc. Natl. Acad. Sci.* **113**, 12156–12161 (2016).

92. Ames, B. N., Elson-Schwab, I. & Silver, E. A. High-dose vitamin therapy stimulates variant enzymes with decreased coenzyme binding affinity (increased Km): Relevance to genetic disease and polymorphisms. *Am. J. Clin. Nutr.* **75**, 616–658 (2002).

93. Hirano, A., Braas, D., Fu, Y. H. & Ptáček, L. J. FAD Regulates CRYPTOCHROME Protein Stability and Circadian Clock in Mice. *Cell Rep.* **19**, 255–266 (2017).

94. Pey, A. L., Megarity, C. F. & Timson, D. J. FAD binding overcomes defects in activity and stability displayed by cancer-associated variants of human NQO1. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1842**, 2163–2173 (2014).

95. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).

96. Auer, P. L. & Lettre, G. Rare variant association studies: Considerations, challenges and opportunities. *Genome Med.* **7**, 1–11 (2015).

97. Martin, H. C. *et al.* Quantifying the contribution of recessive coding variation to developmental disorders Hilary. **1164**, 1161–1164 (2018).

98. Roscoe, B. P., Thayer, K. M., Zeldovich, K. B., Fushman, D. & Bolon, D. N. A. Analyses of the effects of all ubiquitin point mutants on yeast growth rate. *J. Mol. Biol.* **425**, 1363–1377 (2013).

99. Gray, V. E., Hause, R. J. & Fowler, D. M. Analysis of Large-Scale Mutagenesis Data To Assess the Impact of Single Amino Acid Substitutions. **207**, 53–61 (2017).

100. Zvaritch, E. *et al.* An Ryr1I4895T mutation abolishes Ca2+ release channel function and delays development in homozygous offspring of a mutant mouse line. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 18537–18542 (2007).

101. Sun, G. *et al.* Loss of Function Mutation in ELF4 Causes Autoinflammatory and Immunodeficiency Disease in Human. *J. Clin. Immunol.* **42**, 798–810 (2022).

102. Hou, J. W. *et al.* Loss-of-function CFTR p.G970D missense mutation might cause congenital bilateral absence of the vas deferens and be associated with impaired spermatogenesis. *Asian J. Androl.* **25**, 58–65 (2023).

103. Mavor, D. *et al.* Extending chemical perturbations of the ubiquitin fitness landscape in a classroom setting reveals new constraints on sequence tolerance. *Biol. Open* **7**, (2018).

104. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).

105. Findlay, G. M. *et al.* Accurate classification of BRCA1 variants with saturation genome editing. *Nature* **562**, 217–222 (2018).

106. Gupta, K. & Varadarajan, R. Insights into protein structure, stability and function from saturation mutagenesis. *Curr. Opin. Struct. Biol.* **50**, 117–125 (2018).

107. Kroncke, B. M. *et al.* Documentation of an Imperative to Improve Methods for Predicting Membrane Protein Stability. *Biochemistry* **55**, 5002–5009 (2016).

108. Riesselman, A. J., Ingraham, J. B. & Marks, D. S. Deep generative models of genetic variation capture the effects of mutations. *Nat. Methods* **15**, 816–822 (2018).

109. Lienhart, W. D., Gudipati, V. & MacHeroux, P. The human flavoproteome. *Arch. Biochem. Biophys.* **535**, 150–162 (2013).

110. Lucas, T. G., Henriques, B. J. & Gomes, C. M. Conformational analysis of the riboflavin-responsive ETF:QO-p.Pro456Leu variant associated with mild multiple acyl-CoA dehydrogenase deficiency. *Biochim. Biophys. Acta - Proteins Proteomics* **1868**, (2020).

111. Beaver, S. K., Mesa-Torres, N., Pey, A. L. & Timson, D. J. NQO1: A target for the treatment of cancer and neurological diseases, and a model to understand loss of function disease mechanisms. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2019) doi:10.1016/j.bbapap.2019.05.002.

112. Chen, L. Z. *et al.* A refined physical map of the long arm of human chromosome 16. *Genomics* **10**, 308–312 (1991).

113. Siegel, D. & Ross, D. Immunodetection of NAD(P)H:quinone oxidoreductase 1 (NQO1) in human tissues. *Free Radic. Biol. Med.* **29**, 246–253 (2000).

114. Belinsky, M. & Jaiswal, A. K. NAD(P)H:Quinone oxidoreductase1 (DT-diaphorase) expression in normal and tumor tissues. *Cancer Metastasis Rev.* **12**, 103–117 (1993).

115. Nioi, P. & Hayes, J. D. Contribution of NAD (P) H : quinone oxidoreductase 1 to protection against carcinogenesis , and regulation of its gene by the Nrf2 basic-region leucine zipper and the arylhydrocarbon receptor basic helix-loop-helix transcription factors. **555**, 149–171 (2004).

116. Venugopal, R. & Jaiswal, A. K. Nrf2 and Nrf1 in association with Jun proteins regulate antioxidant response element-mediated expression and coordinated induction of genes encoding detoxifying enzymes. *Oncogene* **17**, 3145–3156 (1998).

117. Asher, G., Dym, O., Tsvetkov, P., Adler, J. & Shaul, Y. The crystal structure of NAD(P)H quinone oxidoreductase 1 in complex with its potent inhibitor

dicoumarol. *Biochemistry* (2006) doi:10.1021/bi0600087.

118. Megarity, C. F. *et al.* Negative Cooperativity in NAD(P)H Quinone Oxidoreductase 1 (NQO1). *ChemBioChem* **20**, 2841–2849 (2019).

119. Bianchet, M. A., Faig, M. & Amzel, L. M. Structure and Mechanism of NAD[P]H:Quinone Acceptor Oxidoreductases (NQO). *Methods Enzymol.* **382**, 144–174 (2004).

120. L. Pey, A., F. Megarity, C., Medina-Carmona, E. & J. Timson, D. Natural Small Molecules as Stabilizers and Activators of Cancer-Associated NQO1 Polymorphisms. *Curr. Drug Targets* **17**, 1506–1514 (2016).

121. Preethi, S., Arthiga, K., Patil, A. B., Spandana, A. & Jain, V. Review on NAD(P)H dehydrogenase quinone 1 (NQO1) pathway. *Mol. Biol. Rep.* **49**, 8907–8924 (2022).

122. Lind, C., Hochstein, P. & Ernster, L. DT-diaphorase as a quinone reductase: A cellular control device against semiquinone and superoxide radical formation. *Arch. Biochem. Biophys.* **216**, 178–185 (1982).

123. Thor, H. *et al.* The metabolism of menadione (2-methyl-1,4-naphthoquinone) by isolated hepatocytes. A study of the implications of oxidative stress in intact cells. *J. Biol. Chem.* **257**, 12419–12425 (1982).

124. Wefers, H. & Sies, H. Hepatic low-level chemiluminescence during redox cycling of menadione and the menadione-glutathione conjugate: Relation to glutathione and NAD(P)H:quinone reductase (DT-diaphorase) activity. *Arch. Biochem. Biophys.* **224**, 568–578 (1983).

125. Kensler, T. W., Wakabayashi, N. & Biswal, S. Cell survival responses to environmental stresses via the Keap1-Nrf2-ARE pathway. *Annu. Rev. Pharmacol. Toxicol.* **47**, 89–116 (2007).

126. Beyer, R. E. *et al.* The role of DT-diaphorase in the maintenance of the reduced antioxidant form of coenzyme Q in membrane systems. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 2528–2532 (1996).

127. Siegel, D. *et al.* NAD(P)H:quinone oxidoreductase 1: Role as a superoxide scavenger. *Mol. Pharmacol.* **65**, 1238–1247 (2004).

128. Ross, D. & Siegel, D. Functions of NQO1 in cellular protection and CoQ10 metabolism and its potential role as a redox sensitive molecular switch. *Frontiers in Physiology* (2017) doi:10.3389/fphys.2017.00595.

129. Boothman, D. A., Meyers, M., Fukunaga, N. & Lee, S. W. Isolation of x-ray-inducible transcripts from radioresistant human melanoma cells. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 7200–7204 (1993).

130. Anwar, A. *et al.* Interaction of human NAD(P)H:Quinone oxidoreductase 1 (NQO1) with the tumor suppressor protein p53 in cells and cell-free systems. *J. Biol. Chem.* **278**, 10368–10373 (2003).

131. Asher, G., Lotem, J., Cohen, B., Sachs, L. & Shaul, Y. Regulation of p53 stability and p53-dependent apoptosis by NADH quinone oxidoreductase 1. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 1188–1193 (2001).

132. Asher, G., Tsvetkov, P., Kahana, C. & Shaul, Y. A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73. *Genes Dev.* **19**, 316–321 (2005).

133. Garate, M., Wong, R. P. C., Campos, E. I., Wang, Y. & Li, G. NAD(P)H quinone oxidoreductase 1 inhibits the proteasomal degradation of the tumour suppressor p33ING1b. *EMBO Rep.* **9**, 576–581 (2008).

134. Asher, G. *et al.* p53 hot-spot mutants are resistant to ubiquitin-independent degradation by increased binding to NAD(P)H:quinone oxidoreductase 1. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 15065–15070 (2003).

135. Asher, G., Bercovich, Z., Tsvetkov, P., Shaul, Y. & Kahana, C. 20S proteasomal degradation of ornithine decarboxylase is regulated by NQO1. *Mol. Cell* **17**, 645–655 (2005).

136. Adamovich, Y. *et al.* The Protein Level of PGC-1α, a Key Metabolic Regulator, Is Controlled by NADH-NQO1. *Mol. Cell. Biol.* **33**, 2603–2613 (2013).

137. Oh, E. T. *et al.* NQO1 inhibits proteasome-mediated degradation of HIF-1α. *Nat. Commun.* **7**, 1–14 (2016).

138. Sollner, S. *et al.* Quinone reductase acts as a redox switch of the 20S yeast proteasome. *EMBO Rep.* **10**, 65–70 (2009).

139. Moscovitz, O. *et al.* A Mutually Inhibitory Feedback Loop between the 20S Proteasome and Its Regulator, NQO1. *Mol. Cell* (2012) doi:10.1016/j.molcel.2012.05.049.

140. Di Francesco, A. *et al.* Novel RNA-binding activity of NQO1 promotes SERPINA1 mRNA translation. *Free Radic. Biol. Med.* **99**, 225–233 (2016).

141. Di Francesco, A. *et al.* NQO1 protects obese mice through improvements in glucose and lipid metabolism. *npj Aging Mech. Dis.* **6**, 1–18 (2020).

142. ASHER BEGLEITER, ERICA ROBOTHAM, and M. K. L. Role of NAD(P)H :(Quinone Acceptor) Oxidoreductase in Activation of Mitomycin C under Hypoxia.

143. Choi, E. K. *et al.* Upregulation of NAD(P)H:quinone oxidoreductase by radiation potentiates the effect of bioreductive β-lapachone on cancer cells. *Neoplasia* **9**, 634–642 (2007).

144. Digby, T., Leith, M. K., Thliveris, J. A. & Begleiter, A. Effect of NQO1 induction on the antitumor activity of RH1 in human tumors in vitro and in vivo. *Cancer Chemother. Pharmacol.* **56**, 307–316 (2005).

145. Oh, E. T. & Park, H. J. Implications of NQO1 in cancer therapy. *BMB Rep.* **48**, 609–617 (2015).

146. Bamford, S. *et al.* The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br. J. Cancer* **91**, 355–358 (2004).

147. Lajin, B. & Alachkar, A. The NQO1 polymorphism C609T (Pro187Ser) and cancer susceptibility: A comprehensive meta-analysis. *Br. J. Cancer* **109**, 1325–1337 (2013).

148. Liu, Y. & Zhang, D. The NQO1 C609T Polymorphism and Risk of Lung Cancer: a Meta-analysis. *Asian Pacific J. Cancer Prev.* **12**, 3091–3095 (2011).

149. Traver, R. D. *et al.* Characterization of a polymorphism in NAD(P)H: quinone oxidoreductase (DT-diaphorase). *Br. J. Cancer* **75**, 69–75 (1997).

150. Peng, Q. *et al.* The NQO1 Pro187Ser polymorphism and breast cancer susceptibility : evidence from an updated meta-analysis. (2014)

doi:10.1186/1746-1596-9-100.

151. Goracy, J. *et al.* The analysis of NADPH quinone reductase 1 (NQO1) polymorphism in Polish patients with colorectal cancer. *Biomolecules* **11**, 1–9 (2021).

152. Lienhart, W. *et al.* Collapse of the native structure caused by a single amino acid exchange in human NAD(P)H:quinone oxidoreductase 1. *FEBS J.* **281**, 4691–4704 (2014).

153. Medina-Carmona, E. *et al.* Conformational dynamics is key to understanding loss-of-function of NQO1 cancer-associated polymorphisms and its correction by pharmacological ligands. *Sci. Rep.* **6**, 1–14 (2016).

154. Medina-Carmona, E. *et al.* Site-to-site interdomain communication may mediate different loss-of-function mechanisms in a cancer-associated NQO1 polymorphism. *Sci. Rep.* **7**, 1–18 (2017).

155. Pey, A. L. Biophysical and functional perturbation analyses at cancer-associated P187 and K240 sites of the multifunctional NADP(H):quinone oxidoreductase 1. *Int. J. Biol. Macromol.* **118**, 1912–1923 (2018).

156. Medina-Carmona, E. *et al.* Enhanced vulnerability of human proteins towards disease-associated inactivation through divergent evolution. *Hum. Mol. Genet.* **26**, 3531–3544 (2017).

157. Vankova, P., Salido, E., Timson, D. J., Man, P. & Pey, A. L. A dynamic core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer. *Biomolecules* **9**, 1–17 (2019).

158. Eguch-Ishimae, M. *et al.* The association of a distinctive allele of NAD(P)H:quinone oxidoreductase with pediatric acute lymphoblastic leukemias with MLL fusion genes in Japan. *Haematologica* (2005).

159. Pan, S. shu, Akman, S. A., Forrest, G. L., Hipsher, C. & Johnson, R. The role of NAD(P)H: quinone oxidoreductase in mitomycin C- and porfiromycin-resistant HCT 116 human colon-cancer cells. *Cancer Chemother. Pharmacol.* **31**, 23–31 (1992).

160. Sato, M., Takagi, M. & Mizutani, S. Irradiation-induced p53 expression is

attenuated in cells with NQO1 C465T polymorphism. *J. Med. Dent. Sci.* (2010).

161. Hu, L., Stamberg, J. & Pan, S. The NAD(P)H:quinone oxidoreductase locus in human colon carcinoma HCT 116 cells resistant to mitomycin C. *Cancer Res.* **56**, 5253–5259 (1996).

162. Lienhart, W. D. *et al.* Catalytic competence, structure and stability of the cancer-associated R139W variant of the human NAD(P)H:quinone oxidoreductase 1 (NQO1). *FEBS J.* **284**, 1233–1245 (2017).

163. Lee, W. S., Ham, W. & Kim, J. Roles of NAD(P)H:quinone Oxidoreductase 1 in Diverse Diseases. *Life* **11**, 1–20 (2021).

164. Bolton, J. L., Trush, M. A., Penning, T. M., Dryhurst, G. & Monks, T. J. Role of quinones in toxicology. *Chem. Res. Toxicol.* **13**, 135–160 (2000).

165. Geng, R. *et al.* Oxidative stress-related genetic polymorphisms are associated with the prognosis of metastatic gastric cancer patients treated with epirubicin, oxaliplatin and 5-fluorouracil combination chemotherapy. *PLoS One* **9**, 1–15 (2014).

166. Fagerholm, R. *et al.* NAD(P)H:quinone oxidoreductase 1 NQO1*2 genotype (P187S) is a strong prognostic and predictive factor in breast cancer. *Nat. Genet.* **40**, 844–853 (2008).

167. Lewis, A. *et al.* Treatment of pancreatic cancer cells with dicumarol induces cytotoxicity and oxidative stress. *Clin. Cancer Res.* **10**, 4550–4558 (2004).

168. Beaver, S. K., Mesa-Torres, N., Pey, A. L. & Timson, D. J. NQO1: A target for the treatment of cancer and neurological diseases, and a model to understand loss of function disease mechanisms. *Biochim. Biophys. Acta - Proteins Proteomics* **1867**, 663–676 (2019).

169. Calabrese, V., Bates, T. E. & Giuffrida Stella, A. M. NO Synthase and NO-Dependent Signal Pathways in Brain Aging and Neurodegenerative Disorders: The Role of Oxidant/Antioxidant Balance. *Neurochem. Res.* **25**, 1315–1341 (2000).

170. Stringer, J. L. *et al.* Presence and Induction of the Enzyme NAD(P)H: Quinone Oxidoreductase 1 in the Central Nervous System. *J. Comp. Neurol.* **471**, 289–297

(2004).

171. Schultzberg, M., Segura-Aouilar, J. & Lind, C. Distribution of dt diaphorase in the rat brain: Biochemical and immunohistochemical studies. *Neuroscience* **27**, 763–776 (1988).

172. Floor, E. & Wetzel, M. G. Increased protein oxidation in human substantia nigra pars compacta in comparison with basal ganglia and prefrontal cortex measured with an improved dinitrophenylhydrazine assay. *J. Neurochem.* **70**, 268–275 (1998).

173. Pae, C. U. *et al.* Quinone oxidoreductase (NQO1) gene polymorphism (609C/T) may be associated with tardive dyskinesia, but not with the development of schizophrenia. *Int. J. Neuropsychopharmacol.* **7**, 495–500 (2004).

174. Okada, S. *et al.* No associations between Parkinson's disease and polymorphisms of the quinone oxidoreductase (NQO1, NQO2) genes. *Neurosci. Lett.* **375**, 178–180 (2005).

175. Luo, J. *et al.* Association of the NQO1 C609T polymorphism with Alzheimer's disease in Chinese populations: A meta-analysis. *Int. J. Neurosci.* **126**, 199–204 (2016).

176. Zhu, H. *et al.* The highly expressed and inducible endogenous NAD(P)H:quinone oxidoreductase 1 in cardiovascular cells acts as a potential superoxide scavenger. *Cardiovasc. Toxicol.* **7**, 202–211 (2007).

177. Niemann, B. *et al.* Oxidative Stress and Cardiovascular Risk: Obesity, Diabetes, Smoking, and Pollution: Part 3 of a 3-Part Series. *J. Am. Coll. Cardiol.* **70**, 230–251 (2017).

178. Shyu, H. Y. *et al.* Genotype polymorphisms of GGCX, NQO1, and VKORC1 genes associated with risk susceptibility in patients with large-artery atherosclerotic stroke. *Clin. Chim. Acta* **411**, 840–845 (2010).

179. Han, S. J. *et al.* The C609T variant of NQO1 is associated with carotid artery plaques in patients with type 2 diabetes. *Mol. Genet. Metab.* **97**, 85–90 (2009).

180. Gaikwad, A., Long, D. J., Stringer, J. L. & Jaiswal, A. K. In Vivo Role of NAD(P)H:Quinone Oxidoreductase 1 (NQO1) in the Regulation of Intracellular

Redox State and Accumulation of Abdominal Adipose Tissue. *J. Biol. Chem.*
276, 22559–22564 (2001).

181. Pey, A. L. Phenotypic Modulation of Cancer-Associated Antioxidant NQO1
      Activity by Post-Translational Modifications and the Natural Diversity of the
      Human Genome. *Antioxidants* **12**, (2023).

182. Hornbeck, P. V. *et al.* PhosphoSitePlus: A comprehensive resource for
      investigating the structure and function of experimentally determined post-
      translational modifications in man and mouse. *Nucleic Acids Res.* **40**, 261–270
      (2012).

183. Medina-Carmona, E. *et al.* Phosphorylation compromises FAD binding and
      intracellular stability of wild-type and cancer-associated NQO1: Insights into
      flavo-proteome stability. *Int. J. Biol. Macromol.* **125**, 1275–1288 (2019).

184. Luo, S. *et al.* Akt Phosphorylates NQO1 and Triggers its Degradation,
      Abolishing its Antioxidative Activities in Parkinson's Disease. *J. Neurosci.*
      (2019) doi:10.1523/jneurosci.0625-19.2019.

185. Siegel, D. *et al.* Redox state and the sirtuin deacetylases are major factors that
      regulate the acetylation status of the stress protein NQO1. *Front. Pharmacol.* **13**,
      1–11 (2022).

186. Muñoz, I. G., Morel, B., Medina-Carmona, E. & Pey, A. L. A mechanism for
      cancer-associated inactivation of NQO1 due to P187S and its reactivation by the
      consensus mutation H80R. *FEBS Letters* (2017) doi:10.1002/1873-3468.12772.

187. Medina-Carmona, E. *et al.* Insight into the specificity and severity of pathogenic
      mechanisms associated with missense mutations through experimental and
      structural perturbation analyses. *Hum. Mol. Genet.* **28**, 1–15 (2019).

188. Fontana, A. *et al.* Probing protein structure by limited proteolysis. *Acta Biochim.
      Pol.* **51**, 299–321 (2004).

189. Siegel, D. *et al.* Rapid polyubiquitination and proteasomal degradation of a
      mutant form of NAD(P)H:quinone oxidoreductase 1. *Mol. Pharmacol.* **59**, 263–
      268 (2001).

190. Leaver-Fay, A. *et al.* Rosetta3: An object-oriented software suite for the

simulation and design of macromolecules. *Methods Enzymol.* **487**, 545–574 (2011).

191. Narang, D., Lento, C. & Wilson, D. J. HDX-MS: An analytical tool to capture protein motion in action. *Biomedicines* **8**, 1–20 (2020).

192. Yamada, Y. *et al.* Catalytic Inactivation of Human Phospholipase D2 by a Naturally Occurring Gly901Asp Mutation. *Arch. Med. Res.* **37**, 696–699 (2006).

193. Stevanin, G. Mutation in the Catalytic Domain of Protein Kinase. **61**, (2004).

194. Zhang, Z., Miteva, M. A., Wang, L. & Alexov, E. Analyzing effects of naturally occurring missense mutations. *Comput. Math. Methods Med.* **2012**, (2012).

195. Takamiya, O., Seta, M., Tanaka, K. & Ishida, F. Human factor VII deficiency caused by S339C mutation located adjacent to the specificity pocket of the catalytic domain. *Clin. Lab. Haematol.* **24**, 233–238 (2002).

196. Koukouritaki, S. B. *et al.* Identification and functional analysis of common human flavin-containing monooxygenase 3 genetic variants. *J. Pharmacol. Exp. Ther.* **320**, 266–273 (2007).

197. Caldararu, O., Blundell, T. L. & Kepp, K. P. Three Simple Properties Explain Protein Stability Change upon Mutation. *J. Chem. Inf. Model.* **61**, 1981–1988 (2021).

198. Siegel, D., McGuinness, S. M., Winski, S. L. & Ross, D. Genotype-phenotype relationships in studies of a polymorphism in NAD(P)H:quinone oxidoreductase 1. *Pharmacogenetics* (1999).

199. Sabbir, M. G., Taylor, C. G. & Zahradka, P. CAMKK2 regulates mitochondrial function by controlling succinate dehydrogenase expression, post-translational modification, megacomplex assembly, and activity in a cell-type-specific manner. *Cell Commun. Signal.* **19**, 1–29 (2021).

200. Bharathi, S. S. *et al.* Sirtuin 3 (SIRT3) protein regulates long-chain acyl-CoA dehydrogenase by deacetylating conserved lysines near the active site. *J. Biol. Chem.* **288**, 33837–33847 (2013).

201. Wu, C. *et al.* Global and Site-Specific Effect of Phosphorylation on Protein

Turnover. *Dev. Cell* **56**, 111-124.e6 (2021).

202. Khan, S. H., McLaughlin, W. A. & Kumar, R. Site-specific phosphorylation regulates the structure and function of an intrinsically disordered domain of the glucocorticoid receptor. *Sci. Rep.* **7**, 1–8 (2017).

203. Ponnam, S., Sevrieva, I., Sun, Y. B., Irving, M. & Kampourakis, T. Site-specific phosphorylation of myosin binding protein-C coordinates thin and thick filament activation in cardiac muscle. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 15485–15494 (2019).

204. Schymkowitz, J. *et al.* The FoldX web server: An online force field. *Nucleic Acids Res.* **33**, 382–388 (2005).

205. Sun, C., Zhao, W., Wang, X., Sun, Y. & Chen, X. A pharmacological review of dicoumarol: An old natural anticoagulant agent. *Pharmacol. Res.* **160**, 105193 (2020).

206. Collier, A. C. & Pritsos, C. A. The mitochondrial uncoupler dicumarol disrupts the MTT assay. *Biochem. Pharmacol.* **66**, 281–287 (2003).

207. Betancor-Fernández, I., Timson, D. J., Salido, E. & Pey, A. L. Natural (and Unnatural) Small Molecules as Pharmacological Chaperones and Inhibitors in Cancer. *Target. Traffick. Drug Dev.* 155–190 (2018).

208. Nolan, K. A. *et al.* Coumarin-Based Inhibitors of Human NAD(P)H:Quinone Oxidoreductase-1. Identification, Structure–Activity, Off-Target Effects and In Vitro Human Pancreatic Cancer Toxicity. *J. Med. Chem.* **50**, 6316–6325 (2007).

209. Cheng, J. *et al.* Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science (80-.).* **7492**, (2023).

210. Cresteil, T. & Jaiswal, a K. High levels of expression of the NAD(P)H:quinone oxidoreductase (NQO1) gene in tumor cells compared to normal cells of the same origin. *Biochem. Pharmacol.* **42**, 1021–1027 (1991).