Evolution of Conformational Dynamics Determines the Conversion of a Promiscuous Generalist into a Specialist Enzyme

Taisong Zou,[1] Valeria A. Risso,[2] Jose A. Gavira,[3] Jose M. Sanchez-Ruiz,*[,2] and S. Banu Ozkan*[,1]

[1]Center for Biological Physics, Department of Physics, Arizona State University

[2]Departamento de Química Física, Facultad de Ciencias, Universidad de Granada, Granada, Spain

[3]Laboratorio de Estudios Cristalográficos, Instituto Andaluz de Ciencias de la Tierra (Consejo Superior de Investigaciones Científicas—Universidad de Granada), Granada, Spain

***Corresponding author:** E-mail: sanchezr@ugr.es; banu.ozkan@asu.edu.

Associate editor: Jeffrey Thorne

Abstract

β-lactamases are produced by many modern bacteria as a mechanism of resistance toward β-lactam antibiotics, the most common antibiotics in use. β-lactamases, however, are ancient enzymes that originated billions of years ago. Recently, proteins corresponding to 2- to 3-Gy-old Precambrian nodes in the evolution of Class A β-lactamases have been prepared and shown to be moderately efficient promiscuous catalysts, able to degrade a variety of antibiotics with catalytic efficiency levels similar to those of an average modern enzyme. Remarkably, there are few structural differences (in particular at the active-site regions) between the resurrected enzymes and a penicillin-specialist modern β-lactamase. Here, we propose that the ancestral promiscuity originates from conformational dynamics. We investigate the differences in conformational dynamics of the ancient and extant β-lactamases through MD simulations and quantify the contribution of each position to functionally related dynamics through Dynamic Flexibility Index. The modern TEM-1 lactamase shows a comparatively rigid active-site region, likely reflecting adaptation for efficient degradation of a specific substrate (penicillin), whereas enhanced deformability at the active-site neighborhood in the ancestral resurrected proteins likely accounts for the binding and subsequent degradation of antibiotic molecules of different size and shape. Clustering of the conformational dynamics on the basis of Principal Component Analysis is in agreement with the functional divergence, as the ancient β-lactamases cluster together, separated from their modern descendant. Finally, our analysis leads to testable predictions, as sites of potential relevance for the evolution of dynamics are identified and mutations at those sites are expected to alter substrate-specificity.

Key words: **protein dynamics and structure, ancestral enzyme, molecular dynamics.**

Article

Introduction

It is widely acknowledged that, in many cases, biological function may not be fully understood on the basis of a single static 3D protein structure. Indeed, experimental and computational studies have made increasingly clear that conformational dynamics of protein structures often underlies fundamental molecular phenomena related to biological function, such as enzyme catalysis and molecular recognition (Doucet et al. 2007; Henzler-Wildman et al. 2007; Bahar et al. 2010; Leone et al. 2010; Glembo et al. 2012; Liberles et al. 2012; Sikosek and Chan 2014). It appears plausible then that dynamic features required for function are a product of natural selection, a possibility supported by a number of careful computational studies (Pang et al. 2005; Carnevale et al. 2006; Maguid et al. 2006; Hollup et al. 2011; Glembo et al. 2012; Micheletti 2013; Sikosek and Chan 2014). A particularly interesting possibility in this context is that conformational dynamics plays an important role in protein promiscuity, roughly, the capability of proteins to perform several more or less related molecular tasks. Experimental studies have shown beyond doubt that many proteins are able to perform several functions (Copley 2003; Khersonsky et al. 2006; Babtie et al. 2010; Khersonsky and Tawfik 2010; Garcia-Seisdedos et al. 2012; Duarte et al. 2013). In some cases, the protein is found to display one clearly defined primary function together with several low-level "promiscuous activities." Yet in other cases (e.g., detoxifying enzymes [Zhang et al. 2012]), the protein can efficiently perform several related functions (involving similar chemical transformations and/or substrates) or even clearly different tasks associated with different molecular surfaces or active sites, the so-called protein "moonlighting." It appears reasonable to expect that a special feature of dynamics in functionally relevant protein regions (e.g., conformational diversity, active-site "flexibility," or "deformability," leading to the capability to stabilize different substrates, transition states of leaving groups) may be associated with many cases of protein promiscuity (Khersonsky et al. 2006; Babtie et al. 2010; Khersonsky and Tawfik 2010; Lopez-Canut et al. 2011; Garcia-Seisdedos et al. 2012; Fornili et al. 2013).

The phenomenon of protein promiscuity has important implications both for protein engineering (Kazlauskas 2005;

Nobeli et al. 2009) and for the understanding of molecular evolution (Khersonsky et al. 2006; Babtie et al. 2010; Khersonsky and Tawfik 2010; Garcia-Seisdedos et al. 2012; Duarte et al. 2013). Low, yet significant promiscuous activity levels are often required as a starting point for the directed evolution of high levels of targeted functions of biotechnological interest. Thus, promiscuity is thought to play an essential role in the evolutionary development of new functions. Furthermore, even if some modern enzymes are highly efficient specialists, primordial enzymes were likely promiscuous generalists (Jensen 1976; O'Brien and Herschlag 1999; Khersonsky et al. 2006; Babtie et al. 2010; Khersonsky and Tawfik 2010; Garcia-Seisdedos et al. 2012; Duarte et al. 2013). Many proteins have therefore evolved from generalists into specialists during the course of evolution. Although the molecular mechanisms behind such evolutionary conversions remain poorly understood, we propose that they may be determined to some extent by changes in protein dynamics (Glembo et al. 2012). Certainly, this proposal cannot be tested on the basis of the comparative analysis of modern, evolutionary related proteins, as they do not include the hypothetical promiscuous ancestor. However, recent experimental work (Risso et al. 2013; Risso, Gavira, Gaucher, et al. 2014; Risso, Gavira, Sanchez-Ruiz 2014) on the laboratory resurrection of the ancestors of enzymes involved in antibiotic resistance has recapitulated in the laboratory the evolutionary conversion of generalists into specialist enzymes and demonstrated that such conversion may take place with little change in static X-ray structures, thus suggesting a role for dynamics. As explained below in detail, this recent work on antibiotic resistance proteins (Risso et al. 2013; Risso, Gavira, Gaucher, et al. 2014; Risso, Gavira, Sanchez-Ruiz 2014) provides a unique opportunity to explore the relation among promiscuity, conformational dynamics, and evolution in proteins.

Antibiotic resistance is one of the most serious threats to public health. Bacteria are becoming less and less susceptible to the currently available antibiotics, whereas the development of new antibiotics is becoming more and more difficult and expensive. It is urgent to understand the evolution of antibiotic resistance in order to continually combat bacterial infection (Livermore 1995; Levy and Marshall 2004; Cantón and Coque 2006; Pitout and Laupland 2008; United States Centers for Disease Control and Prevention 2013). Indeed, antibiotic resistance is an ancient phenomenon and resistance genes have been found in uncontaminated environments, such as Alaskan soil (Allen et al. 2009), sediments from the bottom of pacific ocean (Toth et al. 2010), and even 30,000-year-old Beringian permafrost sediments (D'Costa et al. 2011). β-Lactamases in particular probably originated more than 2 Ga, and some of them have been in plasmids for millions of years. To understand the evolution of β-lactamases as well as the evolutionary origin antibiotic resistance, we recently performed an ancestral sequence reconstruction exercise targeting the following Precambrian nodes: The last common ancestor of enterobacteria (ENCA), the last common ancestor of gamma-proteobacteria (GPBCA), the last common ancestor of various Gram-negative bacteria (GNCA), and the last common ancestor of Gram-positive and Gram-negative bacteria (PNCA) (Risso et al. 2013). These ancestors inhabited Earth about 1 Ga (ENCA), 1.5 Ga (GPBCA), 2 Ga (GNCA), and 3 Ga (PNCA) based on the estimates of divergence times. The protein sequences of those ancestors were derived through Bayesian Maximum Likelihood approach in a phylogenetic framework, targeting the Precambrian nodes in the evolution of Class A β-lactamases. The sequence identities of these ancestral proteins range from 53% to 79% in pairwise alignments with TEM-1 β-lactamase, one of their modern descendants. However, despite the significant variations in sequence, they share the canonical lactamase fold and no significant differences are observed at the active site positions when comparing with the extant TEM-1 β-lactamase. More interestingly, laboratory-resurrected β-lactamases corresponding to 2- to 3-Gy-old nodes are highly stable with melting temperate (T_m) about 35° higher than most modern lactamases (Risso et al. 2013). Additionally, they can degrade a variety of antibiotics in vitro with levels of catalytic efficiency similar to that of an average modern enzyme, whereas the modern TEM-1 β-lactamase is clearly a penicillin specialist (Risso et al. 2013). These results supported the thermophilicity of Precambrian life and provided evidence for the evolutionary conversion of generalists into specialists proposed by Jensen (1976) many years ago.

To summarize, laboratory resurrections of Precambrian β-lactamases show dramatically enhanced stability and substrate-promiscuity when compared with some modern decedents, and yet they have very similar 3D structures (Risso et al. 2013, 2014). This prompts the question whether conformational dynamics can provide mechanistic insights about the evolution of β-lactamases. Our approach to probe the relation between substrate promiscuity and dynamics in β-lactamases is based upon the increasingly accepted view that native proteins are to be described as ensembles of more or less diverse conformations. In fact, computational and experimental studies support the notion that, in many cases of substrate promiscuity, the conformations responsible for the different interactions pre-exist in equilibrium in the unligated protein (Pang et al. 2005; Tobi and Bahar 2005; Gerek et al. 2009; Bahar et al. 2010; Gerek and Ozkan 2010; Kar et al. 2010; Khersonsky and Tawfik 2010). This relation between protein interactions and protein dynamics is actually that embodied in the so-called conformational selection model, which has received considerable attention in recent years (Tobi and Bahar 2005; Boehr et al. 2009; Changeux and Edelstein 2011; Munz et al. 2012; Vogt and Di Cera 2012). Inspired by these views, we picture the lactamase molecule in vivo as undergoing transitions between different conformations triggered by random interactions with surrounding solvent molecules, other macromolecules, and so forth. We simulate computationally these random events on the basis of the Perturbation Response Scanning (PRS), which relies on sequentially applying an external random force (i.e., a Brownian kick) on a single residue and recording of the response of other residues (Ikeguchi et al. 2005; Atilgan C and Atilgan AR 2009; Atilgan et al. 2010; Gerek and Ozkan 2011). The PRS results allow us to calculate a metric called

Dynamic Flexibility Index (dfi) that measures the resilience of each given residue to perturbation (Gerek et al. 2013). The special dynamics associated with substrate promiscuity should be revealed by patterns of high *dfi* values in regions close to the active site revealing the deformability required for the binding and catalysis of different ligands. That is, these specific *dfi* patterns would support that the protein exists as an ensemble of conformations displaying the structural variability in the active site region required for efficient binding of substrates of different sizes and shapes. It is relevant to note at this point that, in our recent proposal of the dfi metric in PRS, an Elastic Network Model (ENM) was used to compute the dynamics (Gerek et al. 2013). ENMs describe the protein molecule as a network of nodes connected by uniform springs and have been shown to provide a computationally efficient approach to the calculation of dynamic features that depend on native contact topology, that is, the structure-encoded dynamics (Bahar et al. 2010). However, we are interested here in exploring the differences in dynamics between proteins (ancestral and extant β-lactamases) that share essentially the same static 3D structure as determined by X-ray crystallography. Therefore, we need to go beyond the first-degree approximation (structure-encoded dynamics) and specifically account for the sequence-encoded dynamics within a given structure. To this end, in our application of PRS, we replace the hessian matrix derived from an ENM model with covariance matrices calculated from replica-exchange molecular dynamics (REMD) simulations performed with each specific protein.

The outline of the computational studies reported here is therefore as follows. The three ancestral β-lactamases (corresponding to the nodes PNCA, GNCA, and ENCA) and a modern descendant (TEM-1) are simulated using reservoir REMD (r-REMD), an efficient simulation technique incorporating geometric simulation with REMD algorithm (Roitberg et al. 2007). The simulations provide us with the fundamental dynamics information of those β-lactamases. As we describe in detail further below, the analysis of residue fluctuations indicates that the ancient lactamases are more deformable than TEM-1 lactamase. Furthermore, the results of these MD simulations are used as a basis of a PRS analysis on each β-lactamase from which dfi profiles are calculated. Overall, the conformational dynamics of individual β-lactamases shows changes that are in agreement with the functional divergence: Although the dfi distributions of PNCA and GNCA lactamases are similar to each other and distinctively separated from the functionally divergent TEM-1 lactamase, ENCA lactamase (the most substrate-specific ancient β-lactamase) shows a dfi distribution more similar to that of TEM-1 lactamase. Moreover, TEM-1 lactamase has a more rigid catalytic pocket, suggesting that the shape of the pocket has also evolved toward a specific target, as the catalysis becomes penicillin-specific. Finally, we also analyze the statistical pattern of their dynamics profiles using singular value decomposition (SVD), which enhances the signal-to-noise ratio of the data by expressing them as a linear combination of a few dominant principal components. On the basis of their pairwise distances in the subspace of principal components, a cladogram is constructed to illustrate the evolutionary relationship of these β-lactamases in terms of dynamics. Furthermore, through SVD, we identify likely candidates for mutations that are critical for the dynamics divergence among β-lactamases. Changes of the dynamics caused by mutations at those critical sites might potentially lead to the change of the substrate-specificity of the protein. In summary, our findings suggest that divergent evolution of conformational dynamics best explains the evolution of catalytic function in β-lactamases. Thus, the analysis of the detailed conformational dynamics could help us understand the underlying mechanism of evolutionary change from promiscuous generalists into specialist enzymes.

Results and Discussion

Structure Inspection

The sequences of ENCA, GNCA, and PNCA lactamases substantially differ from the sequences of modern β-lactamases and, in particular, they show 79.2%, 53.6%, and 52.9% sequence identity with the sequence of the modern TEM-1. Despite of the extensive sequence differences, they all share the canonical β-lactamase fold with all-atom RMSDs of 0.53, 0.76, and 0.86 Å with respect to TEM-1 lactamase (fig. 1*A*). Closer inspection of RMSD at individual residue sites reveals minor movement with RMSD ≤ 2 Å in the $\alpha + \beta$ domains of the GNCA and PNCA lactamases corresponding to N-terminal helix and solvent-exposed loops (fig. 1*B*). Moreover, no substantial differences are found in the α-domain and all active site residues occupy canonical space (fig. 1*C*). Therefore, the structural analysis is not sensitive enough to address the cause of the functional divergence, that is, how the β-lactamases evolve from substrate-promiscuous generalists to specialists.

The Relationship between Functional Divergence and Structural Dynamics

Here, we turn to investigate the role of structural dynamics on the functional divergence observed among the β-lactamases. The unbound conformations of the three ancestral β-lactamases (PNCA, GNCA, and ENCA) and a modern descendant (TEM-1) are simulated using r-REMD. r-REMD incorporates the conformations generated by the geometric simulation algorithm FRODA (Wells et al. 2005) as reservoir structures, which in turn increases the efficiency of conformational sampling (Roitberg et al. 2007). We first analyze the root mean square fluctuation (RMSF) of residues for each β-lactamase. RMSF is a measure of the positional deviation of a residue over time from its time-averaged position. Although the structural analysis does not appear to reveal any functionally relevant differences, the RMSF profiles do provide some first indication of such differences. Indeed, the ancient lactamases (PNCA, GNCA, and ENCA) fluctuate somewhat more than TEM-1 lactamase. More interestingly, at the essential active site S70, PNCA and GNCA lactamases show significantly higher fluctuation (i.e., more flexibility) than ENCA and TEM-1 lactamase (fig. 2).

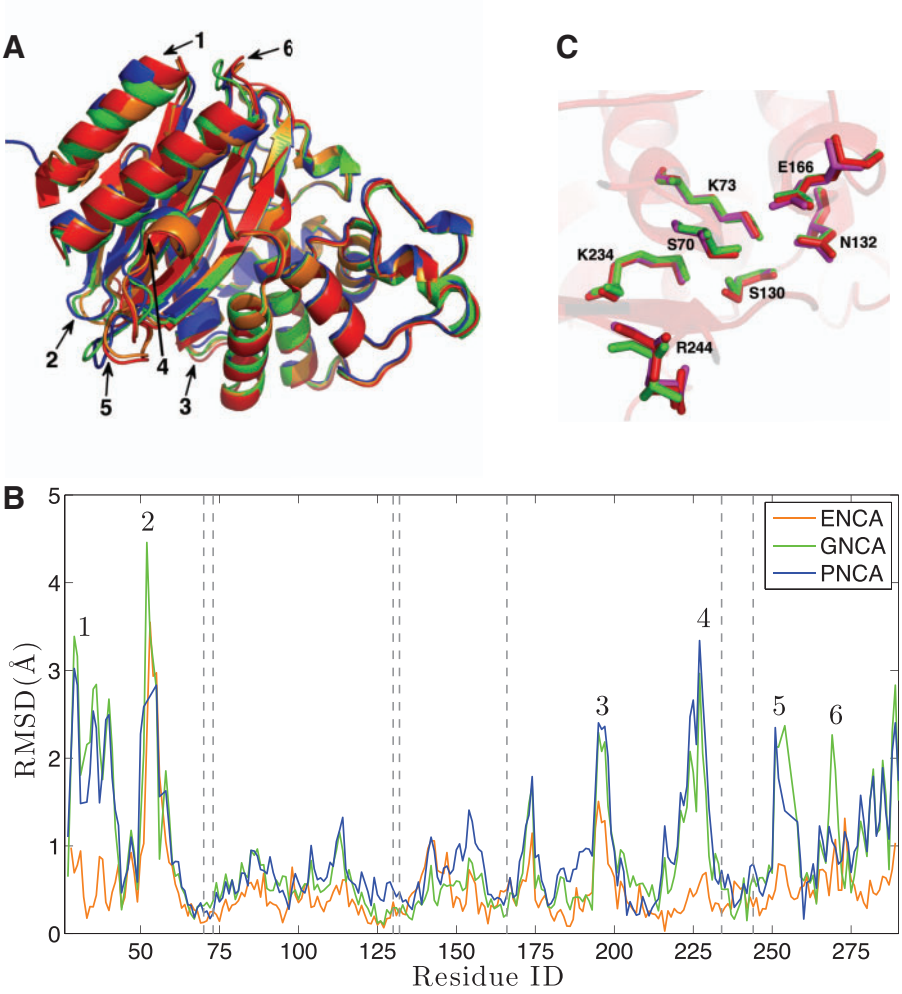

Fig. 1. Structural characterization of laboratory resurrections of Precambrian β-lactamases. (*A*) Structural comparison of the TEM-1 β-lactamase (PDB: 1BTL; red), and the β-lactamases corresponding to the last common ancestor of ENCA (PDB: 3ZDJ; orange), the last common ancestor of various GNCA (PDB: 4B88; green), and the last common ancestor of PNCA (PDB: 4C6Y; blue). (*B*) RMSD of individual residues along the sequence. The vertical dash lines mark the location of active sites. Minor structural differences are seen in the N-terminal helix and solvent-exposed loops (labeled 1–6) (*C*). Close examination of the structural differences at the active site.

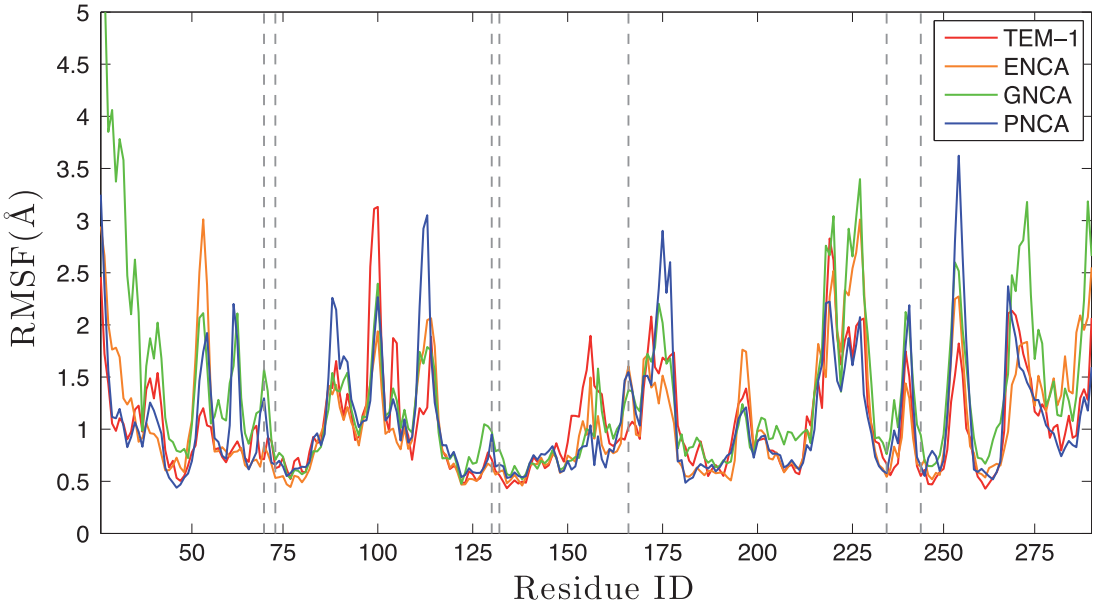

Fig. 2. The RMSF of Cα atoms in TEM-1 (red), ENCA (orange; 1 Gy), GNCA (green; 2 Gy) and PNCA (blue; 3 Gy). The vertical dash lines mark the location of active sites.

RMSF profiles present the structural dynamics of β-lactamase extracted from their equilibrated unbound conformational dynamics. We are, however, more interested in capturing the dynamics profiles of each position as they deviate from unbound equilibrium, as these profiles may shed light on the response of the protein to an approaching substrate as it exerts forces on the protein. For this purpose, we apply PRS approach (see Materials and Methods). In PRS, we introduce perturbations by applying a random external unit force on single residues, and then analyze the residue response fluctuation profile of the rest of the chain using linear response theory. It has been shown in the past that PRS and its derivative are powerful tools to 1) capture conformational changes upon binding (Atilgan C and Atilgan AR 2009; Atilgan et al. 2010), 2) reveal allosteric pathways and identify critical residues that mediate long-range communication (Gerek and Ozkan 2011), 3) generate an ensemble of configuration rapidly for flexible docking that improves binding affinity score (Bolia et al. 2012), and 4) distinguish disease-associated and putatively neutral population variations in human proteome (Gerek et al. 2013). To ensure the isotropicity of perturbation, the Brownian kick is applied at ten different directions for individual sites one at a time. The magnitude of displacement by residue i in respond to the perturbation at residue j is given by the mean square fluctuation $|\Delta R^j|_i$. Then, the perturbation is repeated at all other residues and dynamics flexibility index dfi is normalized average mean square fluctuation of a site upon perturbations of others as shown in equation (4) (see Materials and Methods).

As defined, dfi is a relative value that indicates the average response of any residue site in a protein structure. It measures the resilience of each of individual site to perturbation as it arises from interactions, binding processes or amino acid substitutions. Sites with high dfi are more flexible and prone to "feel" the perturbation of other residues. Furthermore, due to this enhanced flexibility, regions encompassing several high-dfi residues are expected to be more deformable overall. On the other hand, sites with low dfi may absorb and transfer the perturbation throughout the protein in a cascade fashion. They are usually involved with hinge parts of the protein that control the motion like joints in skeleton. Therefore, the dfi could evaluate the contribution of each site to the functionally important dynamics. To eliminate the effect of the global flexibility of different proteins, we here compute the rank of the dfi profile and label it as %dfi. In figures 3*A* and *B*, the %dfi profiles for the four β-lactamases under study are compared. It is useful to examine first the dynamics at the active-site residues participating in catalysis. The catalytic mechanism of class A β-lactamase involves the acylation of the active site S70, followed by deacylation. During this process, a general base is expected to activate the primary catalytic site S70 by accepting the proton from it (Lamotte-Brasseur et al. 1991; Strynadka et al. 1992). Although the identification of the general base as K73 or E166 remains controversial, it is well known that the two sites are critical in this proton transfer event (Chen et al. 1996; Damblon et al. 1996; Atanasov et al. 2000). In addition, several residues, such as S130, N132, K234, R244 (K244 in PNCA/GNCA), are also identified as important for catalysis (Lamotte-Brasseur et al. 1991; Delairesq et al. 1992; Atanasov et al. 2000). According to the %dfi profile, TEM-1 and ENCA lactamases have lower values than PNCA/GNCA at those active site residues. However, it is important to note that in general, it is the positions in the vicinity of the active site that show the most pronounced change in the dynamic profile when comparing ancestral and extant beta-lactamases. In fact, visual inspection of the %dfi profiles identifies four regions (a–d) showing significant flexibility discrepancy among the β-lactamases studied (fig. 3*C*): 1) Region a (residues 57–75) consists of part of helix H2 and a loop region between strand B2 and helix H2, strand B2), 2) region b (residues 123–134) includes a loop region between H4 and H5, 3) region c (residues 140–161) consists of part of helix H6 and a loop region between helix H6 and Ω-loop, and 4) region d (residues 240–267) spans from strands B3 to B5. The dynamics and structural details of those regions are provided in figure 4. Overall, we observe increasing %*dfi* values from TEM-1 to PNCA lactamase in all those regions except the region c, where the trend becomes the opposite. The three regions a, b, and d span the active site and nearby residues, supporting an enhanced active-site deformability in the generalists PNCA and GNCA lactamases (as compared with the specialists ENCA and TEM-1 lactamases), a result which may explain their capability to accommodate antibiotic molecules of different size and shape (fig. 4*A*). This finding also suggests the catalytic specificity in the modern β-lactamase evolved through the decrease of flexibility/deformability in the binding pocket as observed earlier for the evolution of stress hormone receptor (Glembo et al. 2012). That is, PNCA and GNCA lactamases are more deformable around the active site and thus show higher catalytic promiscuity, whereas TEM-1 and ENCA lactamases are more rigid and thus more substrate-specific. Interestingly, region c, which does not include the active-site and nearby residues, shows a trend opposite to that of regions a, b, and d; that is, the deformability of TEM-1 in that region would be higher and there is trend of decreasing deformability from TEM-1 to PNCA. An interesting, although speculative, possibility is that, for this system, there is a correlation between deformability and stability in such a way that a decrease in stability linked to enhanced deformability (i.e., higher *dfi* trend) at the active-site neighborhood is compensated by decreased deformability (i.e., lower *dfi* trend) at other regions.

Investigation of Catalytic Pocket Dynamics

We have also measured the catalytic pocket volume of the β-lactamases using POVME (POcket Volume MEasurer) (Durrant et al. 2011). It fills the regions that encompass the catalytic pocket with equispaced points and then removes those points near protein atoms. The volume of the catalytic pocket is then obtained by counting the remaining points. The pocket volumes in the X-ray structures are 240, 264, 240, and 192 $Å^3$ for TEM-1, ENCA, GNCA, and PNCA lactamases, respectively. Thus, the pocket volume in crystal structure does not show any trend regarding functional specificity

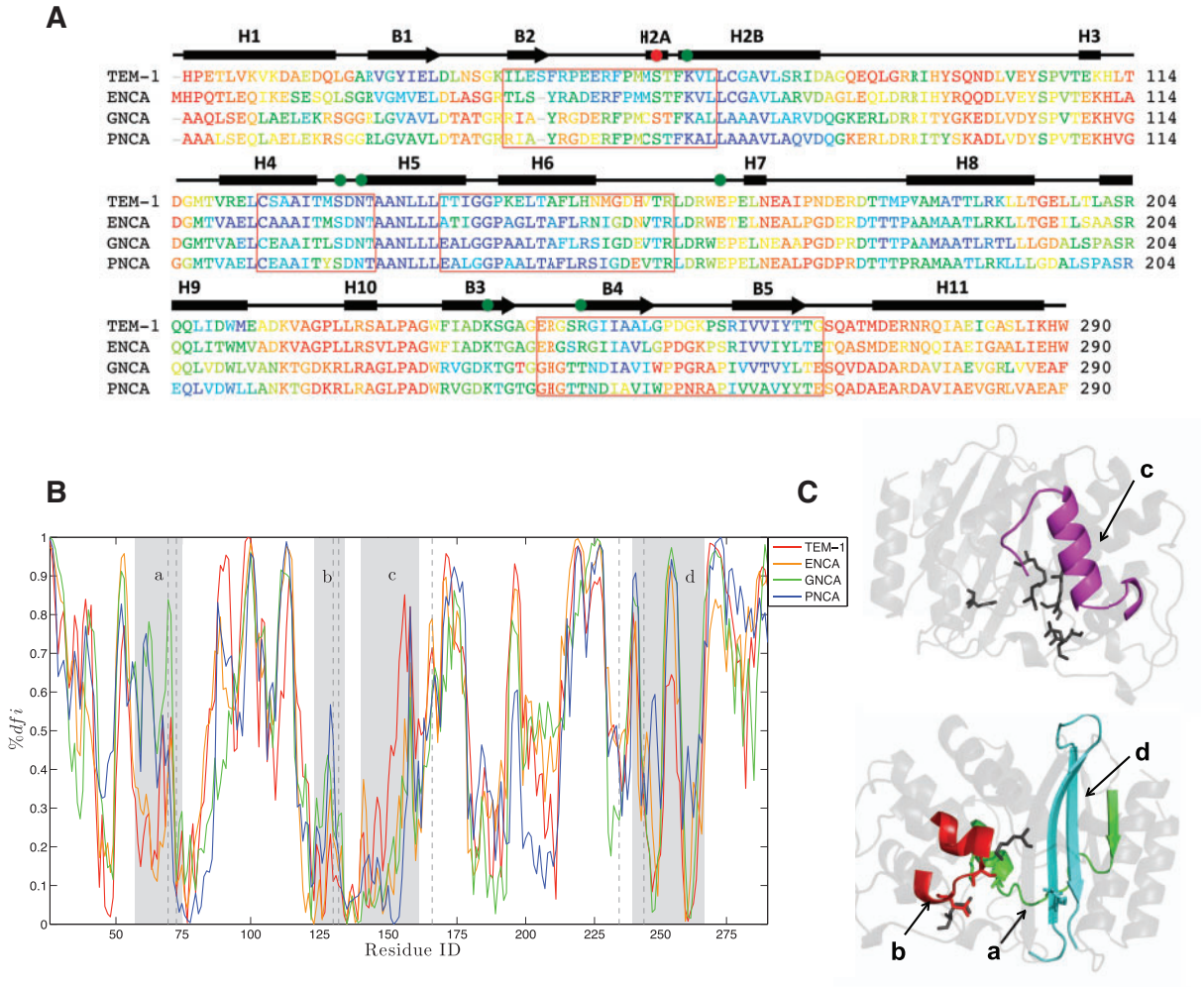

Fig. 3. The dynamics profile of residues in TEM-1 (modern), ENCA (1 Gy), GNCA (2 Gy), and PNCA (3 Gy) β-lactamases. (*A*) The *%dfi* index is mapped onto the multiple sequence alignment of the four β-lactamases. Residues are colored with a spectrum of red to blue, where rigid residues are denoted by blue/green and flexible regions are denoted with red/orange. The primary active site S70 is marked with red dot and other active sites are marked with green dots. Five regions where the β-lactamases show high discrepancy by visual observation are marked with red boxes (region a: residues 57–75; region b: residues 123–134; region c: residues 140–161; region d: residues 240–267). (*B*) The *%dfi* distribution in the four β-lactamases: TEM-1 (red), ENCA (orange), GNCA (green), and PNCA (blue). The vertical dash lines mark the location of active site residues. The five regions with high discrepancy are marked in gray shadow. (*C*) Mapping of the four regions (a–d) with significant dfi difference among β-lactamase to the structure. The active site residues are displayed with sticks. The dynamics and structural details of those regions are shown in figure 4.

versus promiscuity. Then, we analyze how the pocket volume fluctuates and generates volume-size ensemble during the simulations for these four proteins. Figure 5 shows how the pocket volume changes over the last 2-ns r-REMD of each molecular dynamics simulation. Quantitatively, the mean pocket volumes and their fluctuations from the molecular simulations are 170 ± 84, 216 ± 73, 325 ± 117, and 349 ± 112 Å^3 for TEM-1, ENCA, GNCA, and PNCA lactamases, respectively. It indicates that 1) TEM-1 has the smallest catalytic pocket, and the catalytic pockets PNCA and GNCA are larger than ENCA and TEM-1; and 2) the pocket volumes of TEM-1 and ENCA also fluctuate less than GNCA and PNCA. Interestingly, benzylpenicillin is the smallest substrates in the study with the volume 286 Å^3, compared with the volume of cefotaxime 353 Å^3 and that of ceftazidime 437 Å^3 (www.chemicalize.org, last accessed October 19, 2014). Therefore, it is reasonable that PNCA and GNCA can accommodate all these substrates with different size due to their large pockets and also their higher deformability that facilitates the necessary induced conformational changes. We emphasize again that these are dynamic features of the oldest ancestral proteins which are not apparent in the static X-ray structures but that become obvious in the MD simulations. On the other hand, TEM-1 and ENCA are specific likely because through evolution the conformational dynamics and size of the pocket has been shaped for the small penicillin substrate and large substrates such as cefotaxime and ceftazidime cannot access the active site due limited space and lesser deformability.

Clustering Proteins Based on Dynamics Profile and Identify Potentially Function Altering Mutations

Despite the conserved structures of these four β-lactamases, the experimental characterization has shown that their

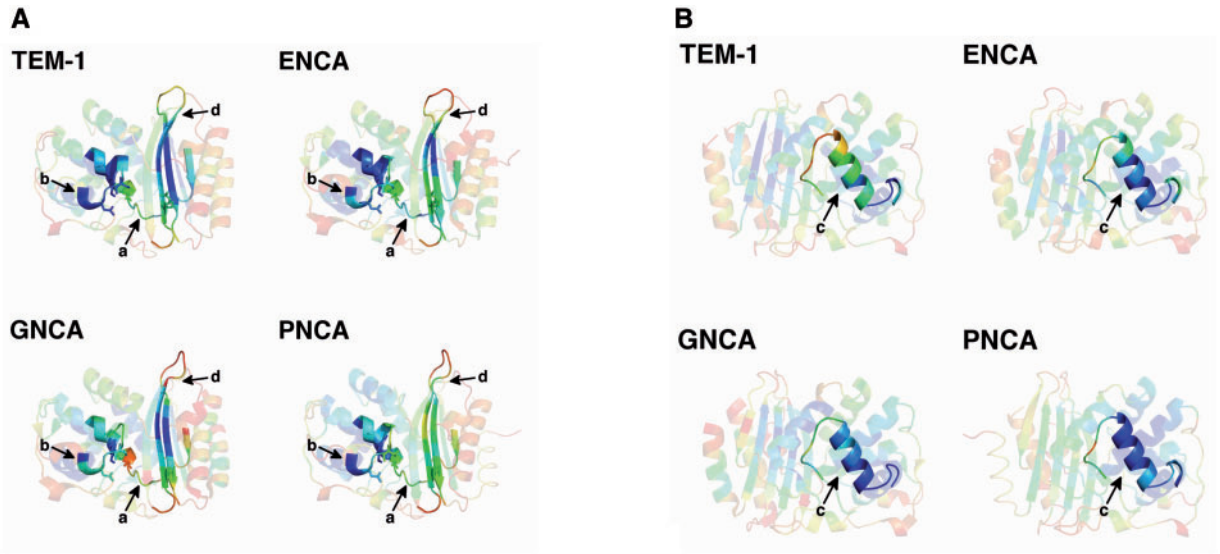

FIG. 4. Close investigation of four regions with significant *dfi* difference among the four β-lactamases studied. These regions are colored with a spectrum of red to blue, where lowest *dfi* regions are denoted by blue and flexible regions are denoted with red/orange. The active sites are displayed with sticks. (*A*) In region a (residues 57–75), TEM-1 and ENCA lactamases have lower dfi profile than GNCA and PNCA lactamases, especially at the active sites S70 and K73, indicating the alteration in conformational dynamics linked to the evolution of this β-lactamases. In region b (residues 123–134), TEM-1 and ENCA lactamases are also more rigid than GNCA and PNCA lactamases, especially at the active sites S130 and N132. Overall, the catalytic pocket, surrounded by regions a, b, and d, exhibits a trend of increased flexibility when going from the specialists (ENCA and TEM1 lactamases) to the ancestral generalists (PNCA and GNCA lactamases). (*B*) In contrast, region c (residues 250–267) seems to evolve toward lower dfi values as we examine the change from PNCA to TEM-1 and, in fact, TEM-1 lactamase is significantly more rigid than the other three proteins in this region.

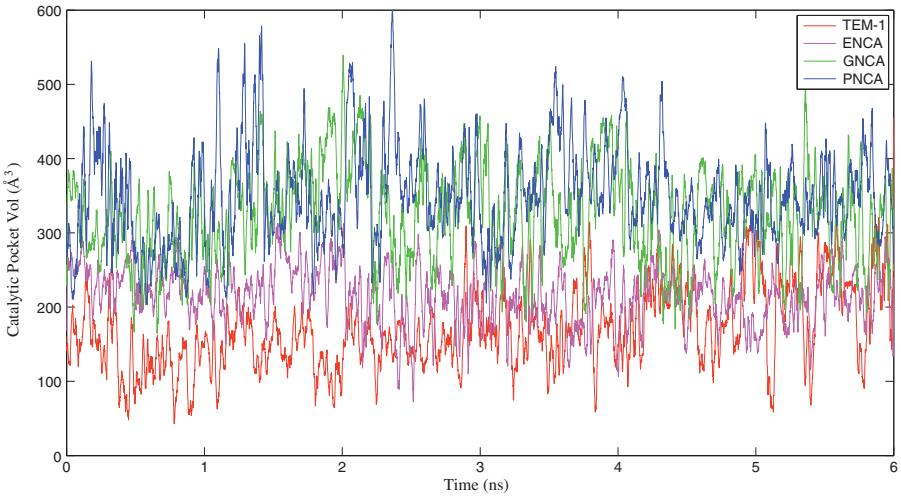

FIG. 5. The volumes of the catalytic pockets of TEM-1, ENCA, GNCA, and PNCA β-lactamases as a function of time for the last 6 ns of an r-REMD simulation. The promiscuous ancestral enzymes PNCA and GNCA lactamases exhibit larger flexibility and pocket volume compared with the specialists ENCA and TEM-1 lactamases.

antibiotic degradation patterns are different, with PNCA and GNCA showing substrate promiscuity (capability to degrade different antibiotics) and TEM-1 and ENCA lactamases being more specific toward penicillins. In order to further clarify the relation between the underlying structural dynamics of these four proteins and the functional differences, we performed an SVD-based clustering analysis of the pattern of dfi profiles for the four β-lactamases. Subsequently, we used the pairwise distances in the left subspace of SVD (fig. 6*A*) to construct the cladogram for clusters shown in figure 6*B*. Interestingly, the β-lactamases are separated into two major branches, with one branch consisting only TEM-1 lactamase (the most modern one), and the remaining three lactamases appearing in the other branch. The larger branch found in this analysis is divided into two subbranches separating ENCA from PNCA and GNCA lactamases. This result shows that PNCA and GNCA lactamases are very similar to each other but further separated from TEM-1 lactamase based on their dynamic characterization.

SVD analysis also enables us to identify mutations that are plausible candidates for being linked to the structural dynamics divergence. The weight of each site, given by its contribution to the top principal components, may be viewed as a measure of the site impact on the dynamic differences

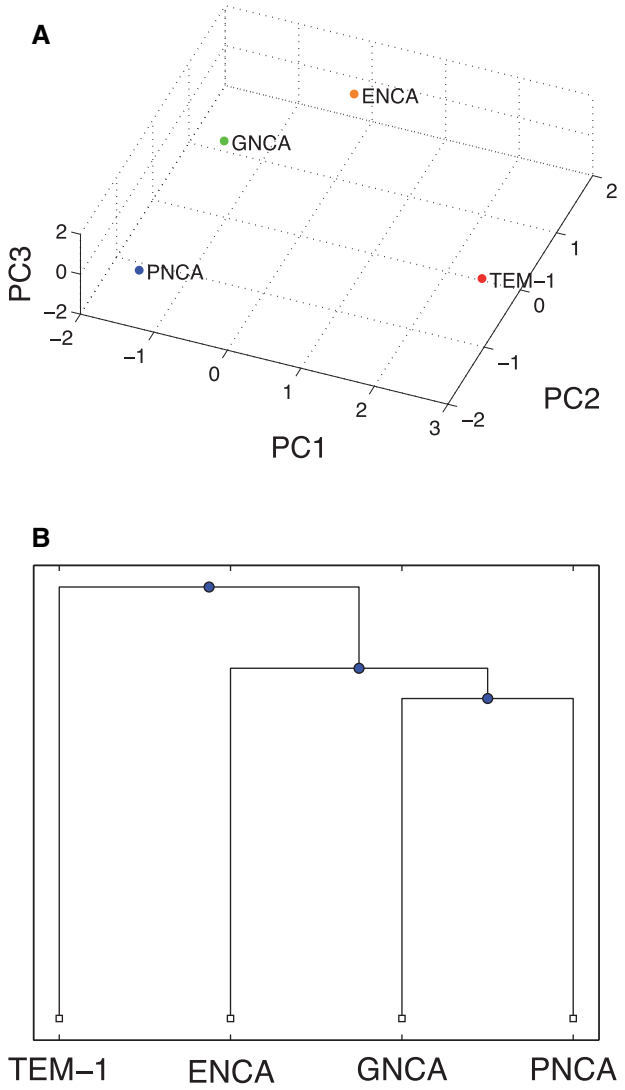

FIG. 6. Clustering of β-lactamases based on the dynamics profiles. (*A*) Distribution of β-lactamases in the subspace formed by top three left-singular vectors (principal components or PC). (*B*) Cladogram of SVD distances for β-lactamases determined from their *dfi* data at 262 residue sites. The structural dynamics clustering agrees with the functional divergence of these enzymes.

between the β-lactamases (see Materials and Methods). Figure 7 displays the weights of all residue sites and marks the statistically critical sites that have large weights. Interestingly, as we also observed from the %dfi profiles, there is a trend for positions around the active site to be involved in the alteration of conformational dynamics and, plausibly, to be linked the evolutionary conversion from promiscuous generalists into specialist enzymes. Overall, sites with large weights in figure 7 are likely important for the change of the dynamics that leads to functional divergence and are, therefore, obvious candidates for mutational studies aimed at engineering substrate specificity.

Conclusion

It is increasingly accepted that protein dynamics underlies biological function and, therefore, that functionally relevant dynamical features in proteins play a major role in protein evolution. Direct evidence of this proposed evolution–dynamics relation has remained, nevertheless, elusive. For instance, several computational comparisons between the dynamics of homolog proteins have indeed found conservation of important dynamic features (Carnevale et al. 2006; Maguid et al. 2006, 2008; Micheletti 2013). However, it has been noted that evolutionary conservation of low-energy collective modes may not reflect their functional importance but rather be a consequence of the robustness of these modes against mutation effects (Echave and Fernandez 2010; Echave 2013). These uncertainties in interpretation are common to the so-called horizontal approaches to molecular evolution, that is, approaches based on the comparison between extant proteins (Harms and Thornton 2010; Ingles-Prieto et al. 2013). On the other hand, several important issues in molecular evolution have been successfully addressed in recent work on the basis of the comparison between extant and ancestral resurrected proteins (Harms and Thornton 2010; Finnigan et al. 2012; Hobbs et al. 2012; Voordeckers et al. 2012; Bar-Rogovsky et al. 2013; Ingles-Prieto et al. 2013; Risso et al. 2013; Kratzer et al. 2014; Risso, Gavira, Sanchez-Ruiz 2014).

Here, we have used this "vertical" approach to explore the relation between protein dynamics and protein evolution. We find that the evolution of conformational dynamics best explains how modern specialists β-lactamases arose from ancestral promiscuous ancestors over a time scale of several billion years. Our results thus provide evidence for the notions that dynamics is crucial for biological function and that variations in sequence can lead to change in dynamics without changing the structure even in functional sites. Therefore, conformational dynamics can be one of the mechanisms. Nature uses to evolve at a molecular level. Furthermore, we have shown how computational analyses can be used to determine sites of potential relevance for the evolution of dynamics thus paving the way for the protein engineering exploitation of protein dynamics

Materials and Methods

Structure Refinement and Simulation

The refinement and equilibrium sampling of ancestral β-lactamases is accomplished with r-REMD (Roitberg et al. 2007). REMD samples the system by molecular dynamics at different temperatures (replicas) and allows the system to attempt exchange between replicas (Sugita and Okamoto 1999). By doing so, systems at high temperature might overcome potential energy barriers and explore a large volume of configuration space. A structure reservoir is prepared and coupled with REMD, a procedure that has shown to improve sampling and capture equilibrium dynamics much more efficiently (Glembo et al. 2012). The system at the highest temperature replica is also allowed to exchange the configuration with the reservoir structure periodically. The configurations in the reservoir are generated using a highly efficient geometric-based sampling technique called FRODA (Wells et al. 2005), which decomposes a protein into a set of small rigid units and interactions are modeled as harmonic constraints. With FRODA and its unfolding version FRODAN (Farrell et al.

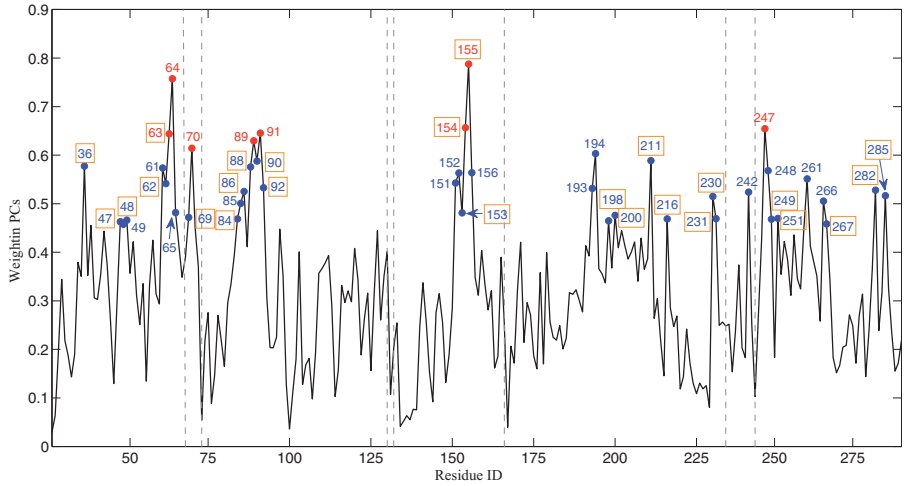

Fig. 7. Weights of residue sites based on their contribution in the top principal components are plotted. The sites whose weights deviate more than twice of standard deviation and one standard deviation from the mean are labeled in red and blue, respectively. The structural dynamics of these sites may contribute the most to the conversion of a promiscuous generalist into a specialist enzyme. The sites where the residue types are not consistent among the four enzymes (i.e., mutational sites) in the four β-lactamases are marked in box.

2010), we generate a large ensemble of partially unfolded conformations. In detail, we first run a restrained simulation for 1.5 ns with 40 replicas from 270 to 450 K in the AMBER96SB force field (Pearlman et al. 1995) with generalized born implicit solvent model (Onufriev et al. 2004). The residue–residue pairs are constrained if their Cα atoms are within 8.0 Å cutoff distance among 90% of reservoir structures. The residue–residue constraints are applied at the Cα atoms of the residues and the force constant is 0.5 kcal/(mol Å^2). After the restrained run, an unrestrained r-REMD with identical parameters is allowed to continue for 15 ns, which gives a total of 600 ns (more than half microsecond) aggregate simulation time. The typical swap likelihood we choose between replicas is approximately 48% (set by appropriate spacing between replica temperatures for an appropriate Boltzmann factor). A convergence analysis is then performed by evaluating the correlation between the successive slowest modes obtained from the covariance matrices of different windows using the MD trajectory. The simulations were considered to have reached convergence when the correlation was above 0.80.

PRS Model and dfi

The canonical PRS model was originally based on the ENM, where the protein is viewed as an elastic network in which each node represents a residue and an harmonic interaction is assigned to a pair of residues if the two residues are within a specified cutoff distance (Atilgan et al. 2001; Atilgan C and Atilgan AR 2009). PRS relies on sequentially applying externally random force (i.e., perturbation like a Brownian kick) on a single residue. The perturbation cascades throughout the residue interaction network and may introduce conformational changes of the protein. The linear responses of other residues are formed as

$$[\Delta \mathbf{R}]_{3N\times 1} = [\Delta \mathbf{H}]^{-1}{}_{3N\times 3N}[\mathbf{F}]_{3N\times 1}, \quad (1)$$

where F is a unit random force on selected residues, H^{-1} is the iinverse of Hessian matrix, and ΔR is the positional displacements of the N residues of the protein in three dimensions.

A disadvantage of the ENM-based PRS model is that the coarse-grained network makes it insensitive to changes arising from the biochemical specificity of amino acids. Therefore, in order to compare the ancestral β-lactamases with similar backbone structures, we replace the ENM basis of PRS with all-atom r-REMD simulations, where the inverse of Hessian matrix is replaced with the covariance matrix G derived from the MD trajectory, that is,

$$[\Delta \mathbf{R}]_{3N\times 1} = [\mathbf{G}]_{3N\times 3N}[\mathbf{F}]_{3N\times 1}. \quad (2)$$

MD simulations take into account long-range interactions as well as the biochemical specificity of amino acids. Thus incorporating MD allows PRS to provide more insights about specific residues beyond the scope of the canonical PRS.

The metric by which PRS quantifies the flexibility of a residue upon the perturbation of other residues is called dfi. To compute dfi, we first apply a unit external force on a single residue. The response vector of positional displacement ΔR is computed by equation (2). To ensure the isotropicity of perturbation, we perform the perturbation in ten directions and take the average of the response vectors. The perturbation is repeated for each residue site and we obtain the perturbation matrix which records the displacement for each residue upon the perturbation of the other residues like

$$[\mathbf{A}]_{N\times N} = \begin{pmatrix} |\Delta R^1|_1 & |\Delta R^2|_1 & \dots & |\Delta R^N|_1 \\ |\Delta R^1|_2 & |\Delta R^2|_2 & \dots & |\Delta R^N|_2 \\ \vdots & \vdots & \ddots & \vdots \\ |\Delta R^1|_{N-1} & |\Delta R^2|_{N-1} & \dots & |\Delta R^N|_{N-1} \\ |\Delta R^1|_N & |\Delta R^2|_N & \dots & |\Delta R^N|_N \end{pmatrix}, \quad (3)$$

where $|\Delta R^j|_i = \sqrt{\langle(\Delta R)^2\rangle}$ denotes the magnitude of the displacement by residue i in response to the perturbation at residue j. A given row of this matrix presents the average displacement of a specific residue from its equilibrium position when other residues are perturbed one at a time, whereas the column shows the response profile of all residues under the perturbation of a specific residue. dfi is defined as the total displacement of residue i induced by perturbations placed on the rest of the residues in the protein, that is, the sum of element in column j of the matrix above, normalized by the total displacement of all residues in the protein. That is,

$$dfi_i = \frac{\sum_{j=1}^{N} |\Delta R^j|_i}{\sum_{i=1}^{N}\sum_{j=1}^{N} |\Delta R^j|_i}. \quad (4)$$

In our analysis we use a 2-ns window frame for computing the covariance matrix of different time intervals of last 6-ns simulations of the lowest replica, and then compute the *dfi* profiles of the residues using each covariance matrix and then compute the average *dfi* per position and ranked by converting to *%dfi*.

Singular Value Decomposition for Clustering and Identifying Functionally Important Dynamics

Singular Value Decomposition (SVD) is a multivariate statistical procedure to elucidate the underlying structure of data. It could be used to increase the signal-to-noise ratio and reduce the redundancy of data. Similar to Principal Component Analysis, SVD transforms data to new subspaces identified by orthonormal bases where the covariance of the data along different orthonormal bases is minimized. It is a powerful tool widely used from information science to biology (Deerwester et al. 1990; Berry et al. 1995; Keskin et al. 2000; West et al. 2001; Kluger et al. 2003).

Let $\mathbf{X}$ denotes an $m \times n$ matrix of interest which contains the information of n subjects characterized by m attributes. The row vector $\mathbf{a}_i$ with n dimension represents the ith attribute in all n subjects, whereas the column vector $\mathbf{b}_j$ with m dimension represents the m attributes for the j subject. In general, SVD decomposes X into the product of three other matrices:

$$[X]_{mxn} = [U]_{mxm}[\Sigma]_{mxn}[V]_{nxn} \quad (5)$$

such that $\mathbf{U}$ and $\mathbf{V}$ have orthonormal columns and Σ is diagonal. The columns of $\mathbf{U}$ (Allen et al. 2009) are left-singular vectors and also the eigenvectors of $\mathbf{X}\mathbf{X}^{\mathrm{T}}$. The columns of $\mathbf{V}$, $\{\mathbf{v}_k\}$, are right-singular vectors and also the eigenvectors of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$. By convention, the diagonal elements of Σ, called singular values of $\mathbf{X}$, are sorted in descending order, that is, $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ $(\sigma_1 \geq \sigma_2 \ldots \geq \sigma_n)$. These diagonal elements represent the variance along the corresponding left-singular and right-singular vectors. Those vectors with large variance are interpreted to be important as they are most relevant to the main characteristics included in the matrix $\mathbf{X}$.

On one hand, the left-singular vectors can be considered as the eigenvectors spanning the new subject subspace. If one wishes to understand the relationship among the subjects, it is necessary to find out the new coordinates of the subjects in this left-singular subspace. The original coordinates of subject j is given by the column vector $\mathbf{b}_j$. Referring to the definition in equation (5), the SVD equation for $\mathbf{b}_j$ is

$$b_j = \sum_{k=1}^{r} v_{jk}\sigma_k u_k, \quad (6)$$

which is a linear combination of the left-singular vectors (Babtie et al. 2010). r denotes the rank of matrix $\mathbf{X}$. According to equation (6), the jth row of $\mathbf{V}\Sigma$, designated as $\mathbf{b}_j'$ gives the coordinate of the subject j in the left-singular subspace $\{\mathbf{u}_k\}$. If $r < m$, the attributes of subject can be captured with fewer variables by $\mathbf{b}_j'$ instead of $\mathbf{b}_j$. Thus, SVD can be used for purpose of dimensional reduction. In this subspace, the distance between two subjects j_1 and j_2 becomes

$$d_{j_1 j_2} = |x_{j_1} - x_{j_2}| = \sqrt{\sum_{k=1}^{r} \left(v_{j_1 k}\sigma_k - v_{j_2 k}\sigma_k\right)^2}. \quad (7)$$

These distances in the subspace provide the basic measure for clustering the subjects. Additionally, the contribution of attributes i in the top left-singular vectors $\{\mathbf{u}_k\}$ is given by a weight

$$w_i = \sum_{k=1}^{r} \sigma_k |u_{ik}|. \quad (8)$$

The weight indicates the significance of the attribute i in the use of distinguishing all subjects.

On the other hand, the right-singular vectors $\{\mathbf{v}_k\}$ can be viewed as the eigenvectors spanning the new attribute subspace. The new coordinates of the attributes in this right-singular subspace reveal the relationship among the attributes. The original coordinates of attribute i are given by the row vector $\mathbf{a}_i$, which can be expressed as a linear combination of the right-singular vectors $\{\mathbf{v}_k\}$

$$a_i = \sum_{k=1}^{r} u_{ik}\sigma_k v_k. \quad (9)$$

Thus, the ith row of $\mathbf{U}\Sigma$, designated as $\mathbf{a}_i'$ gives the coordinate of the attribute i in the right-singular subspace $\{\mathbf{v}_k\}$. The attributes can be grouped together based on their pairwise distance in this subspace, similar to the approach above used in clustering subjects in the left-singular subspace.

Here, SVD analysis is used to classify β-lactamases by examining their dynamics profiles (i.e., dfi values) at different residue sites. The subjects of the study are the four β-lactamases and the attributes are the dfi values. To accommodate β-lactamases with slightly different length, we focus on the 262 residue sites where each β-lactamases has a residue present (i.e., not a gap) in multiple sequence alignment. Therefore in this application of SVD to β-lactamases, each column of $\mathbf{X}$, conventionally denoted as $\mathbf{x}_j$, is a 262-dimensional vector describing the dynamics profile at those residue sites of a given β-lactamases j $(1 \leq j \leq 4)$. We move the origin

to mean of the data by subtracting the mean of row i from each element x_{ij}. The resulting **X** matrix eliminates the generic characteristics of particular residue sites and emphasizes more clearly the differences among dfi patterns of those β-lactamases. As we wish to understand the relationship of the β-lactamases, the signal of interest in this case is the dynamics profile $\mathbf{x}_j$ of β-lactamases j. The dynamic profiles are transformed to the left-singular subspace through SVD, where the modern and ancestral β-lactamases are represented emphasizing their differences. The pairwise distance of β-lactamases in the subspace reveals their dynamics similarities and differences. Moreover, the residue sites with high weight and significant contribution to the top left-singular vectors may account for the major dynamic differences among those β-lactamases. The mutation occurred at those sites may have a large impact on the protein dynamics.

Acknowledgments

This work was supported by grants BIO2012-34937, CSD2009-00088, BIO2010-16800, 1U54GN0945999, and "Factoría Española de Cristalización," Consolider-Ingenio 2010 from Spanish Ministry of Economy and Competitiveness, FEDER Funds. The authors thank the ESRF staff in Grenoble (France) for excellent support during collection of the data required to obtain the structures shown in figure 1. S.B.O. and T.Z. acknowledge XSEDE computing time and CLAS/ASU award.

References

Allen HK, Moe LA, Rodbumrer J, Gaarder A, Handelsman J. 2009. Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J.* 3:243–251.

Atanasov BP, Mustafi D, Makinen MW. 2000. Protonation of the beta-lactam nitrogen is the trigger event in the catalytic action of class A beta-lactamases. *Proc Natl Acad Sci U S A.* 97:3160–3165.

Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J.* 80:505–515.

Atilgan C, Atilgan AR. 2009. Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. *PLoS Comput Biol.* 5:e1000544.

Atilgan C, Gerek ZN, Ozkan SB, Atilgan AR. 2010. Manipulation of conformational change in proteins by single-residue perturbations. *Biophys J.* 99:933–943.

Babtie A, Tokuriki N, Hollfelder F. 2010. What makes an enzyme promiscuous? *Curr Opin Chem Biol.* 14:200–207.

Bahar I, Lezon TR, Yang LW, Eyal E. 2010. Global dynamics of proteins: bridging between structure and function. *Annu Rev Biophys.* 39: 23–42.

Bar-Rogovsky H, Hugenmatter A, Tawfik DS. 2013. The evolutionary origins of detoxifying enzymes: the mammalian serum paraoxonases (PONs) relate to bacterial homoserine lactonases. *J Biol Chem.* 288: 23914–23927.

Berry M, Dumais S, O'Brien G. 1995. Using linear algebra for intelligent information retrieval. *SIAM Rev.* 37:573–595.

Boehr DD, Nussinov R, Wright PE. 2009. The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol.* 5: 789–796.

Bolia A, Gerek ZN, Keskin O, Banu Ozkan S, Dev KK. 2012. The binding affinities of proteins interacting with the PDZ domain of PICK1. *Proteins* 80:1393–1408.

Cantón R, Coque TM. 2006. The CTX-M beta-lactamase pandemic. *Curr Opin Microbiol.* 9:466–475.

Carnevale V, Raugei S, Micheletti C, Carloni P. 2006. Convergent dynamics in the protease enzymatic superfamily. *J Am Chem Soc.* 128: 9766–9772.

Changeux JP, Edelstein S. 2011. Conformational selection or induced fit? 50 years of debate resolved. *F1000 Biol Rep.* 3:19.

Chen CC, Smith TJ, Kapadia G, Wäsch S, Zawadzke LE, Coulson A, Herzberg O. 1996. Structure and kinetics of the beta-lactamase mutants S70A and K73H from *Staphylococcus aureus* PC1. *Biochemistry* 35:12251–12258.

Copley SD. 2003. Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr Opin Chem Biol.* 7:265–272.

Damblon C, Raquet X, Lian LY, Lamotte-Brasseur J, Fonze E, Charlier P, Roberts GC, Frère JM. 1996. The catalytic mechanism of beta-lactamases: NMR titration of an active-site lysine residue of the TEM-1 enzyme. *Proc Natl Acad Sci U S A.* 93:1747–1752.

D'Costa VM, King CE, Kalan L, Morar M, Sung WWL, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R, et al. 2011. Antibiotic resistance is ancient. *Nature* 477:457–461.

Deerwester S, Dumais ST, Furmas GW, Landauer T, Harshaman R. 1990. Indexing by latent semantic analysis. *J Am Soc Inf Sci.* 41: 391–407.

Delairesq M, Labiay R, Samamall J. 1992. Site-directed mutagenesis at the active site of *Escherichia coli* TEM-1 β-lactamase. *J Biol Chem.* 26: 20600–20606.

Doucet N, Savard PY, Pelletier JN, Gagne SM. 2007. NMR investigation of Tyr105 mutants in TEM-1 beta-lactamase: dynamics are correlated with function. *J Biol Chem.* 282:21448–21459.

Duarte F, Amrein BA, Kamerlin SC. 2013. Modeling catalytic promiscuity in the alkaline phosphatase superfamily. *Phys Chem Chem Phys.* 15: 11160–11177.

Durrant JD, De Oliveira CAF, McCammon JA. 2011. POVME: an algorithm for measuring binding-pocket volumes. *J Mol Graph Model.* 29:773–776.

Echave J. 2013. On the evolutionary conservation of protein dynamics: comment on "Comparing proteins by their internal dynamics: exploring structure-function relationships beyond static structural alignments" by Cristian Micheletti. *Phys Life Rev.* 10:31–32; discussion 9-40.

Echave J, Fernandez FM. 2010. A perturbative view of protein structural variation. *Proteins* 78:173–180.

Farrell DW, Speranskiy K, Thorpe MF. 2010. Generating stereochemically acceptable protein pathways. *Proteins* 78:2908–2921.

Finnigan GC, Hanson-Smith V, Stevens TH, Thornton JW. 2012. Evolution of increased complexity in a molecular machine. *Nature* 481:360–364.

Fornili A, Pandini A, Lu HC, Fraternali F. 2013. Specialized dynamical properties of promiscuous residues revealed by simulated conformational ensembles. *J Chem Theory Comput.* 9:5127–5147.

Garcia-Seisdedos H, Ibarra-Molero B, Sanchez-Ruiz JM. 2012. Probing the mutational interplay between primary and promiscuous protein functions: a computational-experimental approach. *PLoS Comput Biol.* 8:e1002558.

Gerek ZN, Keskin O, Ozkan SB. 2009. Identification of specificity and promiscuity of PDZ domain interactions through their dynamic behavior. *Proteins* 77:796–811.

Gerek ZN, Kumar S, Ozkan SB. 2013. Structural dynamics flexibility informs function and evolution at a proteome scale. *Evol Appl.* 6: 423–433.

Gerek ZN, Ozkan SB. 2010. A flexible docking scheme to explore the binding selectivity of PDZ domains. *Protein Sci.* 19:914–928.

Gerek ZN, Ozkan SB. 2011. Change in allosteric network affects binding affinities of PDZ domains: analysis through perturbation response scanning. *PLoS Comput Biol.* 7:e1002154.

Glembo TJ, Farrell DW, Gerek ZN, Thorpe MF, Ozkan SB. 2012. Collective dynamics differentiates functional divergence in protein evolution. *PLoS Comput Biol.* 8:e1002428.

Harms MJ, Thornton JW. 2010. Analyzing protein structure and function using ancestral gene reconstruction. *Curr Opin Struct Biol.* 20: 360–366.

Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, Kern D. 2007. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* 450:913–916.

Hobbs JK, Shepherd C, Saul DJ, Demetras NJ, Haaning S, Monk CR, Daniel RM, Arcus VL. 2012. On the origin and evolution of thermophily: reconstruction of functional Precambrian enzymes from ancestors of Bacillus. *Mol Biol Evol.* 29:825–835.

Hollup SM, Fuglebakk E, Taylor WR, Reuter N. 2011. Exploring the factors determining the dynamics of different protein folds. *Protein Sci.* 20:197–209.

Ikeguchi M, Ueno J, Sato M, Kidera A. 2005. Protein structural change upon ligand binding: linear response theory. *Phys Rev Lett.* 94: 078102.

Ingles-Prieto A, Ibarra-Molero B, Delgado-Delgado A, Perez-Jimenez R, Fernandez JM, Gaucher EA, Sanchez-Ruiz JM, Gavira JA. 2013. Conservation of protein structure over four billion years. *Structure* 21:1690–1697.

Jensen RA. 1976. Enzyme recruitment in evolution of new function. *Annu Rev Microbiol.* 30:409–425.

Kar G, Keskin O, Gursoy A, Nussinov R. 2010. Allostery and population shift in drug discovery. *Curr Opin Pharmacol.* 10:715–722.

Kazlauskas RJ. 2005. Enhancing catalytic promiscuity for biocatalysis. *Curr Opin Chem Biol.* 9:195–201.

Keskin O, Bahar I, Jernigan RL, Beutler JA, Shoemaker RH, Sausville EA, Covell DG. 2000. Characterization of anticancer agents by their growth inhibitory activity and relationships to mechanism of action and structure. *Anticancer Drug Des.* 15:79–98.

Khersonsky O, Roodveldt C, Tawfik DS. 2006. Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr Opin Chem Biol.* 10:498–508.

Khersonsky O, Tawfik DS. 2010. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem.* 79:471–505.

Kluger Y, Basri R, Chang JT, Gerstein M. 2003. Spectral biclustering of microarray data: coclustering genes and conditions. *Genome Res.* 13: 703–716.

Kratzer JT, Lanaspa MA, Murphy MN, Cicerchi C, Graves CL, Tipton PA, Ortlund EA, Johnson RJ, Gaucher EA. 2014. Evolutionary history and metabolic insights of ancient mammalian uricases. *Proc Natl Acad Sci U S A.* 111:3763–3768.

Lamotte-Brasseur J, Dive G, Dideberg O, Charlier P, Frère JM, Ghuysen JM. 1991. Mechanism of acyl transfer by the class A serine beta-lactamase of *Streptomyces albus* G. *Biochem J.* 279:213–221.

Leone V, Marinelli F, Carloni P, Parrinello M. 2010. Targeting biomolecular flexibility with metadynamics. *Curr Opin Struct Biol.* 20:148–154.

Levy SB, Marshall B. 2004. Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med.* 10:S122–S129.

Liberles DA, Teichmann SA, Bahar I, Bastolla U, Bloom J, Bornberg-Bauer E, Colwell LJ, de Koning AP, Dokholyan NV, Echave J, et al. 2012. The interface of protein structure, protein biophysics, and molecular evolution. *Protein Sci.* 21:769–785.

Livermore DM. 1995. beta-Lactamases in laboratory and clinical resistance. *Clin Microbiol Rev.* 8:557–884.

Lopez-Canut V, Roca M, Bertran J, Moliner V, Tunon I. 2011. Promiscuity in alkaline phosphatase superfamily. Unraveling evolution through molecular simulations. *J Am Chem Soc.* 133:12050–12062.

Maguid S, Fernandez-Alberti S, Echave J. 2008. Evolutionary conservation of protein vibrational dynamics. *Gene* 422:7–13.

Maguid S, Fernandez-Alberti S, Parisi G, Echave J. 2006. Evolutionary conservation of protein backbone flexibility. *J Mol Evol.* 63:448–457.

Micheletti C. 2013. Comparing proteins by their internal dynamics: exploring structure-function relationships beyond static structural alignments. *Phys Life Rev.* 10:1–26.

Munz M, Hein J, Biggin PC. 2012. The role of flexibility and conformational selection in the binding promiscuity of PDZ domains. *PLoS Comput Biol.* 8:e1002749.

Nobeli I, Favia AD, Thornton JM. 2009. Protein promiscuity and its implications for biotechnology. *Nat Biotechnol.* 27:157–167.

O'Brien PJ, Herschlag D. 1999. Catalytic promiscuity and the evolution of new enzymatic activities. *Chem Biol.* 6:R91–R105.

Onufriev A, Bashford D, Case DA. 2004. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* 55:383–394.

Pang A, Arinaminpathy Y, Sansom MS, Biggin PC. 2005. Comparative molecular dynamics—similar folds and similar motions? *Proteins* 61: 809–822.

Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, DeBolt S, Ferguson D, Seibel G, Kollman P. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput Phys Commun.* 91:1–41.

Pitout J, Laupland K. 2008. Extended-spectrum β-lactamase-producing Enterobacteriaceae: an emerging public-health concern. *Lancet Infect Dis.* 8:159–166.

Risso VA, Gavira JA, Gaucher EA, Sanchez-Ruiz JM. 2014. Phenotypic comparisons of consensus variants versus laboratory resurrections of Precambrian proteins. *Proteins* 82:887–896.

Risso VA, Gavira JA, Mejia-Carmona DF, Gaucher EA, Sanchez-Ruiz JM. 2013. Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β-lactamases. *J Am Chem Soc.* 135: 2899–2902.

Risso VA, Gavira JA, Sanchez-Ruiz JM. 2014. Thermostable and promiscuous Precambrian proteins. *Environ Microbiol.* 16: 1485–1489.

Roitberg AE, Okur A, Simmerling C. 2007. Coupling of replica exchange simulations to a non-Boltzmann structure reservoir. *J Phys Chem B.* 111:2415–2418.

Sikosek T, Chan HS. 2014. Biophysics of protein evolution and evolutionary protein biophysics. *J R Soc Interface.* 11:20140419.

Strynadka NC, Adachi H, Jensen SE, Johns K, Sielecki A, Betzel C, Sutoh KJM. 1992. Molecular structure of the acyl-enzyme intermediate in β-lactam hydrolysis at 1.7 Å resolution. *Nature* 359:700–705.

Sugita Y, Okamoto Y. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett.* 314:141–151.

Tobi D, Bahar I. 2005. Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proc Natl Acad Sci U S A.* 102:18908–18913.

Toth M, Smith C, Frase H. 2010. An antibiotic-resistance enzyme from a deep-sea bacterium. *J Am Chem Soc.* 132:816–823.

United States Centers for Disease Control and Prevention. 2013. Antibiotic resistance threats in the United States 2013, Washington (DC): CDC.

Vogt AD, Di Cera E. 2012. Conformational selection or induced fit? A critical appraisal of the kinetic mechanism. *Biochemistry* 51: 5894–5902.

Voordeckers K, Brown CA, Vanneste K, van der Zande E, Voet A, Maere S, Verstrepen KJ. 2012. Reconstruction of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovation through gene duplication. *PLoS Biol.* 10:e1001446.

Wells S, Menor S, Hespenheide B, Thorpe MF. 2005. Constrained geometric simulation of diffusive motion in proteins. *Phys Biol.* 2: S127–S136.

West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, Zuzan H, Olson JA, Marks JR, Nevins JR. 2001. Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci U S A.* 98:11462–11467.

Zhang W, Dourado DF, Fernandes PA, Ramos MJ, Mannervik B. 2012. Multidimensional epistasis and fitness landscapes in enzyme evolution. *Biochem J.* 445:39–46.