

**Probing vulnerability of the gp41 C-terminal heptad repeat as target for
miniprotein HIV inhibitors**

Samuel Jurado¹, Christiane Moog², Mario Cano-Muñoz¹, Sylvie Schmidt², Géraldine Laumond², Valentina Ruocco¹, Sara Standoli¹, Daniel Polo-Megías¹, Francisco Conejero-Lara^{1,*} and Bertrand Morel^{1,*}

¹Departamento de Química Física, Instituto de Biotecnología e Unidad de Excelencia de Química Aplicada a Biomedicina y Medioambiente (UEQ), Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain.

²INSERM U1109, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Université de Strasbourg, Strasbourg, France.

***Corresponding authors:**

B. Morel; e-mail: bmorel@ugr.es.

F. Conejero-Lara; tel. +34 958 242371; e-mail: conejero@ugr.es.

Abstract (limit of 250 words)

One of the therapeutic strategies in HIV neutralization is blocking membrane fusion. In this process, tight interaction between the N-terminal and C-terminal heptad-repeat (NHR and CHR) regions of gp41 is essential to promote membranes apposition and merging. We have previously developed single-chain proteins (named covNHR) that accurately mimic the complete gp41 NHR region in its trimeric conformation. They tightly bind CHR-derived peptides and show a potent and broad HIV inhibitory activity *in vitro*. However, the extremely high binding affinity (sub-picomolar) is not in consonance with their inhibitory activity (nanomolar), likely due to partial or temporal accessibility of their target in the virus. Here, we have designed and characterized two single-chain covNHR miniproteins each encompassing one of the two halves of the NHR region and containing two of the four sub-pockets of the NHR crevice. The two miniproteins fold as trimeric helical bundles as expected but while the C-terminal covNHR (covNHR-C) miniprotein is highly stable, the N-terminal counterpart (covNHR-N) shows only marginal stability that could be improved by engineering an internal disulfide bond. Both miniproteins bind their respective complementary CHR peptides with moderate (micromolar) affinity. Moreover, the covNHR-N miniproteins can access their target in the context of trimeric native envelope proteins and show significant inhibitory activity for several HIV pseudoviruses. In contrast, covNHR-C cannot bind its target sequence and neither inhibits HIV, indicating a higher vulnerability of C-terminal part of CHR. These results may guide the development of novel HIV inhibitors targeting the gp41 CHR region.

Keywords

Fusion inhibitors; binding affinity; coiled-coil; envelope glycoprotein; hydrophobic pocket.

Abbreviations

Env, HIV-1 envelope glycoprotein; gp41, glycoprotein subunit 41; gp120, glycoprotein subunit 120; CHR, C-terminal heptad repeat; NHR, N-terminal heptad repeat; CTP, C-terminal pocket; NTP, N-terminal pocket; MP, middle pocket; HP,

hydrophobic pocket; CD, circular dichroism spectroscopy; ITC, isothermal titration calorimetry; DLS, dynamic light scattering.

Highlights

- Design, production and characterization of mini protein HIV inhibitors
- Engineered disulfide bond increases conformational stability
- Protein stabilization enhances inhibitory capacity
- NHR coiled-coil of gp41 is composed of two subdomains with different conformational stability
- Vulnerability of the gp41 CHR region as inhibition target

Introduction.

Infection by the human immunodeficiency virus (HIV) is still among the worst global pandemics, with nearly 1 million deaths per year directly or indirectly associated to the infection and the rate of new infections remains high at 1.7 million per year. However, only a few vaccine candidates have reached phase IIb or III clinical tests with poor or null efficacy. For these reasons, the development of an effective, safe and affordable treatment against HIV and new improved antivirals or microbicides are highly desirable against the infection. During the last decades, antiretroviral therapies have improved considerably the life expectancy of patients infected by HIV. Combined therapies with several drugs (HAART) have raised the genetic barrier for evolution of viral resistance and have been able to reduce the number of copies in the blood stream of patients to undetectable levels, keeping the virus latent integrated in the DNA of infected cells. Nonetheless, the emergence of multidrug-resistant viral strains has been observed in an increasing number of patients ¹ and in some cases serious interactions between drugs have appeared ². It is therefore necessary to expand the repertoire of potent inhibitors targeting different stages of the HIV replication cycle.

Current efforts in this field have been focused on preventing the cell-virus fusion by blocking the HIV envelope protein (Env) ^{3;4}. Env is a non-covalently associated trimer of heterodimers of two glycoprotein subunits, gp120 and gp41 playing crucial roles in HIV cell infection ^{5; 6; 7}. The gp120 sequential binding to CD4 receptor and one of CXCR4 or CCR5 co-receptors triggers a series of large conformational changes in gp41 that lead to membrane fusion and insertion of the viral content into the cell cytoplasm ^{8;9}. In this process, the gp41 subunits insert their N-terminal fusion peptides into the cell membrane and fold into a highly stable 6-helix-bundle (6HB) conformation, in which the N-terminal heptad repeat (NHR) regions form a central helical trimer and the three C-terminal heptad repeat (CHR) regions pack externally over them in an antiparallel orientation ^{10;11}. The energy gain from this folding process serves to bring membranes into close proximity and facilitate their fusion.

Abolishing the whole fusion stage and inhibiting the virus invasion by blocking any single step in the fusion process reveal to be effective strategies. Due to its key role

in membrane fusion, gp41 constitutes a very attractive target for the identification and development of fusion inhibitors. The protein NHR and CHR regions are temporarily exposed in a fusion intermediate state before interacting with each other, so some appropriate molecules able to recognize NHR or CHR regions could interfere with such interaction and act as fusion inhibitors¹². These are mainly CHR peptide mimetics^{13; 14; 15; 16}, artificial D-peptides¹⁷, small compounds^{18; 19} and also antibodies^{18; 19; 20; 21}. Currently, only one inhibitor from this type is approved for clinical use. T20, also known by its generic name as enfuvirtide, is a 36 amino acid polypeptide derived from the CHR and membrane proximal regions of gp41^{13; 22}. Unfortunately, its clinical use has been limited by its short half-life²³ (proteolysis-sensitive and rapid renal filtration) requiring therefore high dosage injections at least twice a day. Moreover, the continuous and expensive treatment generates the appearance of T20-resistant viruses. Subsequently, other peptides and small molecule compounds directed against the coiled-coil structure of NHR have been developed, although very few have reached clinical stage and present various difficulties in their development as suitable drugs^{24; 25; 26; 27}.

Even though peptides derived from the gp41 NHR region inhibit the entry of HIV by interacting with the CHR region, they show much less inhibitory potency due to their low conformational stability and higher tendency to aggregate. However, different strategies have succeeded stabilizing the NHR region in a trimeric helical conformation, with strong improvements in inhibitory potency^{4; 28; 29; 30; 31; 32}.

We have previously designed small protein constructs, called covNHR, consisting of a single polypeptide chain with three helical regions folding together to form an antiparallel trimeric bundle with a similar structure to the NHR region of gp41^{33; 34}. These highly stable and soluble proteins are produced recombinantly by expression in *E. coli* with good yields and do not need any posttranslational modifications. The covNHR proteins bind to peptides derived from the gp41 CHR region with very high affinity and inhibit HIV cell infection in standardized *in vitro* tests with high breadth and potency (IC₅₀: 5-10 nM), for a wide variety of HIV strains^{33; 34}. The strong anti HIV activities of these and other NHR based constructs demonstrate that the CHR region of gp41 is an appealing target for therapeutic strategies.

The NHR binding surface has been described as composed of four different hotspots, namely an N-terminal polar pocket (NTP) ³⁴, a shallow middle pocket (MP) ³³, a deep and prominent hydrophobic pocket (HP), which has been widely used as a drug discovery target ¹⁴, and finally a C-terminal pocket (CTP) adjacent to the HP ³⁵. We have reported very recently a quantitative dissection of the binding thermodynamics between the complementary CHR and NHR regions ³⁶. The binding energy is broadly but unevenly distributed along the interface ¹ and cooperative interactions implying at least two pockets are needed to obtain high affinity in the nM range ³⁶. However, the importance of each of these binding pockets in the fusion inhibition remains unclear. For example, mutations in a covNHR variant abolishing binding at the NTP produce a 10⁵ decrease in affinity for CHR peptides but only a small decrease in inhibitory potency ³⁴.

It is generally accepted that both CHR and NHR regions are exposed in the gp41 prefusion intermediate and become hidden simultaneously with 6HB formation. However, while the formation of an NHR trimeric coiled-coil after receptor engagement is a prerequisite for inhibition by CHR peptide mimetics and analogs ¹⁴, it is unclear whether all or part of the CHR region can be accessed by NHR mimics at an earlier stage.

In this study, we have designed two short versions of covNHR proteins, each containing only two of the four pockets of the NHR crevice. A miniprotein corresponding to the N-terminal part of NHR (covNHR-N) harbors only NTP and MP, whereas a second miniprotein mimicking the C-terminal half of NHR (covNHR-C) exposes only the HP and the CTP. The covNHR-N miniprotein proved to be quite unstable and was stabilized by engineering disulfide-bonds. Their biophysical and calorimetric characterizations have allowed us to measure the thermodynamic contributions to the CHR-NHR interaction in each variant. In addition, the capacity of these covNHR proteins to bind directly several versions of Env was assessed, and their capacity to inhibit different HIV-1 pseudoviruses, including T20-resistant strains was determined. The results improve our understanding of the role of the different NHR sub-sites in the inhibition targeting CHR and reveal that steric hindrances and conformational flexibility of gp41 are important in the vulnerability of the gp41 CHR region to the covNHR proteins. These results could serve as guide for the development of novel HIV-1 fusion inhibitors.

Results.

Design of covNHR miniproteins

We designed covNHR miniproteins harbouring only two of the four sub-pockets of the NHR crevice using the crystallographic structure of the complex between C34 and covNHR-VQ (PDB: 6R2G) as template. This covNHR variant binds C34 with sub-pM affinity³⁴. Three different miniproteins were initially designed, one called covNHR-N containing the two N-terminal pockets (NTP and MP), a second one named as covNHR-M encompassing the two central pockets (MP and HP), and a third protein, designated covNHR-C, with the two C-terminal pockets (HP and CTP). The molecules were built by clipping the NHR helices and connecting the new ends with short loops. The loops were manually built establishing helix capping interactions and maintaining favorable backbone psi and phi angles. While covNHR-C was easily expressed and purified and proved to be highly stable (see below), the initially designed covNHR-N and covNHR-M could not be produced because of aggregation and lack of *E. coli* expression, respectively. The propensity of the sequences to aggregation was tested with AGGRESCAN³⁷. Mutations were added to reduce aggregation propensity and stabilize the proteins with success for covNHR-N. However, several versions of covNHR-M did not show any expression so we abandoned this design. The models were finally energy minimized and submitted to Molecular Dynamics simulations to test their stability (Fig. S1) and the final sequences are provided in Figure 1b. The molecular weights (including the His-tag) are 11909.47 Da for covNHR-N and 12792.82 Da for covNHR-C.

Biophysical characterization of covNHR variants

The covNHR-N and covNHR-C miniproteins (Fig. 1) were produced recombinantly by *E. coli* expression with good yields. CovNHR-N is highly soluble at physiological pH whereas covNHR-C shows some propensity to aggregate at concentration higher than 20 μ M. CovNHR-C is monomeric according to DLS measurements in both pH 2.5 and pH 7.4 buffers (50 mM glycine/HCl, pH 2.5 and 50 mM sodium phosphate buffer, pH 7.4), showing apparent hydrodynamic radii of 1.8-1.9 nm (Fig. S2a and Table S1). While covNHR-N is also monomeric at pH 7.4 ($R_h=1.9$ nm),

its hydrodynamic radius at pH 2.5 is increased up to 3.1 nm, suggesting partial unfolding and/or self-association (Fig. S2b and Table S1).

The two miniproteins show Far-UV CD spectra typical of highly α -helical structure (Fig. 2a), consistent with the designs. The α -helix content of covNHR-N is 64% at pH 7.4, slightly lower than the expected value (72%), but appears to be partially unfolded at acidic pH (Fig. 2a). In contrast covNHR-C exhibits 70% of α -helix at acid and physiological pH, close to the theoretical one (74%).

Subsequently, the thermal stability of each protein was measured by monitoring the CD signal at 222 nm as a function of temperature (Fig. 2b). CovNHR-C protein is extremely stable at both pH 7.4 and pH 2.5, since no unfolding transitions were observed even after heating up to 98°C. In contrast, covNHR-N is relatively unstable at pH 7.4 showing its unfolding transition at 41°C (Fig. 2b and Table S1). This unfolding process appears to be partially reversible upon a second consecutive temperature scan. CovNHR-N is progressively less stable and becomes more expanded upon pH decrease (Table S1). Moreover, its stability and helicity are slightly reduced with an increase in the buffer ionic strength (Fig. S5a).

Stabilization of covNHR-N by engineering an internal disulfide bond.

Since the unfolding temperature of covNHR-N is quite low and close to the physiological temperature, we sought to further stabilize it. First, we carried out a limited proteolysis study of the protein using thermolysin as protease. The protein is quite sensitive to proteolysis at 20°C even at a low 1:10000 protein/protease ratio. Most sensitive cleavage points occur around loop 1 connecting helices 1 and 2 (Fig. S3-S4) suggesting a higher flexibility for this end of the protein. Then, using “Disulfide by Design” web server ³⁸ we found two possibilities for disulfide bond creation connecting different structural elements of the protein by X-Cys mutations (Fig. 3a): Gly32 at loop 1 and either of Thr92 or Arg93, at the C-terminus of the third helix. This region is precisely the less stable part of the protein according to limited proteolysis, so we decided to produce both disulfide-bonded mutants.

Both mutants covNHR-NSS1 (G32C/T92C) and covNHR-NSS2 (G32C/R93C) could be produced by *E. coli* expression with good yields although lower than covNHR-N. All purification steps were made in presence of 10 mM β -mercaptoethanol and a

final oxidation step was carried out by extensive dialysis with buffer without reducing agent. Formation of the disulfide bonds was confirmed in both mutants using Ellman's assay (ThermoFisher).

The structure and stability of the proteins were characterized under different conditions in presence and absence of 5 mM TCEP (Fig. 3b-c and Table S1). Both variants are monomeric in all conditions studied (Fig. S2d). They have α -helical structure similar to covNHR-N and slightly higher in the oxidized state than in the reduced one. The stability of both variants at pH 7.4 in their reduced state is only slightly higher than that of covNHR-N, with increases in T_m of +4°C. However, in the oxidized state the two variants are much more stable, with additional increase of +20°C in T_m . The thermal stability is also strongly increased in all other conditions (Table S1). Moreover, the two variants are much more resistant to proteolysis with thermolysin at 20°C, persisting uncleaved for several hours even in presence of 1:500 protein-to-protease ratio (Fig. S3). These results confirm a very similar and strong stabilization of the proteins produced by the disulfide bond in both variants.

Binding of the CHR peptides to covNHR variants

To characterize the interaction between the covNHR proteins and their respective CHR target regions in gp41, we analyzed the binding of the Y24L peptide (gp41 residues 638-661) to covNHR-N and its two disulfide-stabilized variants, as this peptide spans the N-terminal pocket (NTP) and MP (middle pocket) binding motifs, and the binding of N25S (gp41 residues 616-640) to covNHR-C because it encompasses the binding motifs of the C-terminal pocket (CTP) and the hydrophobic pocket (HP).

Far-UV CD spectra of the protein-peptide mixtures were compared with those of the free molecules (Fig. 4a). Y24L and N25S exhibit spectra typical of unfolded polypeptides. An increase in negative ellipticity was observed in the protein:peptide mixtures indicating acquisition of helical conformation for the peptide as a consequence of binding onto the NHR groove. The negative ellipticity at 222 nm for both complexes is not significantly different suggesting similar α -helical structure in the bound states. Nevertheless, the ellipticity increase is

considerably larger in the case of covNHR-N and its two stabilized variants, indicating a greater acquisition of helical structure in these complexes.

Thermal denaturation experiments of mixtures between Y24L with covNHR-N or the disulfide-bonded variants indicate considerable stabilization of the proteins produced by peptide binding (Fig. S5b-d). The slopes of the transitions increase, indicating an increase in the enthalpy of the denaturation of the complex, as a consequence of the interaction, and an increase in the cooperativity of the unfolding process. Unfortunately, this kind of analysis could not be performed with covNHR-C and N25S because of the extremely high thermal stability of the protein. To test whether the N25S peptide binds to the HP of covNHR-C in the expected conformation, we recorded the near-UV CD spectra of the free and bound states (Fig. 4b). Upon mixing with N25S, the near-UV CD spectrum of covNHR-C exhibits an intense shifted negative ellipticity band centered at 293 nm produced by sidechains stacking of the CHR tryptophan-rich motif onto the HP, as observed in the postfusion structure of gp41³⁹ and in the complexes between C34 and covNHR proteins^{33;34}.

Thermodynamics of peptide binding

To characterize the binding thermodynamics in detail, we performed ITC analysis by direct titration of the protein solutions with their corresponding peptides (Fig 5 a-b). In our recent study, we have shown that Y24L and N25S bind to covNHR-VQ³⁶ with different affinities. While the interaction between Y24L and covNHR-VQ has been found to be relatively tight ($K_d = 90 \pm 7$ nM at 25°C), a higher dissociation constant of 3.0 ± 0.2 μ M was obtained for N25S³⁶.

Here, we have found that Y24L binds to covNHR-N with 1:1 stoichiometry, as expected, since the covNHR proteins were engineered to bind a single peptide unit on the crevice between the two parallel helices³³. The binding shows a considerable negative binding enthalpy and a significantly reduced affinity ($K_d = 0.79 \pm 0.02$ μ M at 25°C) compared that of covNHR-VQ for this peptide at the same temperature (Table 1)³⁶. The binding enthalpy is considerably more negative by about -25 kJ mol⁻¹ (Fig. 5c) and there is a strong increase in binding heat capacity compared to the complete covNHR-VQ protein.

These changes in the thermodynamic magnitudes of binding may be related to a considerable influence of a structural stabilization of covNHR-N upon interaction with the Y24L peptide. This conformational stabilization upon binding was already suggested by the CD thermal unfolding experiments (Fig. S5b).

To further investigate the possible structural changes induced by peptide binding to covNHR-N, NMR ^1H - ^{15}N -HSQC spectra of the labelled protein were recorded in the absence and in complex with unlabelled Y24L peptide (Fig. 6). The spectrum of free protein (at a concentration of 800 μM) presents a poor resolution due to considerable line broadening and low signal dispersion. This indicates a low conformational stability and a certain tendency to aggregate at the high concentration of the NMR experiment. Drastic increases in chemical shift dispersion and signal resolution are observed in the HSQC spectrum in presence of Y24L at a saturating concentration (molar ratio 1:2), indicating strong conformational stabilization induced by complex formation. These data also corroborate perfectly the DLS experiments performed with covNHR-N at different concentrations and in complex with Y24L that also reveal a less expanded structure upon addition of Y24L (Fig. S2c). These experiments confirm a significant conformational instability of the covNHR-N miniprotein.

The thermodynamic magnitudes of Y24L binding are virtually identical for the covNHR-NSS1 and covNHR-NSS2 mutants and quite similar to those of covNHR-N (Fig. 5c and Table 1), except for slightly lower binding enthalpies and significantly less negative heat capacities of binding. This is consistent with their higher conformational stability produced by the disulfide bonds, as demonstrated by the considerable improvement of the chemical shift dispersion observed in the free covNHR-NSS1 HSQC spectrum (Fig. S6). Anyhow, the conformational stabilization produced by the disulfide-bonds has a small impact in the binding affinity of the proteins for the peptide.

The dissociation constant of the complex between N25S and covNHR-C is also relatively high ($K_d = 1.9 \pm 0.3 \mu\text{M}$ at 25°C) but in this case the value is comparable to that of covNHR-VQ with the same peptide ($K_d = 3 \pm 1 \mu\text{M}$ at 25°C). Although the two complexes differ in binding enthalpy by about 15 $\text{kJ}\cdot\text{mol}^{-1}$ at equal temperature (Fig. 5d), their binding heat capacities are identical (Table 1) ³⁶ suggesting a similar mode of binding of N25S to both proteins.

In summary, the miniproteins bind their respective complementary CHR peptides with moderate affinity. While the C-terminal domain encompassing the HP and the CTP retains a high stability and similar interactions with the CHR peptide when isolated from the parent molecule, the N-terminal domain, harbouring the NTP and MP, is intrinsically unstable and its affinity for its CHR target becomes considerably affected compared to covNHR-VQ.

Binding to envelope proteins

To investigate the interactions of the covNHR proteins to their respective gp41 CHR targets in a context more akin to a native-like envelope (Env) spike, we carried out binding experiments with several soluble stabilized Env proteins in precursor and cleaved forms (Table S2). Due to the low thermal stability of covNHR-N, these ELISA experiments were carried out at 25°C.

The three covNHR-N proteins showed a significant level of binding to all Envs (Fig. 7a). The binding level is similar for the three proteins and varies depending on the Env variant but in all cases the binding is weaker than for the parent protein covNHR-VQ. These results indicate that the C-terminal part of the gp41 CHR region from the different Envs in both uncleaved and cleaved trimeric pre-fusion conformations is accessible to interaction by the covNHR proteins encompassing the MP and NTP. In contrast, covNHR-C did not bind to any of the tested Env proteins indicating that the N-terminal region of CHR is less accessible in Envs pre-fusion conformation.

HIV-1 inhibition

Despite the fact that covNHR-C did not show any binding to Env proteins, there is a possibility that the CHR target could become temporarily exposed during the fusion process and accessible to inhibition. The inhibitory activities of the proteins were analyzed using the conventional TZM-bl assay and compared to that of T20. We have used two different HIV-1 pseudovirus strains (SF162 and MW956.26)⁴⁰ and two pseudoviruses designed for resistance to T20 (pNL4.3 and pNL4.3 DIM)³⁴. As shown in Figure 7b and Table 2, covNHR-C did not show detectable inhibitory activity for any tested strain, whereas weak inhibition was obtained for covNHR-N.

Noteworthy, these experiments were carried out at 37°C since physiological temperature is necessary for fusion and infection between viruses and cells. Due to the low thermal stability of covNHR-N ($T_m \sim 41^\circ\text{C}$), the effective concentration of the functional protein is most likely affected at this temperature because of protein unfolding, resulting in overestimated IC_{50} values. This may be confirmed by the significantly higher inhibitory activities of the two disulfide-stabilized variants. Nevertheless, compared to the parent protein covNHR-VQ, containing the four NHR binding pockets, the inhibitory activities of the covNHR-N proteins are more than 10 fold diminished certainly because of their much lower affinity for their CHR target.

The inhibitory potency of T20 is relatively high for the SF162 and MW956.26 strains but much lower for the pNL4-3 strains, which harbor mutations at the NTP pocket conferring resistance to CHR-derived peptides. In contrast, the covNHR-N proteins showed greater inhibitory activity against the T20-resistant strains.

Discussion.

NHR coiled-coil is composed of two subdomains of different intrinsic stability

In this study, we have reported the subdivision of the covNHR-VQ protein into two well-defined miniproteins each representing about a half of the NHR coiled coil. Both proteins have similar molecular size, the same folding topology and are monomeric at physiological pH. However, covNHR-N is significantly less stable than covNHR-C. The complete covNHR-VQ protein already contains several stabilizing mutations, most of them at its N-terminal half³³, but the N-terminal miniprotein needed additional mutations engineering and a disulfide-bond to acquire significant conformational stability. In contrast, the covNHR-C miniprotein is highly stable without any further engineering. This indicates that the C-terminal half of the NHR helix trimer has a much greater intrinsic stability than the N-terminal region.

This is consistent with the crystal structure of a soluble, cleaved Env trimer construct (BG505 SOSIP.664)⁴¹, which has reported that the N-terminal half of NHR is partially disordered, while the C-terminal half is structured as a helix trimer within the Env. However, a more recent cryo-EM structure of a native JRFL

Env trimer with only the C-terminal tail deleted showed the first half of the NHR also in helical conformation ⁴². The authors attributed this difference to the I559P mutation presents in the SOSIP construct, which destabilizes the helix. Nonetheless, the first half of the NHR region is rather flexible and exposed in all prefusion structures and does not form trimer contacts, which are essential for the interaction with its complementary CHR region to acquire a stable 6HB coiled-coil structure. We observe in fact some reminiscency of this in the structural tightening of covNHR-N associated with binding of the Y24L peptide, as observed in the NMR spectra and manifested by a considerable negative binding enthalpy being partially compensated by negative binding entropy. Both magnitudes and the negative binding heat capacity (ΔC_p) are much greater than those observed for the binding of the same peptide to covNHR-VQ ³⁶.

Accessibility of the CHR region for interaction

Despite their different structural stability, both miniproteins are able to interact with their respective CHR peptide targets with similar micromolar-range dissociation constants. Therefore, these miniproteins constitute very convenient probes to test the vulnerability of each half of the CHR region of gp41 as a binding target to block fusion and inhibit HIV-1.

CovNHR-C binds its corresponding CHR peptide (N25S) with an affinity very similar to the parent molecule covNHR-VQ ³⁶. Previous thermodynamic studies performed with the stabilized trimeric NHR construct, IQN17, and short cyclic or D-peptides targetting the gp41 HP have reported K_d values in the same order of magnitude ^{43; 44}. The HP has been reported critical to the stabilization of 6-HB ¹⁰ and, therefore, has represented during decades the most important target for identification of small-molecule HIV-1 fusion and entry inhibitors ¹⁴. In spite that covNHR-C contains the HP and could bind its complementary CHR peptide in solution, it did not interact with any of the Env molecules tested in the ELISA experiments, suggesting an inaccessibility of its target sequence in the context of the prefusion spike.

High-resolution structures of different Env spikes in prefusion conformation have shown that the HP and CTP binding motifs are engaged in a tryptophan clasp involving the side chains of Trp623 (CTP motif), Trp628 and Trp631 (HP motif)

that locks the fusion peptide proximal region in its pre-fusion conformation^{41;42}. This interaction prevents therefore accessibility of covNHR-C binding to its target. In contrast, the three covNHR-N proteins could bind all soluble native Envs tested, although the level of binding measured was lower than that of the parent covNHR-VQ molecule containing the four NHR pockets, consistently with its lower affinity for the Y24L peptide. This indicates that its target region corresponding to the C-terminal part of the $\alpha 9$ helix in prefusion Envs is at least partially accessible. Also, this region is not glycosylated downstream of N651, in contrast to the first half of CHR, which is much more protected by glycans. Moreover, the CHR region near the MPER has been described as highly flexible^{45; 46; 47}. This suggests that covNHR-N proteins may be able to interact with the CHR region without the need of CD4 and co-receptor attachment. A similar observation has been made for 5-Helix^{48; 49}.

Miniproteins as potencial HIV-1 inhibitors

The covNHR-C protein was unable to inhibit any of the HIV-1 strains used in the assays, despite its high stability and the moderate affinity for its target CHR sequence, even though its target region may become transiently accessible during the fusion process. It is possible that micromolar affinity for its target is not sufficient to compensate for this transient exposure.

However, we have found that covNHR-N and its disulfide-stabilized mutants have similar affinities for Y24L as covNHR-C for N25S, and yet they show significant inhibitory activities. Remarkably, the two stable disulfide-bonded mutants reach IC₅₀ values of 30-40 nM for T20-resistant strains and moderate sub-micromolar IC₅₀ values for the other pseudoviruses, well below the measured dissociation constants for the Y24L peptide (low micromolar). This is in contrast with other CHR-targeting inhibitors such as 5-helix and our covNHR-VQ molecule, which show sub-picomolar K_d for CHR peptides but only low-nanomolar IC₅₀ values for virus inhibition.

The lack of correlation between target binding affinity and inhibitory potency in a series of 5-helix and CHR-peptide mutants has been explained by Root and coworkers according to a kinetic dependency of fusion inhibition^{50; 51}. According to their model, inhibitor binding to gp41 would only occur after CD4 activation in the fusion intermediate state, which transiently exposes the target epitope. For

tightly binding inhibitors, IC_{50} is limited to a minimum of k_f/k_{on} , i.e., the ratio between rates of target epitope occlusion and inhibitor association. On the other hand, for low affinity inhibitors, with much higher dissociation rates (k_{off}), binding equilibrium would be established during the window of target epitope exposure and then IC_{50} is correlated to the equilibrium dissociation constant K_D . This results in $IC_{50} \geq K_D$, as observed for 5-helix variants, covNHR-VQ and different CHR-derived peptides. However, for the covNHR-SS1 and covNHR-SS2 proteins, IC_{50} is much lower than K_D (about 5 μ M at 37°C, extrapolated from parameters of Table 1).

A possibility that explains the relatively high inhibitory efficiency of our miniproteins is that CHR accessibility in the fusion intermediate may be considerably higher than for larger molecules, such as 5-helix or covNHR-VQ in the context of the virus. This would give them a fundamental advantage as fusion inhibitors, since less restricted access to their target would provide a broader margin for affinity optimization to improve their inhibitory activity.

A second possibility is that the covNHR-N miniproteins could be able to block fusion at an earlier stage than the CD4-activated fusion intermediate. We have shown here that covNHR-N proteins containing the NTP and MP can bind native Envs, whereas covNHR-C protein cannot and this binding ability correlates with inhibitory activity. Binding to the flexible second half of CHR may perturb the four-helix collar of gp41 that grasps gp120 N- and C-termini, thus producing destabilization of the spike. It has been reported that MPER-targeting neutralizing antibodies, as well as antibodies targeting a conformational epitope mainly composed by HR2 (CHR) regions from two gp41 protomers, act by destabilizing the Env trimer and producing shedding^{52;53}. The possibility that this mechanism also applies to the covNHR-N miniproteins deserves further investigation.

Therefore, there is considerable margin for improvement of the covNHR-N miniproteins by engineering new modifications increasing their stability, as well as by improving the modest binding affinity for their target adding, for instance, new motifs targeting the nearby MPER region.

In conclusion, we have designed, produced and characterized two single chain covNHR miniproteins each mimicking one of the halves of the gp41 NHR region.

These miniproteins fold autonomously and represent subdomains of NHR, with very different intrinsic stability. Although both covNHR proteins can bind their respective complementary CHR peptides with similar affinity, the C-terminal protein cannot bind its target in soluble prefusion Env spikes and does not show any HIV-1 inhibitory activity in vitro. In contrast, the N-terminal proteins show significant capacity to bind the Env spikes and are capable to inhibit HIV-1 with a relatively high potency, suggesting that their small size confers them a particular advantage to access their target in the virion context. These results show a special vulnerability of the flexible C-terminal end of the gp41 CHR region to novel fusion inhibitors.

Materials and Methods.

Protein and peptide samples

The NHR and CHR gp41 sequences used in this work are described in Figure 1. The reference gp41 sequence was taken from the full gp160 precursor glycoprotein of the HIV-1 BRU isolate (Swiss-Prot entry sp|P03377|ENV_HV1BR). The covNHR miniproteins were computationally designed using SwissPDBviewer⁵⁴ and YASARA software⁵⁵. The DNA encoding the protein sequences were synthesized and cloned into pET303 expression vectors (ThermoFisher Scientific, Waltham, USA). To facilitate purification by Ni-Sepharose affinity chromatography, the protein sequences were histidine tagged at the C-terminus with the sequence GGGGSHHHHHH. The covNHR proteins were produced and purified following the protocol previously described³³. Synthetic CHR peptides, both N-acetylated and C-amidated, were acquired from Genecust (Luxembourg), with a purity > 95%. Protein and peptide concentrations were determined by UV absorption measurements at 280 nm using the extinction coefficients calculated according to their respective amino acid sequences with the ExPasy ProtParam server (<https://web.expasy.org/protparam/>)⁵⁶.

Circular dichroism

CD spectra were recorded in a Jasco J-715 spectropolarimeter (Jasco, Tokyo, Japan) equipped with Peltier-thermostatic cell holder. Measurements of the far-UV

CD spectra (260–200 nm) were made with a 1 mm path length quartz cuvette at a protein concentration of ~20 μM . Spectra were recorded at a scan rate of 100 nm/min, 1 nm step resolution, 1 s response, and 1 nm bandwidth. The resulting spectra were usually the average of five scans and the percentage of α -helical structure was estimated from the far-UV CD spectra as described elsewhere⁵⁷. In thermal melting experiments, the CD signal was monitored as a function of temperature at 222 nm.

Near-UV CD spectra (350–250 nm) were measured at a protein concentration of ~15 μM using a 5 mm cuvette and were typically the average of 20 scans. Each spectrum was corrected by baseline subtraction using the blank spectrum obtained with the buffer and finally the CD signal was normalized to molar ellipticity ($[\theta]$, in $\text{deg}\cdot\text{dmol}^{-1}\cdot\text{cm}^2$). The interaction experiments with CHR peptides were carried at a 1:2 molar ratio between the proteins and the corresponding peptide.

Dynamic light scattering

The particle sizes of the covNHR proteins were assessed by DLS measurements using a DynaPro MS-X instrument (Wyatt, Santa Barbara, CA). Dynamics software (Wyatt Technology Corporation, Santa Barbara, CA, USA) was used in data collection and processing. Sets of DLS data were measured at 25°C with an average number of 50 acquisitions and an acquisition time of 10 seconds.

Limited proteolysis

Proteolysis experiments were carried out using thermolysin as protease (Sigma, St Louis, MO) in 50 mM HEPES, 100 mM NaCl, 5 mM CaCl_2 pH 7.2. Aliquots of the covNHR proteins at 30 μM concentration, previously dialyzed in the buffer and equilibrated at the desired temperature, were mixed with thermolysin at different protein:protease ratios and incubated in a thermoblock at 20°C. The reaction was stopped at different times by adding EDTA to reach 20 mM concentration and the sample aliquots were immediately frozen. The samples were analyzed by SDS-PAGE in Tris/Tricine buffers and by HPLC coupled to ESI-MS in a WATERS LCT Premier XE instrument equipped with a time-of-flight analyzer. Fragments were identified using the ExPasy Findpept tool (<https://web.expasy.org/findpept/>).

Isothermal titration calorimetry

ITC measurements were carried out in a Microcal VP-ITC calorimeter (Malvern Instruments, Worcestershire, UK). The protein solutions were titrated with 25 injections of 5 μL of the peptide solution at 480 sec intervals. Protein concentration in the cell was around 20 μM , while the ligands in the syringe were typically at 200-300 μM . The experiments were carried out in 50 mM sodium phosphate buffer, pH 7.4. As a blank, an independent experiment with only buffer in the calorimeter's cell was performed with the same peptide solution to determine the corresponding heats of dilution. The experimental thermograms were baseline corrected and the peaks were integrated to determine the heats produced by each ligand injection. Finally, each heat was normalized per mole of added ligand. The resulting binding isotherms were fitted using a binding model of identical and independent sites, allowing the determination of the binding constant, K_b , the binding enthalpy, ΔH_b , and the binding stoichiometry, n , for each interaction. From these values, the Gibbs energy and entropy of binding could be derived as $\Delta G_b = -RT \cdot \ln K_b$ and $T \cdot \Delta S_b = \Delta H_b - \Delta G_b$. Binding heat capacities were determined from the slope of the dependences of the binding enthalpies measured at different temperatures.

NMR spectroscopy

NMR experiments were performed on a Varian Direct Drive 600 MHz spectrometer (Varian, California, USA). The protein samples were labelled with ^{15}N following a protocol established elsewhere⁵⁸. The protein was dissolved at 0.8 mM in 50 mM sodium phosphate pH 7.4, containing 10% D_2O . To prepare a protein-peptide mixture, the lyophilized peptide was dissolved in H_2O at a concentration of 1.6 mM and the pH was adjusted to pH 7.4. Then this peptide solution was further lyophilized. Subsequently, the newly lyophilized peptide was dissolved with the protein solution to reach a final concentration of 1.6 mM. ^1H - ^{15}N 2D-HSQC spectra were acquired with 2K points and 32 increments and processed with NMRPipe⁵⁹.

Binding to HIV-1 envelope spikes

The capacity of the covNHR proteins to bind soluble HIV-1 envelope proteins (Env) was determined by ELISA. Briefly, 96-well ELISA plates (Maxisorp, Nunc) were coated at 4 °C overnight with various Envs (Table S2) in 0.1 M bicarbonate buffer (pH 9.6). After saturation with 2% BSA, 0.05% Tween in PBS for 1.5 h at 25°C, 0.01 µM of covNHR molecules (100 µL diluted in 1% BSA 0.05% Tween solution) were added and incubated for 2 h at room temperature. The plate was then washed five times and covNHR binding was detected with 100µL anti-6X Histag antibody conjugated to horseradish peroxidase (HRP) (Abcam) at 1/10000 dilution incubated for 1 hour at room temperature. Antibody binding was then revealed with tetramethylbenzidine (TMB) substrate buffer, the reaction was stopped with 1M H₂SO₄ and optical density was read at 450 nm with a Molecular Device Plate Reader equipped with SoftMax Pro 6 program. Background binding was measured in plates without Env and subtracted from the data. The percentage of binding was calculated using the readings with wells coated with His-tagged Env incubated with PBS buffer instead of covNHR molecules as control 100% binding.

HIV-1 inhibitory assays

The inhibition of HIV replication was determined using the conventional TZM-bl assay measured as a function of reductions in Tat-regulated Firefly luciferase (Luc) reporter gene expression⁶⁰. Pseudoviruses expressing different Env were tested for HIV inhibitory potential⁴⁰. The IC₅₀, the concentration (in nM) of inhibitor inducing a 50% decrease in relative luminometer units (RLU), corresponding to a 50% decrease in virus replication was calculated by non-linear regression using a sigmoidal Hill function, as implemented in Origin software (Originlab, Northampton, MA).

Acknowledgements.

This work has been funded by the Spanish Ministry of Economy and Competitiveness (grant: BIO2016-76640-R), ANRS and the Vaccine Research Institute for the Investissements d'Avenir program to CM and by the European Fund for Research and Development from the European Union. SJ acknowledges a short-term fellowship from Boehringer-Ingelheim Fonds. MCM was a recipient of a

contract from the Program for Warranty of Young Employment of the European Social Fund. VR and SS acknowledge ERASMUS+ exchange grants from the European Union.

Figure Legends.

Figure 1. Design of covNHR miniproteins. (a) Crystallographic structure of covNHR-VQ in complex with C34 (pdb: 6R2G). CovNHR-VQ is represented in golden color and the C34 peptide in blue and red. The subdomains corresponding to covNHR-N and covNHR-C are respectively underlined by the green and magenta dashed boxes. Models of covNHR-C (magenta) and covNHR-N (green) are shown. (b) Sequence alignment of the covNHR proteins. Residues in bold indicate mutations engineered in the original design of covNHR-VQ. Mutations in red were engineered in this work. (c) Sequences of the CHR peptides used for binding experiments in this study. Residues corresponding to the different binding pocket motifs are colored as follows: CTP in orange, HP in cyan, MP in blue and NTP in red.

Figure 2. Secondary structure and thermal stability of covNHR-N and covNHR-C. (a) Far UV CD spectra of covNHR-N (black lines) and covNHR-C (grey lines) at pH 2.5 (dashed lines) and pH 7.4 (solid lines). (b) Thermal unfolding of covNHR-N (black line) and covNHR-C (grey line) followed by monitoring the CD signal at 222 nm.

Figure 3. Engineering a disulfide bond into covNHR-N protein. (a) Model of covNHR-NSS1 (G32C/T92C) showing the location of the residues chosen for mutations. (b) Far UV CD spectra of covNHR-NSS1 (black lines) and covNHR-SS2 (grey lines) in the presence (dashed lines) and absence (solid lines) of 5 mM TCEP. (c) Comparison of the thermal unfolding of reduced (black circle symbols) and oxidized (black square symbols) covNHR-NSS2 with covNHR-N (grey symbols) followed by monitoring the CD signal at 222 nm. Symbols represent the experimental data and the solid lines correspond to the best fitting carried out using a two states unfolding model.

Figure 4. Interaction between covNHR proteins with the corresponding peptides monitored by far-UV and near-UV CD. (a) Far UV CD spectra of free covNHR proteins and 1:2 mixtures of the proteins with the corresponding CHR peptide. Spectra are normalized per mole and represented as follow: 1:2 mixtures (solid lines); free covNHR variants (dotted lines); free peptide (dashed lines); sum of the spectra of the free molecules (dashed-dotted lines). Protein concentrations used were 20 μ M. (b) Near UV CD spectra of isolated covNHR-C and in complex with N25S at a ratio 1:2.

Figure 5. Isothermal titration calorimetry experiments of CHR binding. Thermograms and ITC binding isotherms of Y24L binding to covNHR-N (a) and of N25S binding to covNHR-C (b) at 20°C. The symbols correspond to the experimental heats and the lines represent the fittings using a binding model of n identical and independent sites. (c-d). Comparison of the thermodynamic binding parameters of Y24L (c) and N25S (d) to the covNHR proteins at 25°C. The values have been calculated from the parameters taken from Table 1.

Figure 6. ^1H - ^{15}N HSQC spectra of covNHR-N free (a) and bound with an excess of Y24L at a ratio 1:2 (b).

Figure 7. (a) Binding of covNHR proteins to different soluble Envs measured by ELISA. CovNHR binding to Env was detected using anti-Histag Ab as primary antibody. Background binding was measured without Env and subtracted from the data. 100% positive control was measured with wells directly coated with a His-tagged Env. Data correspond to mean \pm S.D. values of three independent measurements. (b) HIV-1 inhibitory activity of covNHR miniproteins and T20 on a T20 resistant variant. In vitro inhibition of pNL4-3 DIM pseudovirus infection of TZM-bL cells by fusion inhibitors added at different concentrations. Data are the mean \pm S.D. of three independent measurements. Regression curves were plotted using a non-linear regression using a sigmoidal Hill function, as implemented in Origin software (Originlab, Northampton, MA).

References.

1. Johnson, L. M., Horne, W. S. & Gellman, S. H. (2011). Broad distribution of energetically important contacts across an extended protein interface. *J Am Chem Soc* **133**, 10038-41.
2. Tseng, A. & Foisy, M. (2012). Important Drug-Drug Interactions in HIV-Infected Persons on Antiretroviral Therapy: An Update on New Interactions Between HIV and Non-HIV Drugs. *Curr Infect Dis Rep* **14**, 67-82.
3. Chan, D. C. & Kim, P. S. (1998). HIV entry and its inhibition. *Cell* **93**, 681-4.
4. Eckert, D. M. & Kim, P. S. (2001). Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region. *Proc Natl Acad Sci U S A* **98**, 11187-92.
5. White, T. A., Bartesaghi, A., Borgnia, M. J., Meyerson, J. R., de la Cruz, M. J., Bess, J. W., Nandwani, R., Hoxie, J. A., Lifson, J. D., Milne, J. L. & Subramaniam, S. (2010). Molecular architectures of trimeric SIV and HIV-1 envelope glycoproteins on intact viruses: strain-dependent variation in quaternary structure. *PLoS Pathog* **6**, e1001249.
6. Ward, A. B. & Wilson, I. A. (2017). The HIV-1 envelope glycoprotein structure: nailing down a moving target. *Immunol Rev* **275**, 21-32.
7. Munro, J. B. & Mothes, W. (2015). Structure and Dynamics of the Native HIV-1 Env Trimer. *J Virol* **89**, 5752-5.
8. Eckert, D. M. & Kim, P. S. (2001). Mechanisms of viral membrane fusion and its inhibition. *Annu Rev Biochem* **70**, 777-810.
9. Melikyan, G. B. (2011). Membrane fusion mediated by human immunodeficiency virus envelope glycoprotein. *Curr Top Membr* **68**, 81-106.
10. Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**, 263-73.
11. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. (1997). Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387**, 426-30.
12. Yi, H. A., Fochtman, B. C., Rizzo, R. C. & Jacobs, A. (2016). Inhibition of HIV Entry by Targeting the Envelope Transmembrane Subunit gp41. *Curr HIV Res* **14**, 283-94.

13. Kilby, J. M., Hopkins, S., Venetta, T. M., DiMassimo, B., Cloud, G. A., Lee, J. Y., Alldredge, L., Hunter, E., Lambert, D., Bolognesi, D., Matthews, T., Johnson, M. R., Nowak, M. A., Shaw, G. M. & Saag, M. S. (1998). Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. *Nat Med* **4**, 1302-7.
14. Chan, D. C., Chutkowski, C. T. & Kim, P. S. (1998). Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. *Proc Natl Acad Sci U S A* **95**, 15613-7.
15. Eron, J. J., Gulick, R. M., Bartlett, J. A., Merigan, T., Arduino, R., Kilby, J. M., Yangco, B., Diers, A., Drobnes, C., DeMasi, R., Greenberg, M., Melby, T., Raskino, C., Rusnak, P., Zhang, Y., Spence, R. & Miralles, G. D. (2004). Short-term safety and antiretroviral activity of T-1249, a second-generation fusion inhibitor of HIV. *J Infect Dis* **189**, 1075-83.
16. He, Y., Cheng, J., Lu, H., Li, J., Hu, J., Qi, Z., Liu, Z., Jiang, S. & Dai, Q. (2008). Potent HIV fusion inhibitors against Enfuvirtide-resistant HIV-1 strains. *Proc Natl Acad Sci U S A* **105**, 16332-7.
17. Welch, B. D., VanDemark, A. P., Heroux, A., Hill, C. P. & Kay, M. S. (2007). Potent D-peptide inhibitors of HIV-1 entry. *Proc Natl Acad Sci U S A* **104**, 16828-33.
18. Frey, G., Rits-Volloch, S., Zhang, X.-Q., Schooley, R. T., Chen, B. & Harrison, S. C. (2006). Small molecules that bind the inner core of gp41 and inhibit HIV envelope-mediated fusion. *Proc Natl Acad Sci U S A* **103**, 13938-13943.
19. Lu, L., Yu, F., Cai, L., Debnath, A. K. & Jiang, S. (2016). Development of Small-molecule HIV Entry Inhibitors Specifically Targeting gp120 or gp41. *Curr Top Med Chem* **16**, 1074-90.
20. Corti, D., Langedijk, J. P., Hinz, A., Seaman, M. S., Vanzetta, F., Fernandez-Rodriguez, B. M., Silacci, C., Pinna, D., Jarrossay, D., Balla-Jhagihorsingh, S., Willems, B., Zekveld, M. J., Dreja, H., O'Sullivan, E., Pade, C., Orkin, C., Jeffs, S. A., Montefiori, D. C., Davis, D., Weissenhorn, W., McKnight, A., Heeney, J. L., Sallusto, F., Sattentau, Q. J., Weiss, R. A. & Lanzavecchia, A. (2010). Analysis of memory B cell responses and isolation of novel monoclonal antibodies with neutralizing breadth from HIV-1-infected individuals. *PLoS One* **5**, e8805.

21. Miller, M. D., Geleziunas, R., Bianchi, E., Lennard, S., Hrin, R., Zhang, H., Lu, M., An, Z., Ingallinella, P., Finotto, M., Mattu, M., Finnefrock, A. C., Bramhill, D., Cook, J., Eckert, D. M., Hampton, R., Patel, M., Jarantow, S., Joyce, J., Ciliberto, G., Cortese, R., Lu, P., Strohl, W., Schleif, W., McElhaugh, M., Lane, S., Lloyd, C., Lowe, D., Osbourn, J., Vaughan, T., Emini, E., Barbato, G., Kim, P. S., Hazuda, D. J., Shiver, J. W. & Pessi, A. (2005). A human monoclonal antibody neutralizes diverse HIV-1 isolates by binding a critical gp41 epitope. *Proc Natl Acad Sci U S A* **102**, 14759-64.
22. Wild, C., Oas, T., McDanal, C., Bolognesi, D. & Matthews, T. (1992). A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition. *Proc Natl Acad Sci U S A* **89**, 10537-41.
23. Poveda, E., Briz, V. & Soriano, V. (2005). Enfuvirtide, the first fusion inhibitor to treat HIV infection. *AIDS Rev* **7**, 139-47.
24. He, Y. (2013). Synthesized peptide inhibitors of HIV-1 gp41-dependent membrane fusion. *Curr Pharm Des* **19**, 1800-9.
25. Steffen, I. & Pohlmann, S. (2010). Peptide-based inhibitors of the HIV envelope protein and other class I viral fusion proteins. *Curr Pharm Des* **16**, 1143-58.
26. He, Y., Xiao, Y., Song, H., Liang, Q., Ju, D., Chen, X., Lu, H., Jing, W., Jiang, S. & Zhang, L. (2008). Design and evaluation of sifuvirtide, a novel HIV-1 fusion inhibitor. *J Biol Chem* **283**, 11126-34.
27. Dwyer, J. J., Wilson, K. L., Davison, D. K., Freel, S. A., Seedorff, J. E., Wring, S. A., Tvermoes, N. A., Matthews, T. J., Greenberg, M. L. & Delmedico, M. K. (2007). Design of helical, oligomeric HIV-1 fusion inhibitor peptides with potent activity against enfuvirtide-resistant virus. *Proc Natl Acad Sci U S A* **104**, 12772-7.
28. Root, M. J., Kay, M. S. & Kim, P. S. (2001). Protein design of an HIV-1 entry inhibitor. *Science* **291**, 884-8.
29. Louis, J. M., Bewley, C. A. & Clore, G. M. (2001). Design and properties of N(CCG)-gp41, a chimeric gp41 molecule with nanomolar HIV fusion inhibitory activity. *J Biol Chem* **276**, 29485-9.

30. Louis, J. M., Nesheiwat, I., Chang, L., Clore, G. M. & Bewley, C. A. (2003). Covalent trimers of the internal N-terminal trimeric coiled-coil of gp41 and antibodies directed against them are potent inhibitors of HIV envelope-mediated cell fusion. *J Biol Chem* **278**, 20278-85.
31. Bianchi, E., Joyce, J. G., Miller, M. D., Finnefrock, A. C., Liang, X., Finotto, M., Ingallinella, P., McKenna, P., Citron, M., Ottinger, E., Hepler, R. W., Hrin, R., Nahas, D., Wu, C., Montefiori, D., Shiver, J. W., Pessi, A. & Kim, P. S. (2010). Vaccination with peptide mimetics of the gp41 prehairpin fusion intermediate yields neutralizing antisera against HIV-1 isolates. *Proc Natl Acad Sci U S A* **107**, 10655-60.
32. Chen, X., Lu, L., Qi, Z., Lu, H., Wang, J., Yu, X., Chen, Y. & Jiang, S. (2010). Novel recombinant engineered gp41 N-terminal heptad repeat trimers and their potential as anti-HIV-1 therapeutics or microbicides. *J Biol Chem* **285**, 25506-15.
33. Crespillo, S., Camara-Artigas, A., Casares, S., Morel, B., Cobos, E. S., Mateo, P. L., Mouz, N., Martin, C. E., Roger, M. G., El Habib, R., Su, B., Moog, C. & Conejero-Lara, F. (2014). Single-chain protein mimetics of the N-terminal heptad-repeat region of gp41 with potential as anti-HIV-1 drugs. *Proc Natl Acad Sci U S A* **111**, 18207-12.
34. Jurado, S., Cano-Munoz, M., Morel, B., Standoli, S., Santarossa, E., Moog, C., Schmidt, S., Laumond, G., Camara-Artigas, A. & Conejero-Lara, F. (2019). Structural and Thermodynamic Analysis of HIV-1 Fusion Inhibition Using Small gp41 Mimetic Proteins. *J Mol Biol* **431**, 3091-3106.
35. He, Y., Cheng, J., Li, J., Qi, Z., Lu, H., Dong, M., Jiang, S. & Dai, Q. (2008). Identification of a critical motif for the human immunodeficiency virus type 1 (HIV-1) gp41 core structure: implications for designing novel anti-HIV fusion inhibitors. *J Virol* **82**, 6349-58.
36. Jurado, S., Cano-Munoz, M., Polo-Megias, D., Conejero-Lara, F. & Morel, B. (2020). Thermodynamic dissection of the interface between HIV-1 gp41 heptad repeats reveals cooperative interactions and allosteric effects. *Arch Biochem Biophys* **688**, 108401.

37. Conchillo-Sole, O., de Groot, N., Aviles, F., Vendrell, J., Daura, X. & Ventura, S. (2007). AGGRESKAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics* **8**, 65.
38. Craig, D. B. & Dombkowski, A. A. (2013). Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* **14**, 346.
39. Peisajovich, S. G., Blank, L., Epand, R. F., Epand, R. M. & Shai, Y. (2003). On the interaction between gp41 and membranes: the immunodominant loop stabilizes gp41 helical hairpin conformation. *J Mol Biol* **326**, 1489-501.
40. deCamp, A., Hraber, P., Bailer, R. T., Seaman, M. S., Ochsenbauer, C., Kappes, J., Gottardo, R., Edlefsen, P., Self, S., Tang, H., Greene, K., Gao, H., Daniell, X., Sarzotti-Kelsoe, M., Gorny, M. K., Zolla-Pazner, S., LaBranche, C. C., Mascola, J. R., Korber, B. T. & Montefiori, D. C. (2014). Global panel of HIV-1 Env reference strains for standardized assessments of vaccine-elicited neutralizing antibodies. *J Virol* **88**, 2489-507.
41. Pancera, M., Zhou, T., Druz, A., Georgiev, I. S., Soto, C., Gorman, J., Huang, J., Acharya, P., Chuang, G. Y., Ofek, G., Stewart-Jones, G. B., Stuckey, J., Bailer, R. T., Joyce, M. G., Louder, M. K., Tumba, N., Yang, Y., Zhang, B., Cohen, M. S., Haynes, B. F., Mascola, J. R., Morris, L., Munro, J. B., Blanchard, S. C., Mothes, W., Connors, M. & Kwong, P. D. (2014). Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* **514**, 455-61.
42. Lee, J. H., Ozorowski, G. & Ward, A. B. (2016). Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. *Science* **351**, 1043-8.
43. Cole, J. L. & Garsky, V. M. (2001). Thermodynamics of peptide inhibitor binding to HIV-1 gp41. *Biochemistry* **40**, 5633-41.
44. Sia, S. K., Carr, P. A., Cochran, A. G., Malashkevich, V. N. & Kim, P. S. (2002). Short constrained peptides that inhibit HIV-1 entry. *Proc Natl Acad Sci USA* **99**, 14664-14669.
45. Guttman, M., Garcia, N. K., Cupo, A., Matsui, T., Julien, J. P., Sanders, R. W., Wilson, I. A., Moore, J. P. & Lee, K. K. (2014). CD4-induced activation in a soluble HIV-1 Env trimer. *Structure* **22**, 974-84.
46. Stadtmueller, B. M., Bridges, M. D., Dam, K. M., Lerch, M. T., Huey-Tubman, K. E., Hubbell, W. L. & Bjorkman, P. J. (2018). DEER Spectroscopy Measurements Reveal Multiple Conformations of HIV-1 SOSIP Envelopes

- that Show Similarities with Envelopes on Native Virions. *Immunity* **49**, 235-246 e4.
47. Pan, J., Peng, H., Chen, B. & Harrison, S. C. (2020). Cryo-EM Structure of Full-length HIV-1 Env Bound With the Fab of Antibody PG16. *J Mol Biol* **432**, 1158-1168.
 48. Koshiya, T. & Chan, D. C. (2003). The prefusogenic intermediate of HIV-1 gp41 contains exposed C-peptide regions. *J Biol Chem* **278**, 7573-9.
 49. Root, M. J. & Hamer, D. H. (2003). Targeting therapeutics to an exposed and conserved binding element of the HIV-1 fusion protein. *Proc Natl Acad Sci U S A* **100**, 5016-21.
 50. Kahle, K. M., Steger, H. K. & Root, M. J. (2009). Asymmetric deactivation of HIV-1 gp41 following fusion inhibitor binding. *PLoS Pathog* **5**, e1000674.
 51. Steger, H. K. & Root, M. J. (2006). Kinetic dependence to HIV-1 entry inhibition. *J Biol Chem* **281**, 25813-21.
 52. Ruprecht, C. R., Krarup, A., Reynell, L., Mann, A. M., Brandenberg, O. F., Berlinger, L., Abela, I. A., Regoes, R. R., Gunthard, H. F., Rusert, P. & Trkola, A. (2011). MPER-specific antibodies induce gp120 shedding and irreversibly neutralize HIV-1. *J Exp Med* **208**, 439-54.
 53. Lee, J. H., Leaman, D. P., Kim, A. S., Torrents de la Pena, A., Sliepen, K., Yasmeen, A., Derking, R., Ramos, A., de Taeye, S. W., Ozorowski, G., Klein, F., Burton, D. R., Nussenzweig, M. C., Pognard, P., Moore, J. P., Klasse, P. J., Sanders, R. W., Zwick, M. B., Wilson, I. A. & Ward, A. B. (2015). Antibodies to a conformational epitope on gp41 neutralize HIV-1 by destabilizing the Env spike. *Nat Commun* **6**, 8167.
 54. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-23.
 55. Krieger, E. & Vriend, G. (2014). YASARA View - molecular graphics for all devices - from smartphones to workstations. *Bioinformatics* **30**, 2981-2.
 56. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook* (Walker, J. M., ed.), pp. 571-607. Humana Press, Totowa, NJ.

57. Luo, P. & Baldwin, R. L. (1997). Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **36**, 8413-21.
58. Weber, D. J., Gittis, A. G., Mullen, G. P., Abeygunawardana, C., Lattman, E. E. & Mildvan, A. S. (1992). NMR docking of a substrate into the X-ray structure of staphylococcal nuclease. *Proteins* **13**, 275-87.
59. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-93.
60. Sarzotti-Kelsoe, M., Bailer, R. T., Turk, E., Lin, C. L., Bilska, M., Greene, K. M., Gao, H., Todd, C. A., Ozaki, D. A., Seaman, M. S., Mascola, J. R. & Montefiori, D. C. (2014). Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *J Immunol Methods* **409**, 131-46.

Conceptualization : SJ, CM, FCL, BM

Methodology : FCL

Formal analysis : SJ, MCM, FCL, BM

Investigation : SJ, CM, MCM, SS, GL, VR, SS, DPM, FCL, BM

Resources : FCL, CM

Data Curation : SJ, CM, FCL, BM

Writing - Original Draft : FCL, BM

Writing - Review & Editing : SJ, CM, MCM, VR, SS, DPM, FCL, BM

Supervision : FCL, BM

Project administration : FCL, BM

Funding acquisition : CM, FCL

Table 1. Thermodynamic parameters of binding of gp41 CHR peptides to covNHR proteins measured by ITC.

Protein	Peptide	Temp. (°C)	K_d (μM)	ΔH_b ($\text{kJ}\cdot\text{mol}^{-1}$)	n	ΔC_{pb} ($\text{kJ}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$)
covNHR-VQ ^a	Y24L	25	0.090 ± 0.007	-63 ± 1	0.83	-1.6 ± 0.1
	N25S	25	3 ± 1	-34 ± 2	1	-1.5 ± 0.2
covNHR-N	Y24L	10	0.23 ± 0.02	-46 ± 0.4	0.86	-2.8 ± 0.2
		15	0.29 ± 0.02	-61 ± 0.6	0.84	
		20	0.44 ± 0.02	-76 ± 1.4	0.8	
		25	0.79 ± 0.02	-88 ± 0.5	0.91	
covNHR-NSS1	Y24L	10	0.27 ± 0.02	-47.6 ± 0.5	1.04	-2.0 ± 0.4
		15	0.37 ± 0.02	-57.2 ± 0.4	0.83	
		20	0.5 ± 0.7	-66.0 ± 1.2	1.04	
		25	1.1 ± 0.2	-79.6 ± 2.1	1.1	
covNHR-NSS2	Y24L	10	0.21 ± 0.01	-48.3 ± 0.3	0.99	-2.1 ± 0.3
		15	0.29 ± 0.02	-59.5 ± 0.5	0.94	
		20	0.51 ± 0.02	-69.5 ± 0.4	1.16	
		25	1.1 ± 0.3	-85 ± 4	1.0	
covNHR-C	N25S	10	0.24 ± 0.4	-28 ± 0.8	0.84	-1.5 ± 0.2
		15	0.31 ± 0.06	-32 ± 1	0.8	
		20	0.48 ± 0.09	-42 ± 1.4	1.03	
		25	1.9 ± 0.3	-49 ± 2.6	0.94	

Errors correspond to 95% confidence intervals of the fittings.

a. Data taken from Jurado *et al.*, 2020³⁶.

Table 2. In vitro HIV-1 inhibition by covNHR proteins and T20. Inhibitory activity (IC_{50} nM \pm S.D. of triplicates) was measured with the standard TZM-bl assay using different pseudoviruses.

Pseudovirus	covNHR-N	covNHR-NSS1	covNHR-NSS2	covNHR-C	covNHR-VQ ^b	T20 ^b
<i>pNL4-3 XCS</i>	740 \pm 100	37 \pm 6	39 \pm 4	-	2.0 \pm 0.3	182 \pm 25
<i>pNL4-3 (DIM)^a</i>	284 \pm 15	32 \pm 7	37 \pm 3	-	3.1 \pm 0.3	490 \pm 260
<i>SF162</i>	-	162 \pm 14	126 \pm 26	-	8 \pm 2	51 \pm 1
<i>MW965.26</i>	740 \pm 260	188 \pm 36	220 \pm 9	-	2 \pm 1	4 \pm 1

a. T20-resistant strain

b. Data taken from Jurado *et al.*, 2019³⁴

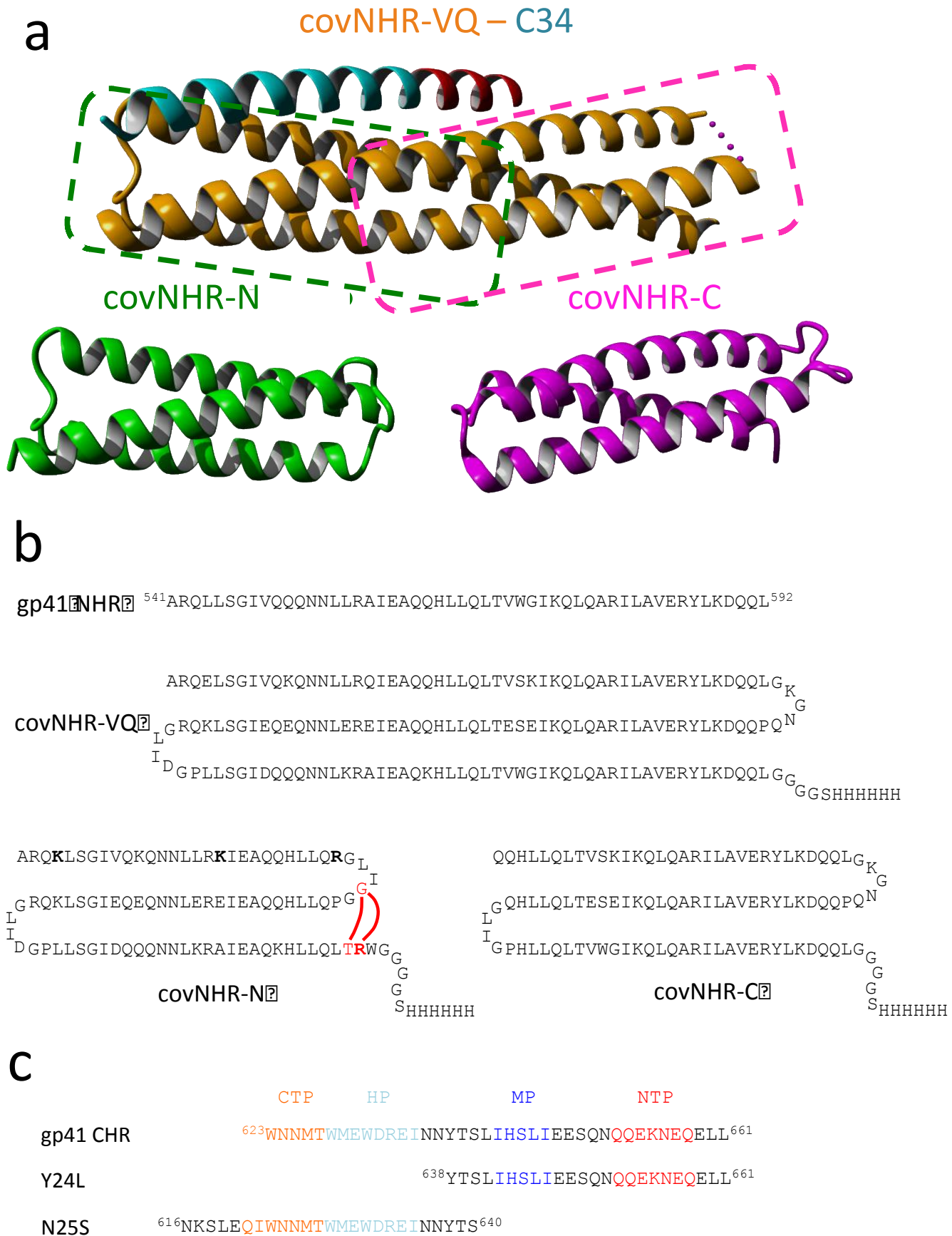
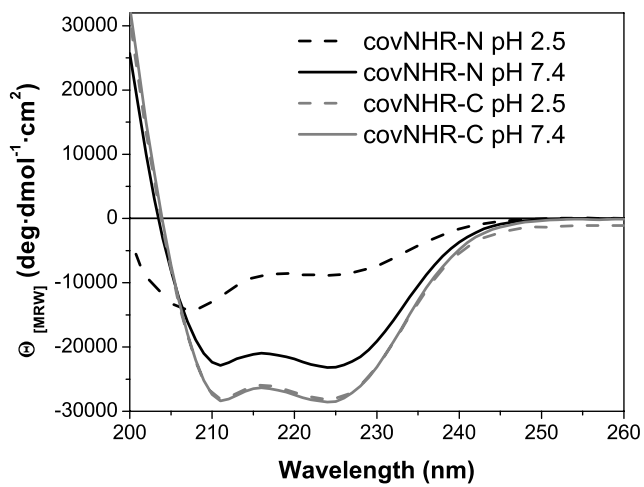


Figure 2

a



b

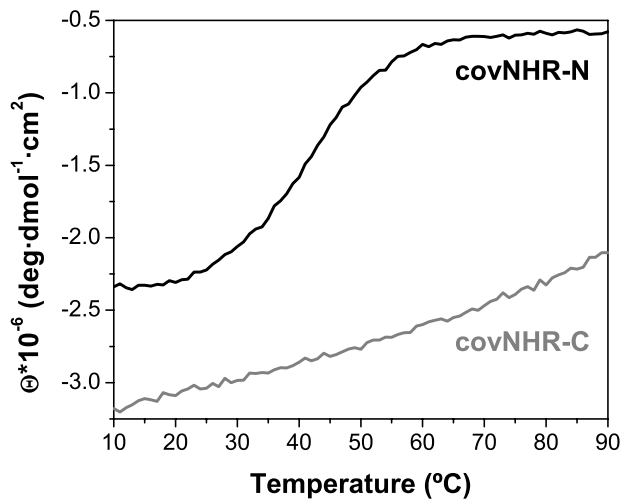


Figure 3

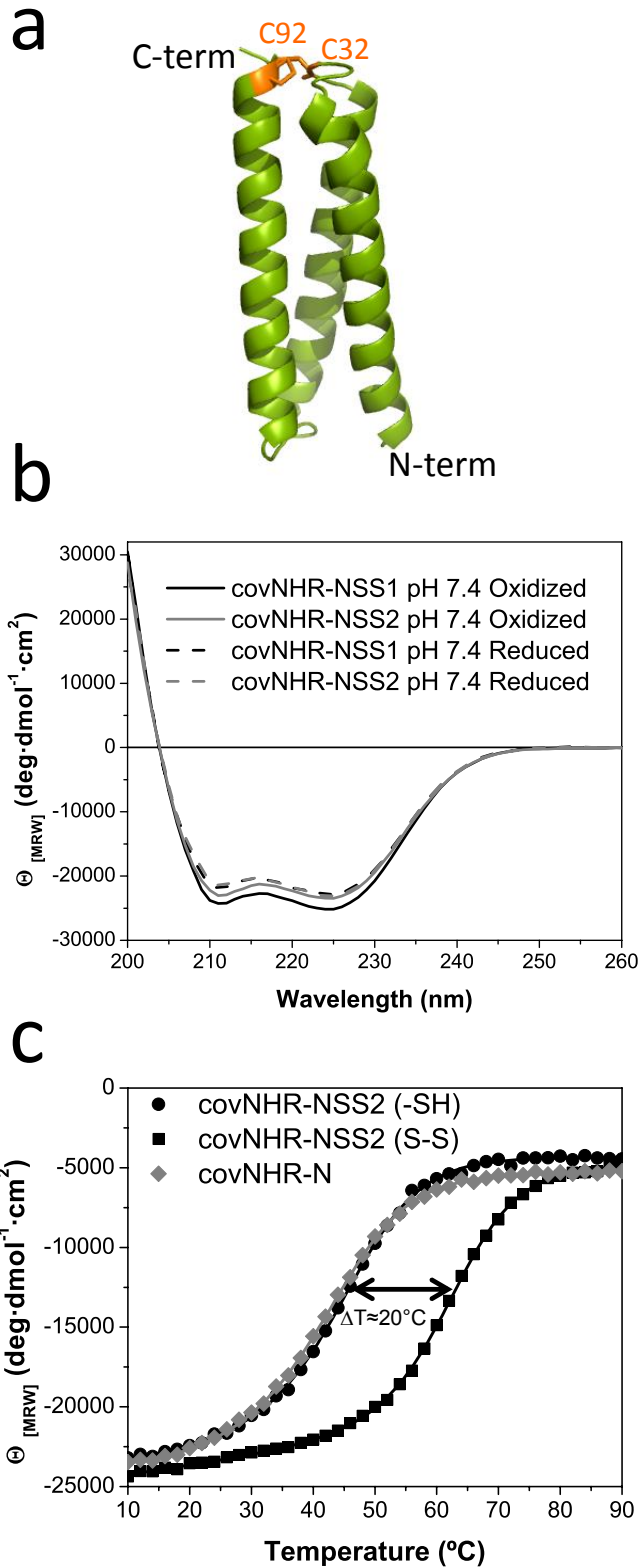
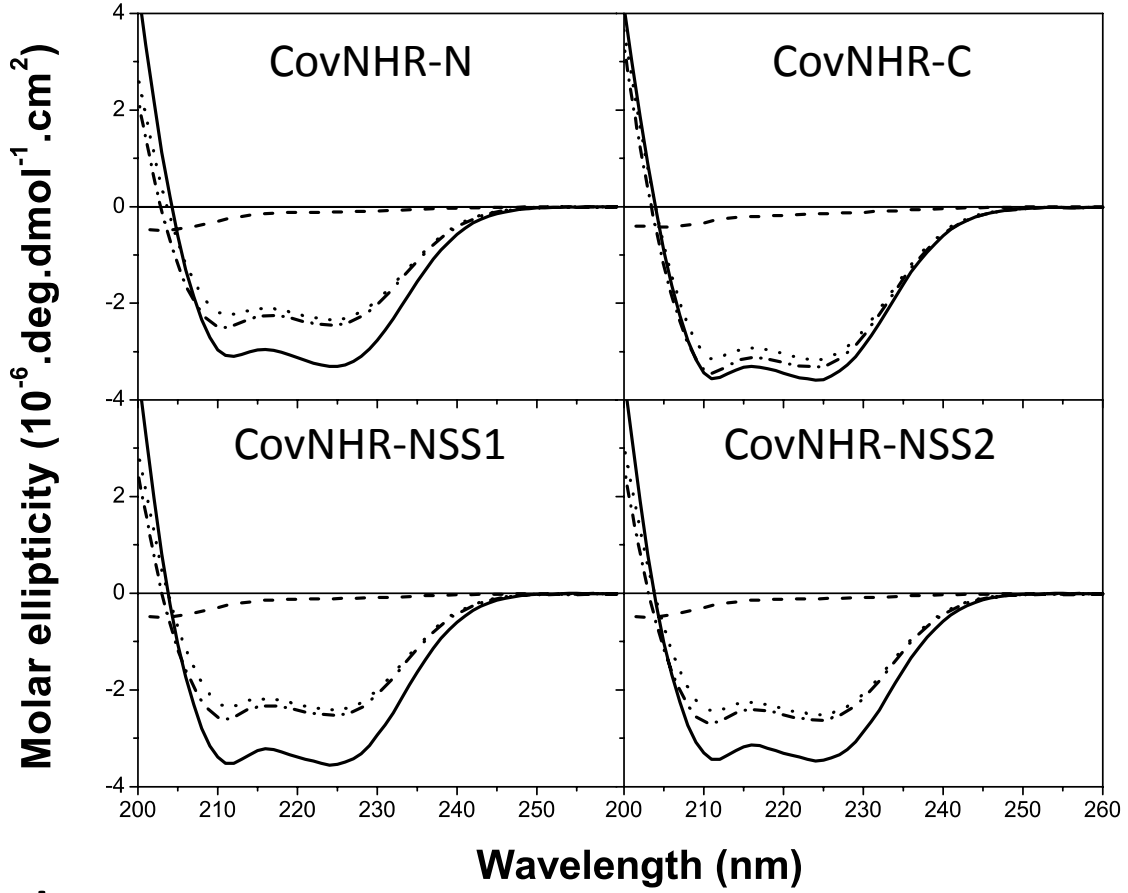


Figure 4

a



b

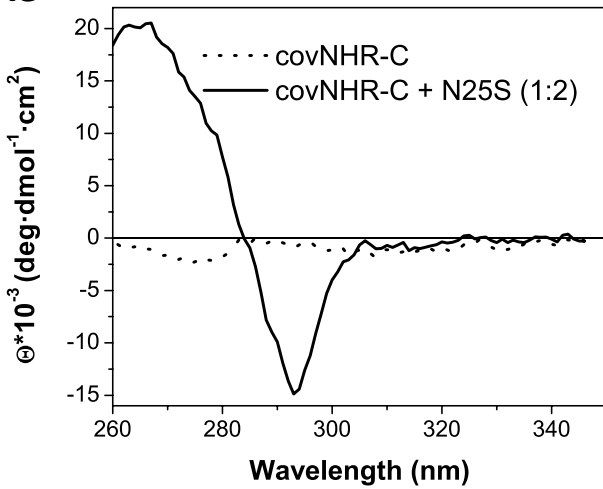


Figure 5

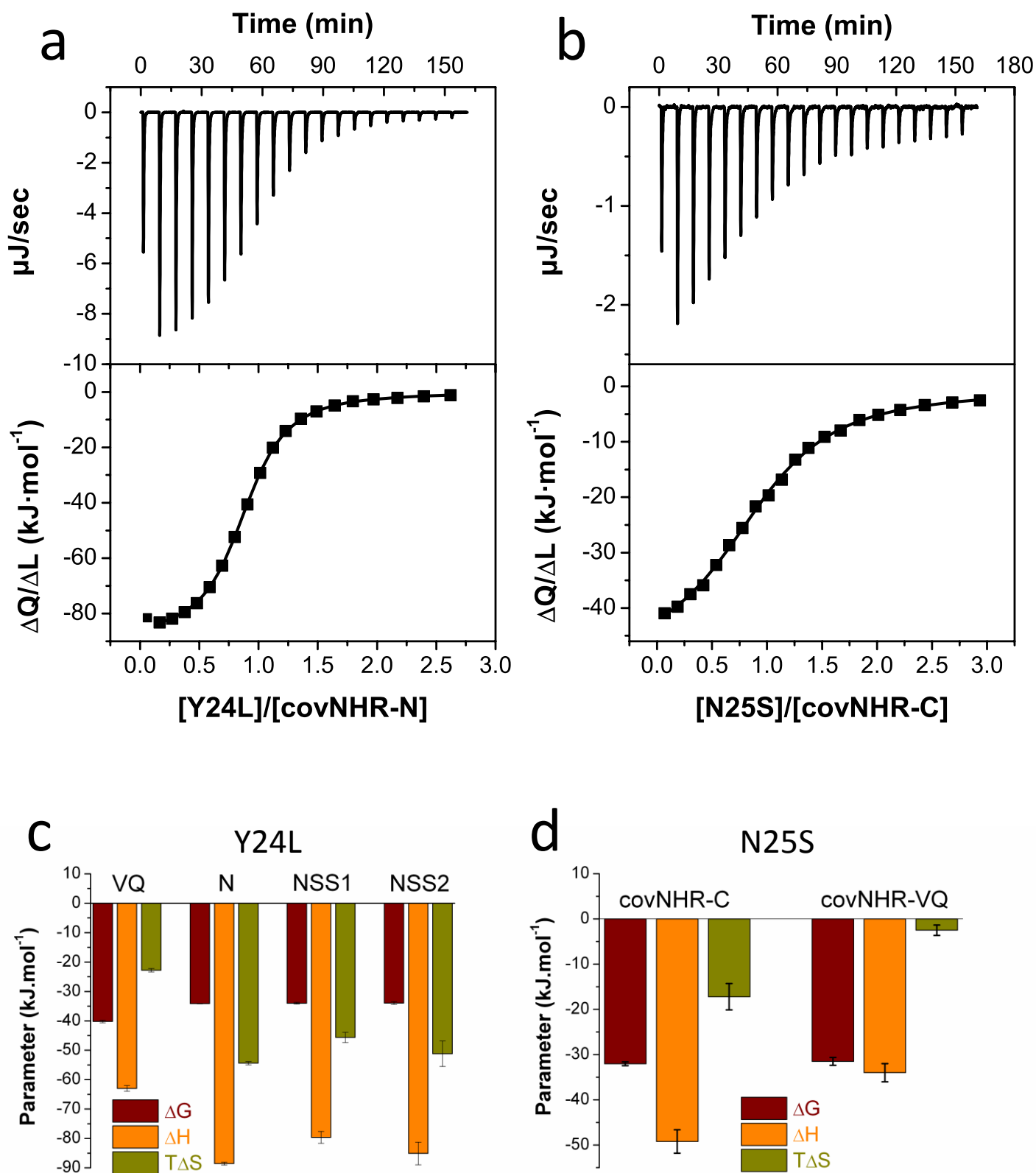


Figure 6

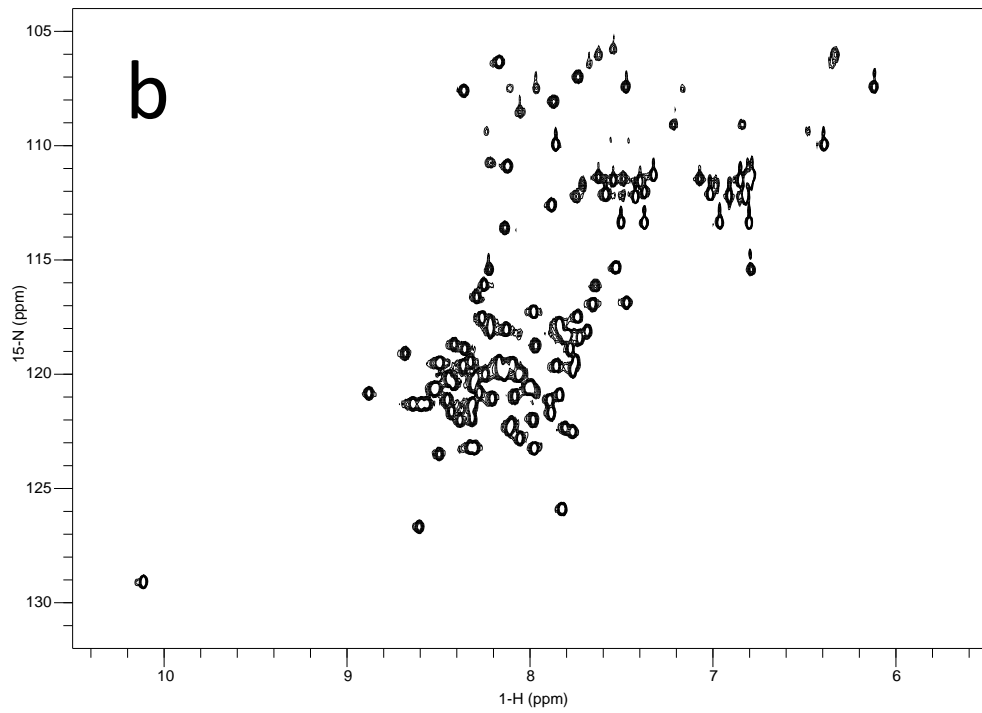
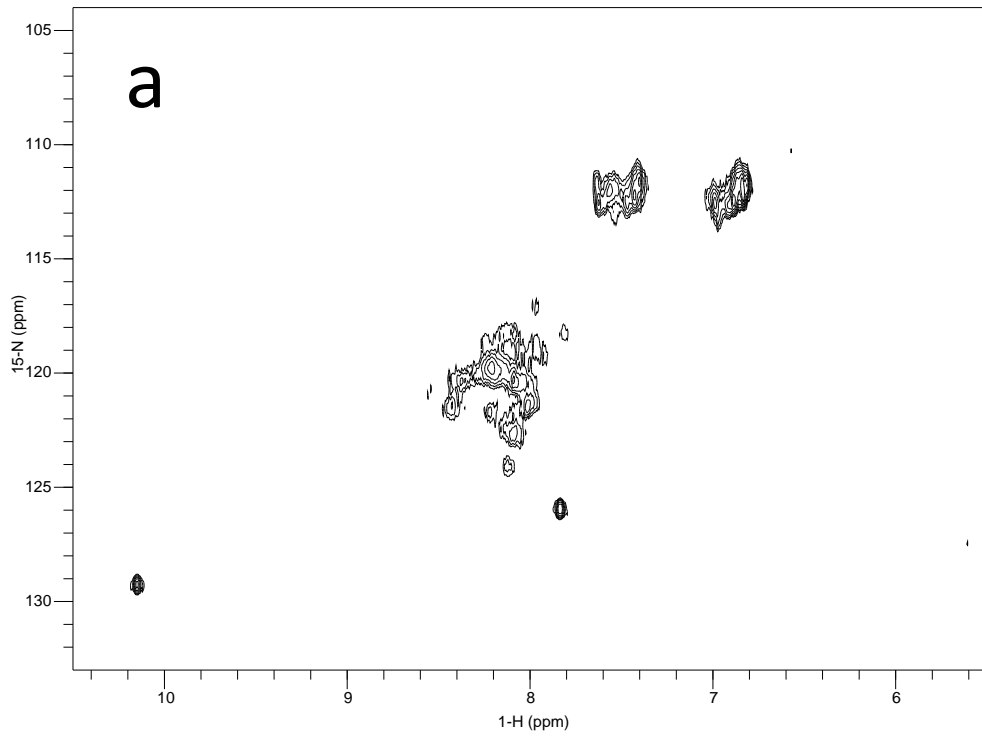
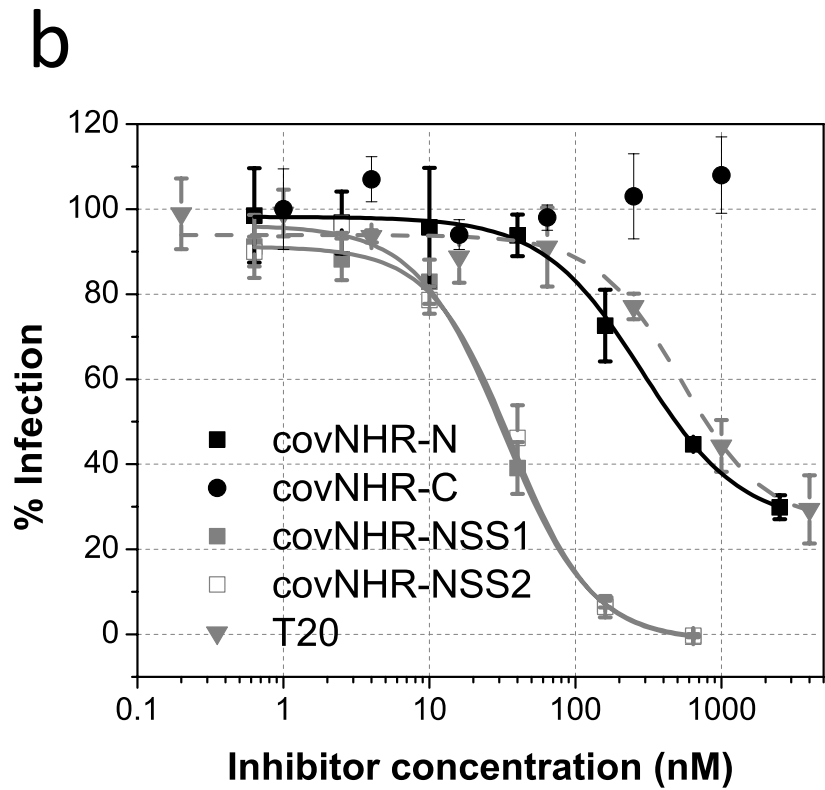
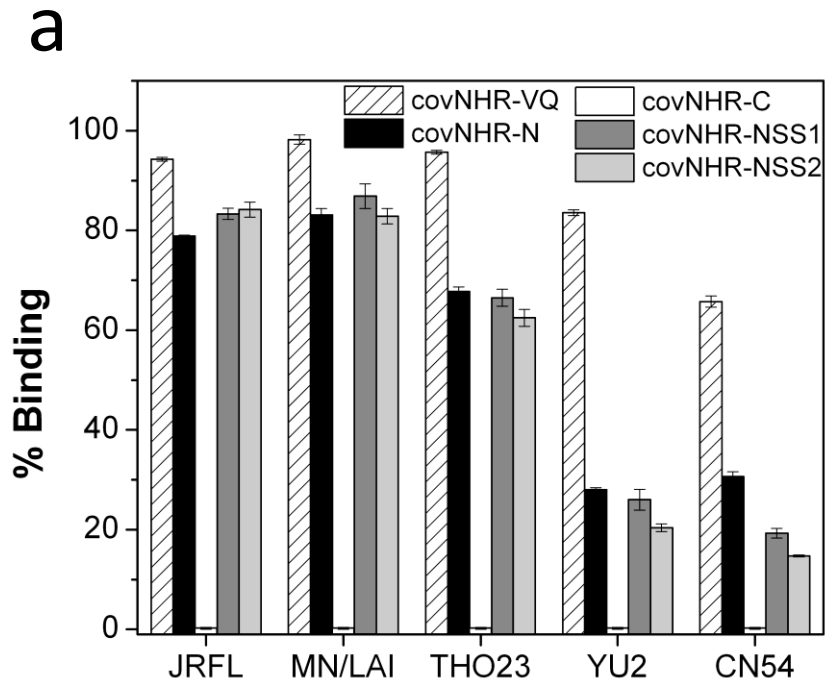


Figure 7





[Click here to access/download](#)

Supplementary Material (To be Published)
Supporting Information_Jurado_revised.docx

