JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

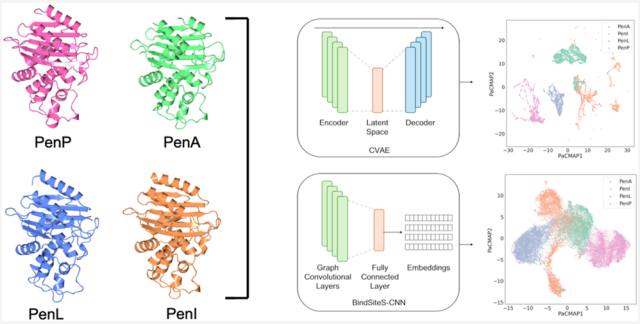
pubs.acs.org/jcim
Article

Evolutionary Dynamics and Functional Differences in Clinically Relevant Pen β-Lactamases from *Burkholderia* spp.

Jing Gu, Pratul K. Agarwal, Robert A. Bonomo, and Shozeb Haider*

Cite This: *J. Chem. Inf. Model.* 2025, 65, 5086−5098
Read Online

ABSTRACT: Antimicrobial resistance (AMR) is a global threat, with *Burkholderia* species contributing significantly to difficult-to-treat infections. The Pen family of β-lactamases are produced by all *Burkholderia* spp., and their mutation or overproduction leads to the resistance of β-lactam antibiotics. Here we investigate the dynamic differences among four Pen β-lactamases (PenA, PenI, PenL and PenP) using machine learning driven enhanced sampling molecular dynamics simulations, Markov State Models (MSMs), convolutional variational autoencoder-based deep learning (CVAE) and the BindSiteS-CNN model. In spite of sharing the same catalytic mechanisms, these enzymes exhibit distinct dynamic features due to low sequence identity, resulting in different substrate profiles and catalytic turnover. The BindSiteS-CNN model further reveals local active site dynamics, offering insights into the Pen β-lactamase evolutionary adaptation. Our findings reported here identify critical mutations and propose new hot spots affecting Pen β-lactamase flexibility and function, which can be used to fight emerging resistance in these enzymes.



INTRODUCTION

Antimicrobial resistance (AMR) is a leading cause of death worldwide.[1] The World Health Organisation highlighted that AMR was the direct reason for 1.27 million global deaths in 2019 and contributed indirectly to 4.95 million deaths.[2,3] Besides the high mortality rate, AMR is also estimated by the World Bank to result in the loss of up to US$ 3 trillion in gross domestic product (GDP) annually by 2030 and cost US$ 1 trillion financial burden to the healthcare by 2050.[4]

Burkholderia is a genus of Gram-negative bacteria that includes over 30 species of mammalian pathogens, some of which are clinically important to humans.[5,6] *Burkholderia cepacian* complex (BCC) and *Burkholderia gladioli* (*B. gladioli*) can infect individuals with cystic fibrosis (CF) and chronic granulomatous disease and cause difficult-to-treat chronic infections. *Burkholderia pseudomallei* (*B. pseudomallei*) is the etiologic agent of an often-fatal disease melioidosis, found in tropical and subtropical regions.[7−9] Outbreaks of BCC were reported in Hong Kong in 2020 and in the U.K. between 2023 to 2024.[10,11] In the US, as reported by CDC, *Burkholderia* related cases have been consistently reported between 2020 and 2024.[12,13]

β-lactam antibiotics such as Meropenem and ceftazidime are commonly recommended therapies for *Burkholderia* infections.[14] However, the potency of these β-lactam antibiotics is declining because of the expression of bacterial β-lactamases neutralize the effects of these drugs. All *Burkholderia* spp. can produce class A Pen β-lactamases, whose overproduction and mutations cause resistance to β-lactam antibiotics such as ceftazidime and amoxicillin-clavulanic acid.[5,8] PenA is a β-lactamase identified from the member of BCC named *Burkholderia multivorans*. It was identified to be a carbapenemase most similar to KPC-2.[6,15−18] PenI produced by *B. pseudomallei* is also referred to as the soluble form of PenA and was shown to possess Extended-spectrum β-lactamase (ESBL) properties.[7,16,19] PenL is a β-lactamase highly conserved in pathogenic *Burkholderia* spp. such as*B. pseudomallei*,*B. mallei*, and*B. cenocepacia* (called PenA in previous reports by H.S. Kim's group).[20−24] PenL from*B. thailandensis* strain E264 confers resistance to amoxicillin.[20] In *Burkholderia* spp., ceftazidime was also used for the selection of ceftazidime-resistant strains of *B. thailandensis* that express PenL.[25] PenP from *Bacillus licheniformis* is a narrow-spectrum class A β-lactamase that hydrolyses β-lactam antibiotics via the transient formation of an acyl-enzyme complex.[26,27]

Although all Pen β-lactamases share a common catalytic mechanism (Figure S1), their substrate profiles for a large number of clinically available antibiotics are very different (Table S1).[16] By acquiring single or multiple mutations at certain strategic "hot spots", these enzymes experience rapid

Received: February 11, 2025
Revised: April 22, 2025
Accepted: April 28, 2025
Published: May 2, 2025

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

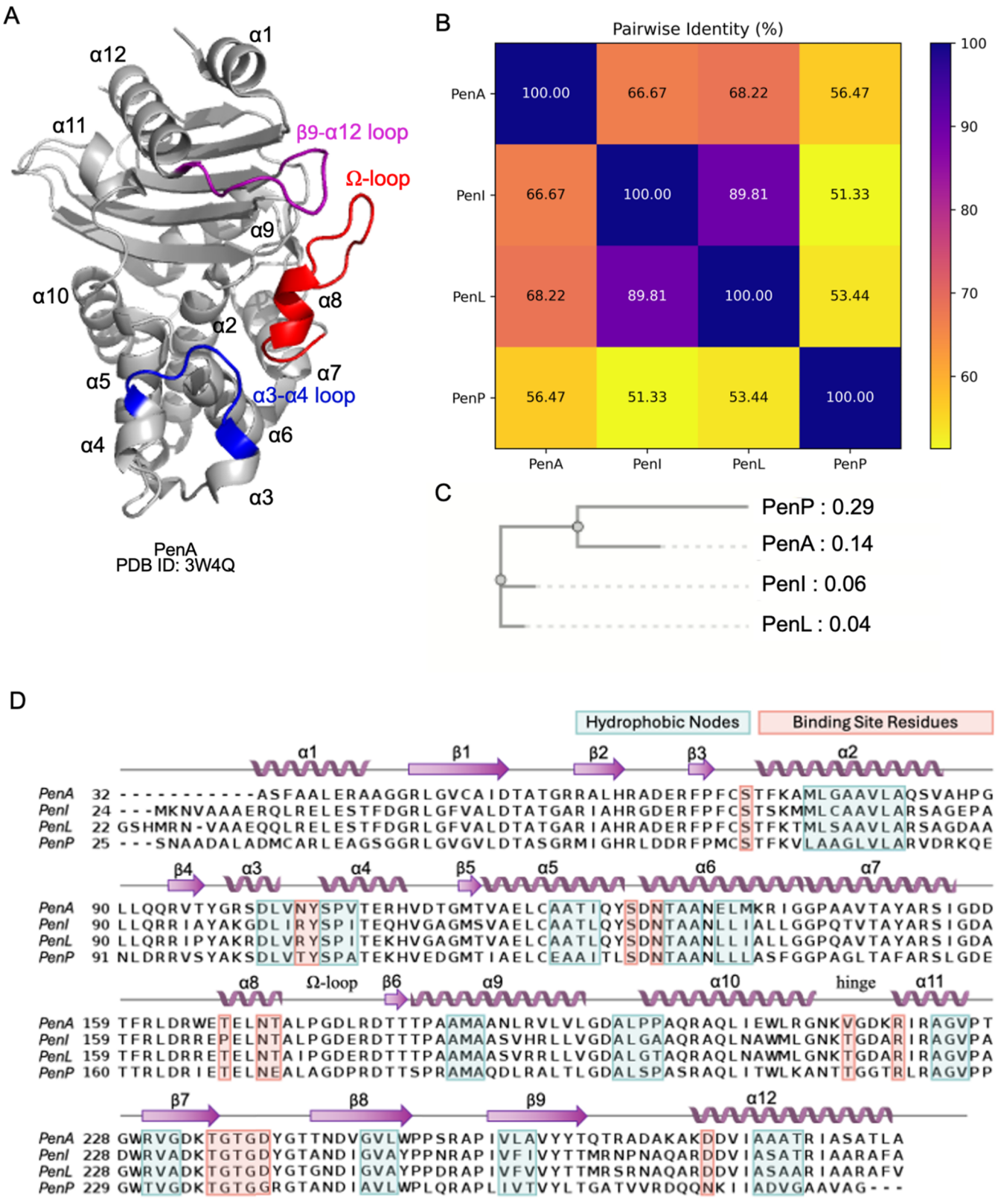


Figure 1. Structure and sequence of Pen β-lactamases. (A) Crystal structure of PenA β-lactamase (PDB ID: 3W4Q). (B) Pairwise sequence identity of the four Pen β-lactamases. (C) Phylogenetic tree of four Pen β-lactamases based on the multiple sequence alignment. The numbers indicate the branch lengths that represent the evolutionary distance between the nodes. (D) Multiple sequence alignments between PenA, PenI, PenL and PenP with secondary structure element annotations. The cyan boxes indicate the hydrophobic nodes, while orange boxes represent binding site residues.

molecular evolution over a period of several years, leading to drastic expansion of their substrate profiles toward new generations of antibiotics.[26] To reveal the indispensable structural and functional information related to such hot

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

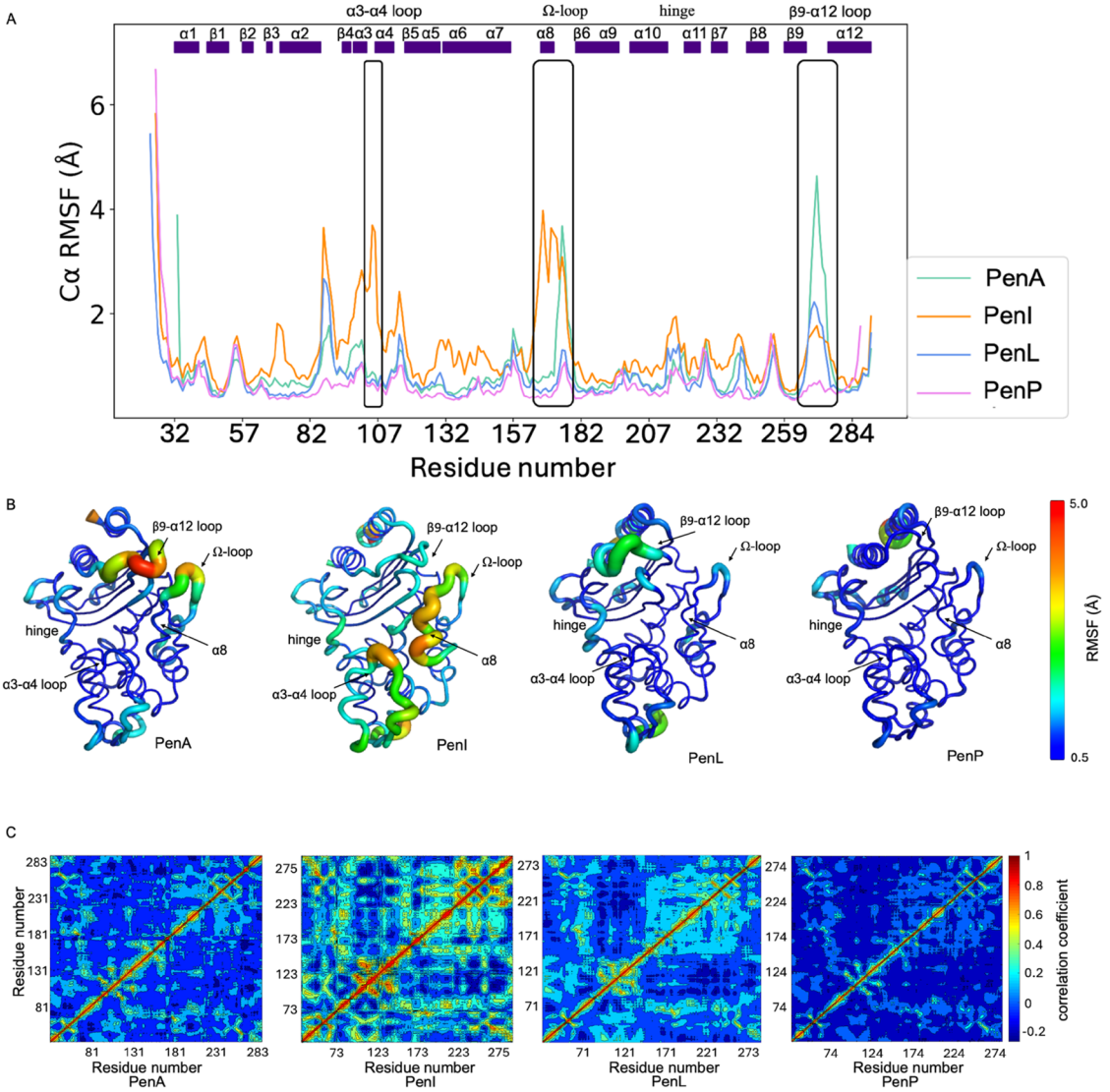


Figure 2. Conformational dynamics in Pen β-lactamases. (A) RMSF plots. (B) The putty structures of Pens. The putty structures are colored based on the RMSF with the range clipped into 0.5 to 5.0 Å. (C) Dynamic Cross Correlation Matrix plots.

spots, the sequence and structural differences between PenA (PDB ID: 3W4Q), PenI (PDB ID: 3W4P), PenL (PDB ID: 5GL9) and PenP (PDB ID: 6NIQ) were studied. The main focus was on investigating the dynamic differences between the four closely related Pen β-lactamases, using machine learning enabled molecular simulations, Markov State Models (MSMs), and deep learning. More specifically, this work aims to differentiate and categorize conformations through convolutional variational autoencoder-based deep learning (CVAE) and the trained BindSiteS-CNN model, integrating binding site local similarity information into the representation of global protein dynamic properties. The results presented here highlight key residues and sites on Pen β-lactamases that have a potential impact on the flexibility and dynamic stability of the structure, eventually leading to the evolutionary functional differences.

RESULTS AND DISCUSSION

Sequence and Structure Evolution. The Pen family of enzymes belong to class A β-lactamases according to the homology-based Ambler classification and have conserved secondary structures (Figure1A).[5] However, there is considerable variation in their sequence identities, ranging from 51.33 to 89.81% (Figure1B). PenI and PenL share the highest sequence identity at 89.81%, which is consistent with being functionally similar as Extended-spectrum β-lactamases (ESBL). PenP and the other three Pen β-lactamases do not share very high sequence identity. The phylogenetic analysis revealed a possible evolutionary pattern among the Pen β-lactamases (Figure1C). The constructed phylogenetic tree posits PenP and PenA to be evolutionary close to each other, suggesting a common ancestor between them. Nevertheless, Pen β-lactamases have conserved structural features like hydrophobic nodes and binding site residues, similar to that observed in all class A β-lactamases (Figure1D).

Structural and Dynamic Differences. The root-mean-square fluctuation (RMSF) and deviation (RMSD) is often used as a measure to highlight flexibility in protein dynamics.[28] There are considerable dynamic structural differences observed from RMSF and RMSD analysis between the four Pen β-lactamases (Figures 2 and S2). The MSF comparison between the Pen enzymes was generated, using PenA as the reference

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

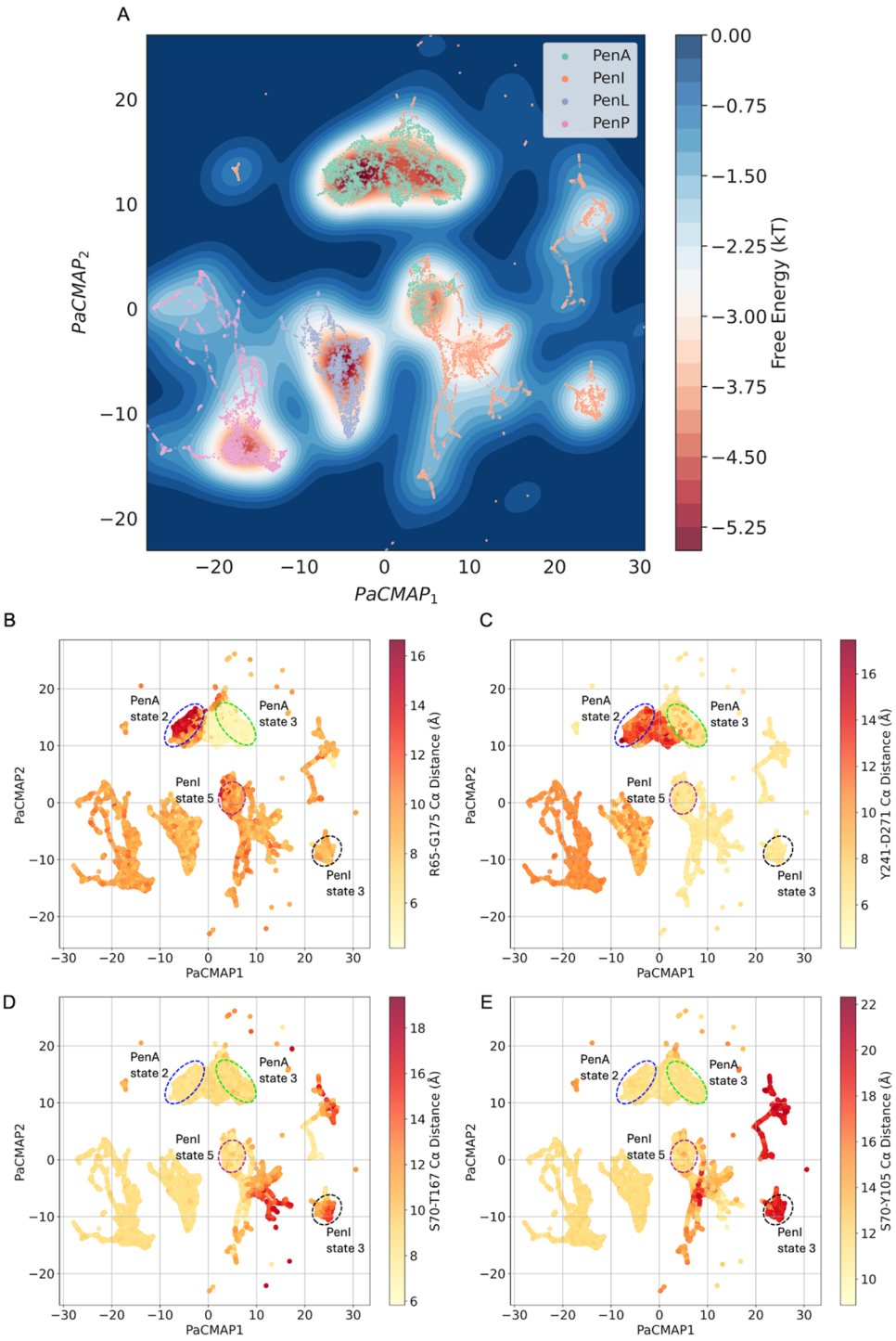


Figure 3. Convolutional variational autoencoder (CVAE)-based deep learning analysis. CVAE-learned features of high-dimensional data represented in 2D following PaCMAP treatment. (A) free energy landscape. (B) Cα distances of R65-G175 represent the conformation of the central section of the Ω-loop. (C) Cα distances of Y241-D271 represent the conformation of β9-α12 loop. (D) Cα distances of S70−P167 of α8-helix. (E) Cα distances of S70−Y105 of the α3-α4 loop.

structure (Figure 2A). The result has also been visualized using RMSF-colored putty plots to highlight regions of differential flexibility (Figure 2B). The conventional RMSD fitting approach, which uses all Cα atoms, is ineffective in distinguishing between regions of high and low mobility in β-lactamases. To differentiate these regions, we used MDLovoFit analysis to perform RMSD fitting using a fraction (φ) of Cα atoms. Beyond this fraction, there is a sharp rise in the RMSD value for the remaining Cα atoms.[29] Such analysis can be effectively used to observe the least (blue) and the most mobile atoms (red) (Figure S2). Different RMSD and the corresponding fraction of aligned atoms (φ) were used for the four Pen β-lactamases (PenA: RMSD = 0.9 Å, φ = 0.79; PenI: RMSD = 1.0 Å, φ = 0.67; PenL: RMSD = 0.7 Å, φ = 0.72; PenP: RMSD = 0.55 Å, φ = 0.70). The analysis identifies several common dynamic motifs including the α3-α4 loop, Ω-

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

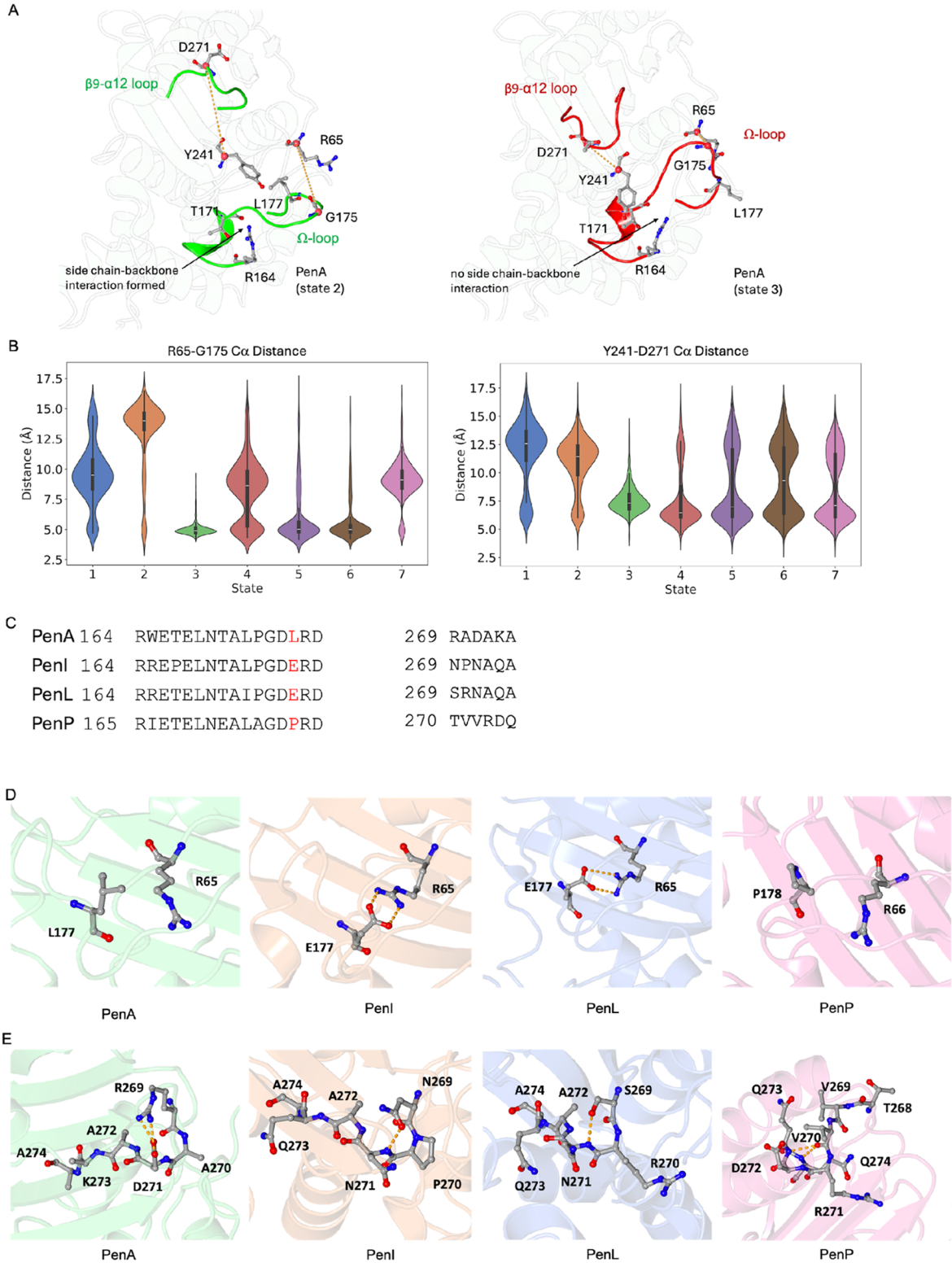


Figure 4. B9-α12 loop and Ω-loop in Pen β-lactamases. (A) The open conformation of the central section of the Ω-loop and the β9-α12 loop (state 2) and the closed conformation (state 3) in PenA. (B) Cα distance of R65-G175 and Y241-D271 in all PenA states. (C) Local sequence alignments of the Ω-loop and the β9-α12 loop of Pens. (D) The interactions of L177 in PenA, E177 in PenI and PenL, P178 in PenP. E177-R65 in PenI and PenL forms a salt-bridge to further stabilize the central section of the Ω-loop. This interaction is absent in PenA (L177) and PenP (P178) (E) The interactions in the β9-α12 loop of Pens.

loop (R164-D179), hinge region, α8-helix and the β9-α12 loop. Between the 4 enzymes, PenI is the most dynamic, while PenP is the most stable. The structure of PenA and PenI display similar dynamics with flexibility in the β9-α12 loop and the Ω-loop. While some flexibility is observed in α3-α4 loop in PenI, this is absent in PenA. To explore the correlated motions within the four Pen β-lactamases, we computed the dynamic cross-correlation maps (DCCMs) (Figures 2C and S3).[30] In

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

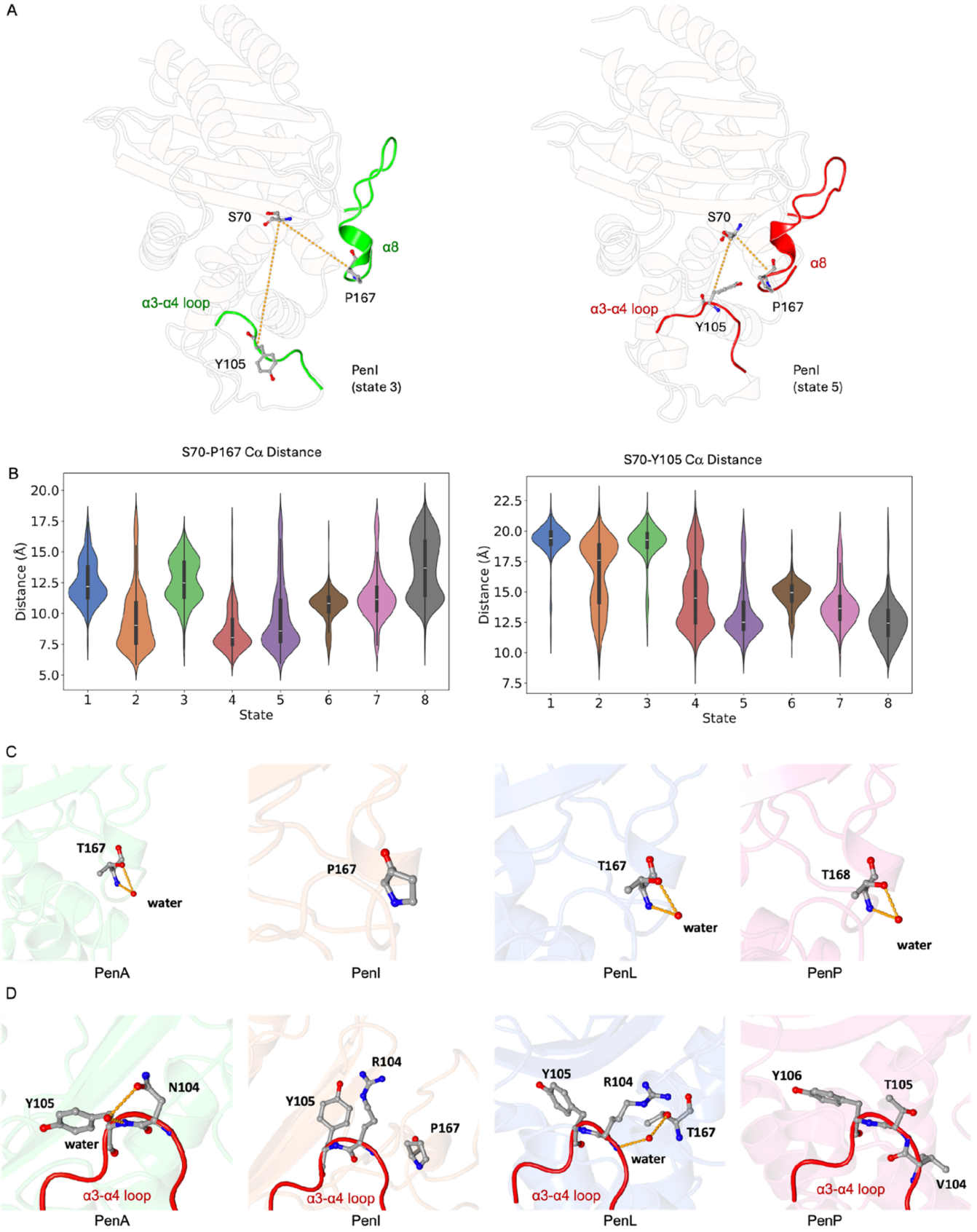


Figure 5. α8 helix and α3-α4 loop of Pen β-lactamases. (A) The open conformation of the central section of α8 helix and α3-α4 loop (state 3) and the closed conformation (state 5) of PenI. (B) Cα distance between S70−P167 and S70−Y105 in all PenI states. (C) T167 in PenA, P167 in PenI, T167 in PenL, T168 in PenP. When Thr is present, as in PenA, PenL and PenP, the secondary structure of the α8 helix is stable, and the χ1 angle of the Thr side chain is stabilized by a structural water molecule. In PenI (P167) this interaction cannot be formed. (D) The conformation and key residues of the α3-α4 loop (red) in all Pens.

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

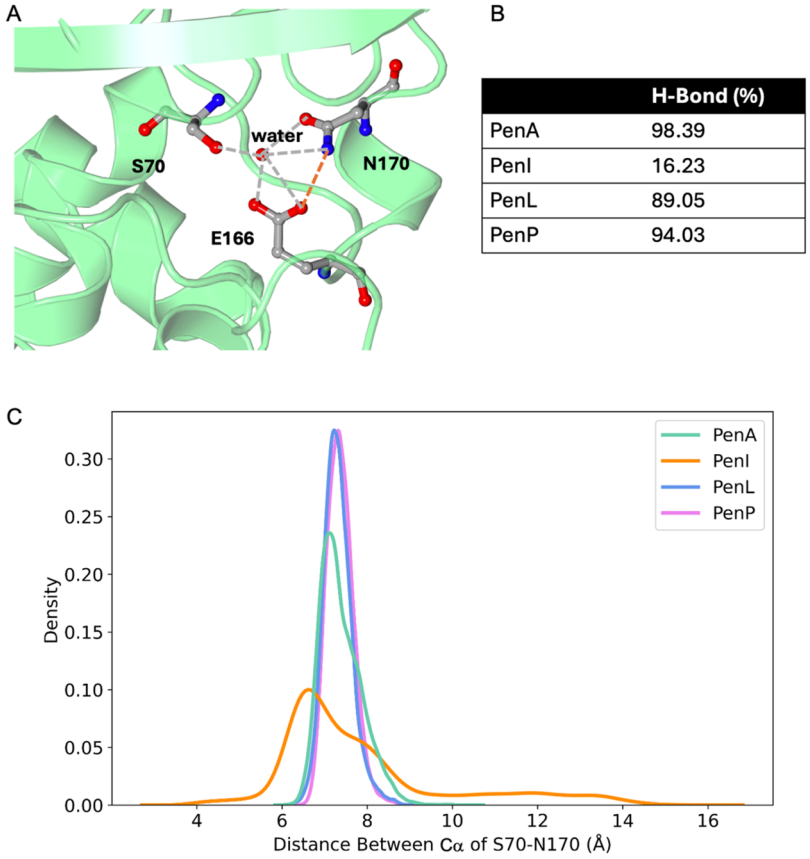


Figure 6. (A) Residues E166 and N170 of the Ω-loop position a water molecule used during acylation and deacylation steps. The E166-N170 H-bond is highlighted in orange, while other H-bonds within the network are shown in gray. (B) The E166-N170 H-Bond ratio during the MD simulation. (C) The distance between the Cα of S70 and N170.

these figures, the blue regions represent weak or slightly negative correlations, while light blue to green regions indicate moderate positive correlations. Positive correlations represented with red color imply residues moving in the same direction. There are more significant positive correlations identified in PenI (Figure S3), which suggest enhanced communication and cooperative motions across its structure while the least correlations in PenP reflect its stability observed.

To get a better understanding of the conformational differences, we further investigated the dynamics of these motifs in detail with MSMs and CVAE.

Markov State Models and Convolutional Variational Autoencoder. MSMs[31] of the four Pen β-lactamases were successfully built using the backbone torsions of all residues and the χ1 angle from the residues of the hydrophobic nodes, α3-α4 loop, β9-α12 loop and the Ω-loop as input features (Figures S4−S7). The structural results were collected from each PCCA distribution state. For each state, 1000 frames were saved as the further input of CVAE-based deep learning analysis.

To further compare significant differences in structural conformations, the unsupervised CVAE-based deep learning approach was used.[32] Distance matrices of Cα of the hydrophobic node, α3-α4 loop, β9-α12 loop and the Ω-loop residues were calculated from the saved frames of all states of the four Pen β-lactamases and stacked together as the input data of CVAE. During the training stage of the CVAE, parallel experiments with the third to the 30th latent dimension were run. The model selected for decoding was the one with the lowest loss, which was with the 28th latent dimension.

A free energy landscape was generated based on the 2D PaCMAP representation of the CVAE latent space (Figure 3A). The four Pen β-lactamases show different dynamics and can be clustered separately in different PaCMAP spaces while some conformations of PenA and PenI are clustered in the same space, which indicates that they share similar dynamics with these states. The "closed/open" conformations of Ω-loop and β9-α12 loop in PenA and the α8-helix and the α3-α4 loop in PenI are defined by calculating the representative pairwise Cα distances: R65-G175 (Figure 3B), Y241-D271 (Figure 3C) in PenA, S70−P167 (Figure 3D), S70−Y105 (Figure 3E) in PenI. Since there are mutants with different amino acids in Pens, the calculated distances are of the Cα pairs at the same position based on the structural alignment.

Referring to the distribution of states from the MSMs (Figure S8), state 2 of PenA represents the "open−open" conformation of the middle of the Ω-loop and the β9-α12 loop, while state 3 of PenA represents the "closed−closed" conformation of the Ω-loop and the β9-α12 loop (Figure 4A). State 3 of PenI represents the "open−open" conformation of the α8-helix and the α3-α4 loop, while state 5 of PenI represents the "closed−closed" conformation of the α8-helix and the α3-α4 loop (Figure 5A).

The central section of the Ω-loop (A172-R178) and the β9-α12 loop in PenA displays high flexibility with significant open and closed conformations (Figure 4A). The Cα distance between R65-G175 can be used to represent the open/closed

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

conformation of the Ω-loop, since R65 is a stable site. Similarly measuring the Cα distance between Y241-D271 highlight the dynamics of the β9-α12 loop (Figure 4B). The central section of the Ω-loop in PenA and PenI is more dynamic than that of PenL and PenP (A173-R179) (Figure 2B). L177 in PenA also contributes to the dynamics in PenA (Figure 4C,D). The absence of E177-R65 interaction (Figure 4D) in PenA allows for a more dynamic central section of the Ω-loop. In the open state in PenA, the side chain-backbone interaction between R164 and T171 is formed and further stabilizes the Ω-loop. Since the Ω-loop is reported to be very important to the binding of the substrates to β-lactamases, this extra flexibility might be one of the determinants of different substrate preferences of PenA. This hypothesis is supported by the findings of Papp-Wallace et al.,[16] who reported that PenA exhibits significantly higher efficiency than PenI across various β-lactam substrates. In contrast, Wong et al.[26] reported that engineered β-lactamases with enhanced Ω-loop flexibility can indeed lead to extended substrate profile.

Residue 269 positioned within the β9-α12 loop is poorly conserved between PenA (R269), PenI (N269), PenL (S269) and PenP (V270) enzymes (Figure 4C,E). In PenA, R269 forms an ion pair interaction with the side chain of D271 (Figure 4E). Similarly, the shorter side chains in PenI (N269) and PenL (S269) forms an H-bond interaction with the mainchain nitrogen of N271. In PenP, the structural arrangement of this loop is such that the side chain of T269 is unable to make any interactions. This structural arrangement instead allows the mainchain carbonyl oxygen of V270 to make two hydrogen bonds with the mainchain nitrogen atom of R272 and Q274 (Figure 4E). The residues to the α12-helix of PenP can form additional H-bond between the backbone of N270 and the backbone of D273 and Q274 (Figure 4E) that makes the α12-helix of PenP longer than that of PenA, PenI and PenL. This difference might be the reason for the higher stability of the β9-α12 loop in PenP.

The α8-helix is a short 3_{10} helix and just precedes the Ω-loop. The calculation of the Cα distance of S70−P167 can represent the conformation of the α8-helix since S70 is a stable site that can be used as the reference, while that of S70−Y105 can show the dynamic of the α3-α4 loop (Figure 5A,B). This helix in PenI is highly dynamic compared to that in PenA, PenL and PenP (Figure 2B). In PenI, a Proline is present at position 167, while this residue is a Threonine in PenA, PenL and PenP (Figure 5C). Proline is considered a potent breaker of α-helices and β-sheets.[33] The presence of P167 in PenI significantly destabilizes the α8-helix. The dynamic differences observed in the α3-α4 loop can also be explained by comparing PenI with other Pens (Figure 5D). In PenI, P167 is unable to make any interactions. However, T167 with a hydroxyl side chain can form interactions with the backbone of R104 via a stable bridging water molecule in PenL.

Residues E166 and N170 of the Ω-loop position a water molecule used during the acylation and deacylation of β-lactams (Figure 6A).[16,27] The presence of P167 in the α8-helix in PenI leads to a shift in the position N170 and subsequently results in an increase in the hydrogen bonding distance between E166 and N170. The hydrogen bond is formed only 16.2% of the simulation time (Figure 6B). Furthermore, the distance between N170 and the catalytic S70 also increases, thereby contributing to the overall increase in instability of the backbone in an indirect way (Figure 6C). The higher flexibility observed in PenI may explain the reason why PenI was also called the soluble form of PenA.[7] This might result in the lower catalytic efficiency of PenI due to the weakening of interactions within the hydrogen bonding network associated with locating key water molecules within the binding site. In particular, Papp−Wallace et al.[16] reported that in PenI, the side chain of E166 and N170 adopt alternative conformations at both pH7.5 and 9.5, leading to weakened hydrogen bonding to the diacylation water and a reduced occupancy compared to PenA. While substrate-bound simulations would be required to validate the mechanistic details, the current structural evidence provides a consistent explanation for the observed kinetic differences.

The mutations that have been observed in the Pen family of β-lactamases, are also consistent with the evolutionary trend observed in other class A β-lactamases. The P167T substitution has been reported to cause ceftazidime resistance in CTX-M-23, another class A ESBL.[34] With T167, PenA has a one dilution higher level of ceftazidime resistance than PenI, which possesses P167.[16] Furthermore, PenL containing T167 can be expressed in ceftazidime-resistant strains of *B. thailandensis*.[25] There is also evidence that PenP with T167 can bind with ceftazidime.[26] The P167G substitution included in a triple mutant of another class A β-lactamase TEM-1, W165Y/E166Y/P167G, or the laboratory mutant P167S have also been reported to alter the conformation of the active site and results in ceftazidime hydrolysis.[35,36] These all suggest that position 167 is a hot spot in Pen β-lactamases that may alter conformational dynamics and impact function.

P174 was previously identified as one of the conserved residues within class A β-lactamases.[37] The emergence of A175 in the corresponding position in PenP indicates another possibility of substitution at this position. The flexibility of the Ω-loop is closely related to substrate selectivity in class A β-lactamases, and since PenP is a narrow spectrum β-lactamase, we speculate that P174 plays a key role in expanding the substrate range of the enzyme.

Residues at position 104 within Pen β-lactamases (N104 in PenA, R104 in PenI and PenL, T105 in PenP) have also been identified at analogous position in other representative class A β-lactamases as a hot spot. The point mutation D104E identified in SHV-1 β-lactamase can stabilize a key binding loop in the interface of SHV-1 and β-lactamase inhibitor protein (BLIP) by increasing the volume of the side chain of E104.[38] In TEM-1, the substitution of residue 104 has also been reported to produce extended-spectrum resistance and be related to the level of ceftazidime resistance.[36]

Residue W105 was suggested to be important to the binding of β-lactams to KPC-2 and other class A β-lactamases.[39] Previous studies of TEM-1 have also identified Y105 as a substitution hot spot.[40] In L2 β-lactamases, the equivalent position is an H118Y substitution. This has been shown to affect the local environment within the active site.[41] Although Y105 is currently conserved in Pen β-lactamases, the orientation of Y105 has been reported to decrease the catalytic activity of PenI by obstructing the active site and preventing the β-lactam from binding.[16] Based on our hypothesis that class A β-lactamases might share a similar evolutionary trend, we can predict that position 105 could be a potential hot spot in Pen β-lactamases.

BindSiteS-CNN Based Binding Site Comparison. BindSiteS-CNN is a Spherical Convolutional Neural Network (S-CNN) model trained to provide analyses of similarities in proteins based on their local physicochemical properties of

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

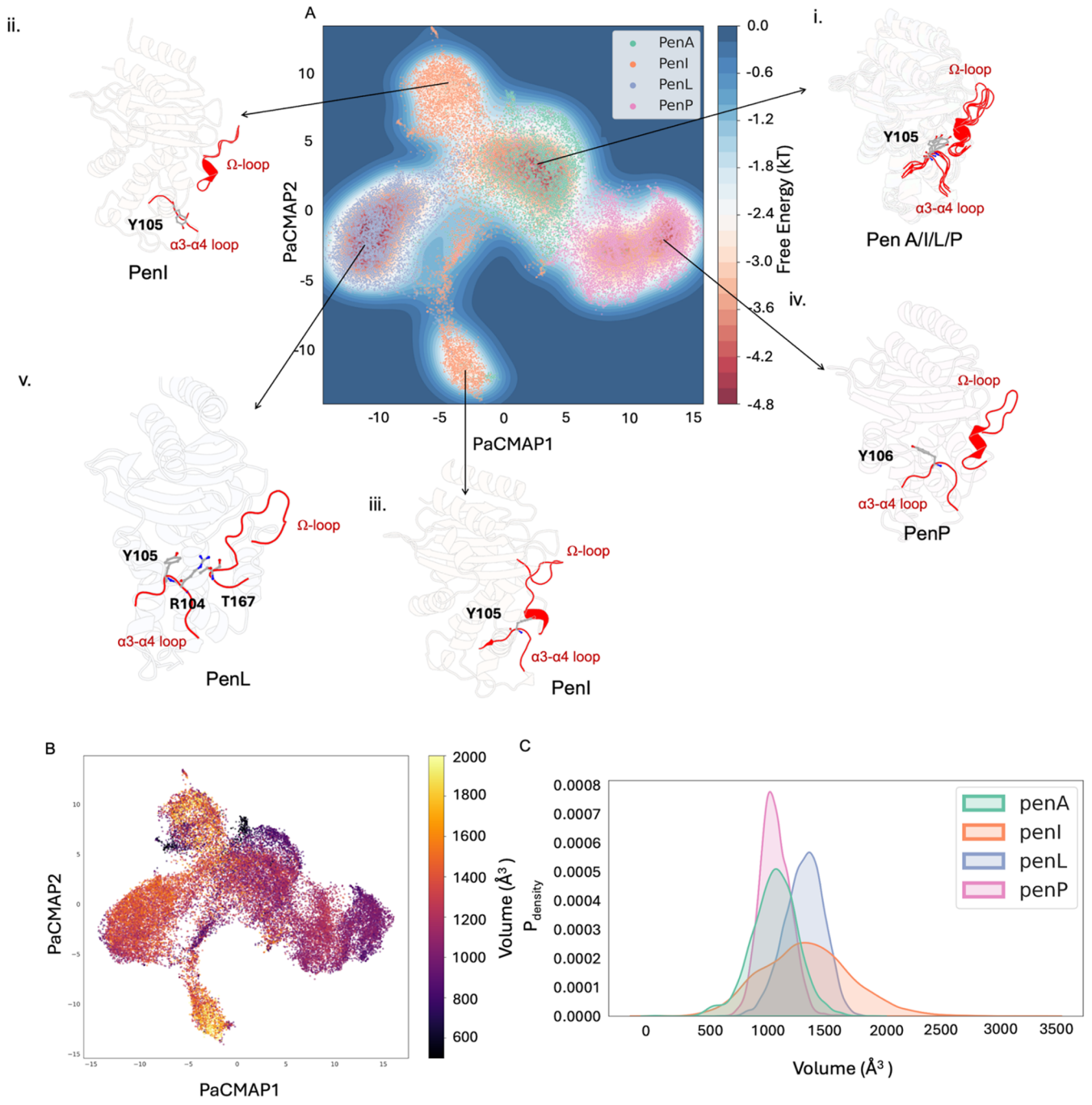


Figure 7. (A) BindSiteS-CNN-based high-dimensional embeddings represented in 2D with PaCMAP with free energy. The α3-α4 loop and the Ω-loop are highlighted in red. (i) Conformations of PenA, PenI, PenL and PenP extracted from the same energy basin. (ii) The representative conformation of PenI with α3-α4 loop in an outward extended conformation. (iii) The representative conformation of PenI with α8-helix segment of the Ω-loop moving away from the active site. (iv) The representative conformation of PenP (v) Conformation extracted from the main basin of PenL. (B) BindSiteS-CNN-based high-dimensional embeddings represented in 2D with PaCMAP colored by binding pocket volume. (C) The calculated binding pocket volume of Pen β-lactamases.

their binding sites.[42] It has been applied successfully to the study of L2 β-lactamases.[41] Our study is driven by the hypothesis that enzymes with similar structural features in the binding site will cluster together. The binding site residues were selected as highlighted in Figure 1D. Furthermore, the embeddings from the BindSiteS-CNN deep learning model can be visualized in 2D with PaCMAP.[43] The input frames were taken from every 24th frame of the 298 trajectories in each system (PenA, PenI, PenL, and PenP).

Conformations of PenA, PenI, PenL and PenP extracted from the same energy basin with the most common area of overlap display high structural similarity (Figure7(i)). This can also be considered as the most representative cluster of PenA. The following comparisons are referred to this set of stable conformations. PenI is most flexible around the active site within the studied family of Pen β-lactamases. There are two significant clusters of PenI (Figure7ii,iii). In the first cluster (CL1), the α3-α4 loop orients away from the active site,

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

resulting in an enlarged active site (Figure7ii,B). The calculated volume of the binding site within the simulation is 1074.34 ± 204.16 Å^3 for PenA, 1353.24 ± 389.74 Å^3 for PenI, 1329.58 ± 170.27 Å^3 for PenL, and 1079.47 ± 131.60 Å^3 for PenP (Figure 7C). In cluster 2 (CL2), the α8-helix moves away from the active site. This also results in an increase in the volume of the active site. In both clusters, the local features of the active site changes. (Figure7ii,iii,B). The structure of the active site of PenP is stable and less flexible, resulting in a smaller binding site. In PenP, the side chain of Y106 orients toward the hinge region rather than the Ω-loop (Figure7(iv)). This might explain the narrow-spectrum activity of PenP. The exemplar conformation of PenL highlights the stable interactions between the side chain of T167 with the backbone of R104 via a bridging water molecule.

The PaCMAP of BindSiteS-CNN based embeddings highlight the differences between the local features of the active site in Pen β-lactamases. This analysis also reveals the local changes within the active site of Pen β-lactamases.

CONCLUSIONS

To explore the dynamic differences between Pen β-Lactamases (PenA, PenI, PenL and PenP), this study employed enhanced sampling MD simulations, MSMs, CVAE and the BindSiteS-CNN deep learning model. Although the members are from the same Pen β-Lactamase family, we nevertheless captured several significant dynamic differences between them. The β9-α12 loop and Ω-loop are more dynamic in PenA due to the occurrence of L177 and R269. The α3-α4 loop, the Ω-loop, and the α8-helix are flexible in PenI. The emergence of P167 in PenI significantly improves its flexibility and solubility by disrupting the secondary structure of α8-helix. This alters the interactions made with the conserved water molecule required in the deacylation step of catalysis. The high level of structural stability of PenP is also consistent with its narrow spectrum activity. The application of the BindSiteS-CNN model to compare the local active site dynamics of Pen β-Lactamases revealed the difference between all four simulated systems. The potential hot spots highlighted in this study may serve as a guide for further understanding of the biological functions and the evolutionary relationship of Pen β-Lactamases.

MATERIALS AND METHODS

Structural Models. The crystal structures of PenA (PDB ID 3W4Q),[16] PenI (PDB ID 3W4P),[16] PenL (PDB ID 5GL9),[24] and PenP (PDB ID 6NIQ) were downloaded from the Protein Data Bank.

Sequence Alignment and Phylogenetic Tree Generation. Amino acid sequences of PenA, PenI, PenL and PenP were derived from the structural files through systematic parsing of the respective PDB structure files through pdb2fasta functionality. A multiple sequence alignment was conducted utilizing Clustal Omega employing its default parameters.[44] The resulting alignment was subsequently visualized and interpreted with Jalview version 2.11.3.3 to furnish a graphical depiction of sequence congruities for improved comprehension.[45]

Systems Preparation and Adaptive Sampling MD Simulations. Molecular dynamics simulations of PenA, PenI, PenL and PenP were conducted with the following protocol. The initial system preparation utilized the PlayMolecule ProteinPrepare Web Application at pH 7.4.[46] All heteroatoms were excised from the PDB files. ProteinPrepare autonomously executed pK_a calculations and optimized hydrogen bonds, while simultaneously assigning charges and protonating the PDB file within the high-throughput molecular dynamics (HTMD) framework.[47] Charge assignments were based on the local environment of the protonated structure, optimizing its hydrogen-bonding network. Utilizing tleap from the Amber MD package,[48] input files containing detailed information about atoms, bonds, angles, dihedrals, and initial atom positions were generated. The Amberff14SB force field was applied,[49] and each system was solvated in TIP3P water model within a cubic box,[50] maintaining a minimum 10 Å distance from the nearest solute atom, and neutralized with Na^+ and Cl^- ions. Prepared systems were initially minimized through 1,000 iterations of steepest descent and subsequently equilibrated for 5 ns under NPT conditions at 1 atm, employing the ACEMD engine within the HTMD framework.[47,51] The temperature was steadily increased to 300 K with a time step of 4 fs, using rigid bonds, a 9 Å cutoff, and particle mesh Ewald summations for long-range electrostatics.[52,53] During equilibration, the protein backbone was restrained with a spring constant of 1 kcal mol^{-1}Å^{-2}, while the Berendsen barostat controlled pressure and velocities were based on the Boltzmann distribution.[54] The production phase, conducted in the NVT ensemble, used a Langevin thermostat with 0.1 ps damping and a hydrogen mass repartitioning scheme, allowing a 4 fs time step and recording trajectory frames every 0.1 ns, all culminating in a final, unrestrained production step to observe natural system dynamics. The simulations were run until a minimum of 298 trajectories were obtained, with each trajectory counting 600 frames and sampling a cumulative 17.88 μs for each system.

Markov State Models. Pyemma v2.5.12 was used to build the MSMs.[31] Backbone dihedral angles (φ and ψ) of all residues and the φ1 angle from the residues of the hydrophobic nodes, α3-α4 loop, β9-α12 loop and the Ω-loop of Pens were selected as the input features. The featurized trajectories were projected onto selected independent components (ICs) using TICA. The produced projections can show the maximal autocorrelation for a given lag time. The chosen ICs were then clustered into selected clusters using k-means. In this way, each IC was assigned to the nearest cluster center. The lag time was chosen to build the MSM with metastable states according to the implied time scales (ITS) plot. After passing the Chapman−Kolmogorov (CK) test within confidence intervals, the MSM was defined as good. This indicates the model highly agrees with the input data, and it is statistically significant for use. Bayesian MSM was used to build the final model in the system. The net flux pathways between microstates, starting from state 1, were calculated using Transition Path Theory (TPT) function. The pathways all originate from state 1, as it shows the lowest stationary probability (the highest free energy) in the system. This is why state 1 is a reasonable starting point to illustrate all the relevant kinetic transitions through the full FE landscape. The structural results were selected from each Perron cluster-cluster analysis (PCCA) distribution.

Deep Conformational Clustering Using CVAE. The utilization of CVAE was implemented in a systematic manner to reveal the dynamic difference of Pen β-lactamases. Data derived from each protein system such as the pairwise distances were computed using MDAnalysis and MDTraj.[55,56] Detailed analyses involved the formulation of pairwise distance

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

maps extracted from all states of the MSM of each system. The focus was directed toward hydrophobic nodes, α3-α4 loop, β9-α12 loop and the Ω-loop residues featuring the constraint of Cα atom distances $\leq$ 8 Å, which were recorded as nonzero values in a specified three-dimensional matrix. The cumulative data, represented as 74 × 74 distance matrices, were consolidated into a unified 3D matrix for every system, accompanied by a label file containing pertinent metadata.

The CVAE's encoding section was structured with an 80:20 validation ratio and underwent training across 100 epochs. Dimensions spanning from 3 to 30 were explored, eventually settling on the 28th dimension, which exhibited the minimal loss, for the model's architecture. Continuous oversight was maintained for potential overfitting. For the decoding component, matrices and label files derived from Pens were used to assess the model's performance and to discern the clustering patterns of the conformations inherent to these systems.

The PaCMAP algorithm was then employed to reduce the dimensionality of the decoded embedding into two dimensions, thereby simplifying visualization.[57] This, when combined with the free energy landscape, proved instrumental in isolating distinctive conformations from energetically favorable regions.

BindSiteS-CNN Based Binding Site Comparisons. BindSiteS-CNN was employed to capture the differences between the active site local features of the four systems: PenA, PenI, PenL and PenP. The methodology encompassed binding pocket surface preparation and BindSiteS-CNN model processing and was adopted from Scott et al.[42] The samples were taken every 24th frame of the 298 trajectories in each system.

In the binding pocket surface preparation phase, the binding pocket surface of each frame was generated with side chain atoms of the binding site residues as the filtering reference. The computed pocket surface meshes with vertices enriched with physicochemical information describing the hydrophobicity, electrostatic potential and interaction-based classification of surface-exposed atoms lining the pocket were saved as PLY files and integrated as part of an in-house β-lactamases active pocket database.

During the BindSiteS-CNN model processing stage, the prepared 3D pocket mesh objects were fed into the trained BindSiteS-CNN model as input data. The embeddings of all input frames from the BindSiteS-CNN model as output data were saved out with labels into one pkl file. PaCMAP[57] has been used to visualize their distribution in the descriptor space by downgrading high-dimensional embeddings to 2D, those dots represent similar binding sites would cluster together.

Structural Analysis. The trajectories of the molecular simulations were meticulously aligned to their corresponding structures with MDAnalysis[55] and MDTraj.[56] The stride of frames within these trajectories was retrieved using the identical set of tools. To elucidate the general dynamics features inherent in the trajectories, calculations were performed again leveraging the functions within the MDTraj and MDAnalysis packages and MDLovofit.[29] For a more visual and intuitive understanding, the trajectories were loaded into the PyMOL Molecular Graphics System (http://www.pymol.org). This tool also facilitated the superimposing of structures and enabled a comprehensive conformational comparison. After delineating the spatial variations between distinct conformations, visual representations were generated via the Protein Imager.[58] Additionally, the Matplotlib package[59] in Python was employed for all statistical and graphical representations, including plots and figures, to present the data in a comprehensive and interpretable manner.

■ ASSOCIATED CONTENT

Data Availability Statement

All simulation data can be downloaded from the DOI 10.5281/zenodo.14843919.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.5c00271.

The catalytic mechanism (Figure S1); kinetic data (Table S1); Core Cα RMSD (Figure S2); Dynamic Cross-Correlation Maps (Figure S3); MSMs of Pen β-lactamases (Figure S4−S7); and CVAE based deep learning analysis of each Pen β-lactamase (Figure S8) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Shozeb Haider − *UCL School of Pharmacy, University College London, London WC1N 1AX, U.K.; University of Tabuk (PFSCBR), Tabuk 47512, Saudi Arabia; UCL Center for Advanced Research Computing, University College London, London WC1H 9RL, U.K.;* orcid.org/0000-0003-2650-2925; Email: Shozeb.haider@ucl.ac.uk

Authors

Jing Gu − *UCL School of Pharmacy, University College London, London WC1N 1AX, U.K.;* orcid.org/0009-0003-1172-9454

Pratul K. Agarwal − *High-Performance Computing Center, Oklahoma State University, Stillwater, Oklahoma 74078-1010, United States;* orcid.org/0000-0002-3848-9492

Robert A. Bonomo − *Research Service, Department of Veterans Affairs Medical Center and Clinician Scientist Investigator, Department of Veterans Affairs Medical Center, Louis Stokes Cleveland, Cleveland, Ohio 44106, United States; Department of Molecular Biology and Microbiology and Department of Medicine, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, United States; Departments of Pharmacology, Biochemistry, and Proteomics and Bioinformatics, CaseWestern Reserve University School of Medicine,, Cleveland, Ohio 44106, United States; CWRU-Cleveland VAMC Centerfor Antimicrobial Resistance and Epidemiology (Case VA CARES), Cleveland, Ohio 44106, United States;* orcid.org/0000-0002-3299-894X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.5c00271

Author Contributions

J.G.: Formal analysis, validation, investigation, visualization, writing—original draft; P.K.A.: Methodology, investigation; Robert.A.B.: Conceptualization, supervision, writing—review and editing; S.H.: Conceptualization, methodology, supervision, validation, writing—review and editing.

Notes

The authors declare the following competing financial interest(s): RAB reports grants from Entasis, Merck, Wockhardt, Shionogi, and Venatorx outside the submitted work. All the other authors have no conflict of interest.

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

REFERENCES

(1) Ikuta, K. S.; Sharara, F.; et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **2022**, *399* (10325), 629−655.

(2) Murray, C. J. L.; Ikuta, K. S.; Sharara, F.; Swetschinski, L.; Aguilar, G. R.; Gray, A.; Han, C.; Bisignano, C.; Rao, P.; Wool, E.; et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **2022**, *399* (10325), 629−655.

(3) Antimicrobial Resistance, 2023. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance.

(4) Bank, W.. Drug-Resistant Infections: A Threat to Our Economic Future, 2017. https://www.worldbank.org/en/topic/health/publication/drug-resistant-infections-a-threat-to-our-economic-future.

(5) Papp-Wallace, K. M.; Shapiro, A. B.; Becka, S. A.; Zeiser, E. T.; LiPuma, J. J.; Lane, D. J.; Panchal, R. G.; Mueller, J. P.; O'Donnell, J. P.; Miller, A. A. In Vitro Antibacterial Activity and In Vivo Efficacy of Sulbactam-Durlobactam against Pathogenic Burkholderia Species. *Antimicrob. Agents Chemother.* **2021**, *65* (3), No. e01930-20.

(6) Becka, S. A.; Zeiser, E. T.; LiPuma, J. J.; Papp-Wallace, K. M. The Class A beta-Lactamase Produced by Burkholderia Species Compromises the Potency of Tebipenem against a Panel of Isolates from the United States. *Antibiotics* **2022**, *11* (5), No. 674.

(7) Randall, L. B.; Dobos, K.; Papp-Wallace, K. M.; Bonomo, R. A.; Schweizer, H. P. Membrane-Bound PenA β-Lactamase of Burkholderia pseudomallei. *Antimicrob. Agents Chemother.* **2016**, *60* (3), 1509−1514.

(8) Schweizer, H. P. Mechanisms of antibiotic resistance in Burkholderia pseudomallei: implications for treatment of melioidosis. *Future Microbiol.* **2012**, *7* (12), 1389−1399.

(9) Rholl, D. A.; Papp-Wallace, K. M.; Tomaras, A. P.; Vasil, M. L.; Bonomo, R. A.; Schweizer, H. P. Molecular Investigations of PenA-mediated beta-lactam Resistance in Burkholderia pseudomallei. *Front. Microbiol.* **2011**, *2*, No. 139.

(10) Cheng, L.-K.; Chau, S. K.; Chan, W. S.; Chen, J. H.; Wong, B. K.; Fung, K. S. An outbreak of Burkholderia cepacia complex exit site infection among peritoneal dialysis patients caused by contaminated spray dressing. *Infect. Prev. Pract.* **2024**, *6* (2), No. 100359.

(11) Agency, U. H. S. Health Protection Report Volume 18, Issue 3, News − 14 March 2024, 2024. https://www.gov.uk/government/publications/health-protection-report-volume-18-2024/hpr-volume-18-issue-3-news-14-march-2024.

(12) Deida, A. A. V.; Spicer, K.; McNamara, K. X.; et al. Burkholderia multivorans Infections Associated with Use of Ice and Water from Ice Machines for Patient Care Activities—Four Hospitals, California and Colorado, 2020−2024. *MMWR Morbidity and Mortality Weekly Report* **2024**, *73*, 883−887.

(13) Hudson, M. J.; P, S.; Mathers, A.; et al. Outbreak of Burkholderia stabilis Infections Associated with Contaminated Nonsterile, Multiuse Ultrasound Gel — 10 States, May−September 2021. *MMWR Morb. Mortal. Wkly. Rep.* **2022**, *71*, 1517−1521.

(14) Chirakul, S.; Norris, M. H.; Pagdepanichkit, S.; Somprasong, N.; Randall, L. B.; Shirley, J. F.; Borlee, B. R.; Lomovskaya, O.; Tuanyok, A.; Schweizer, H. P. Transcriptional and post-transcriptional regulation of PenA β-lactamase in acquired Burkholderia pseudomallei β-lactam resistance. *Sci. Rep.* **2018**, *8* (1), No. 10652.

(15) Trépanier, S.; Prince, A.; Huletsky, A. Characterization of the penA and penR genes of Burkholderia cepacia 249 which encode the chromosomal class A penicillinase and its LysR-type transcriptional regulator. *Antimicrob. Agents Chemother.* **1997**, *41* (11), 2399−2405.

(16) Papp-Wallace, K. M.; Taracila, M. A.; Gatta, J. A.; Ohuchi, N.; Bonomo, R. A.; Nukaga, M. Insights into beta-lactamases from Burkholderia species, two phylogenetically related yet distinct resistance determinants. *J. Biol. Chem.* **2013**, *288* (26), 19090−19102.

(17) Poirel, L.; Rodriguez-Martinez, J.-M.; PléSiat, P.; Nordmann, P. Naturally Occurring Class A ß-Lactamases from the Burkholderia cepacia Complex. *Antimicrob. Agents Chemother.* **2009**, *53* (3), 876−882.

(18) Becka, S. A.; Zeiser, E. T.; Marshall, S. H.; Gatta, J. A.; Nguyen, K.; Singh, I.; Greco, C.; Sutton, G. G.; Fouts, D. E.; LiPuma, J. J.; Papp-Wallace, K. M. Sequence heterogeneity of the PenA carbapenemase in clinical isolates of Burkholderia multivorans. *Diagn. Microbiol. Infect. Dis.* **2018**, *92* (3), 253−258.

(19) Papp-Wallace, K. M.; Becka, S. A.; Taracila, M. A.; Winkler, M. L.; Gatta, J. A.; Rholl, D. A.; Schweizer, H. P.; Bonomo, R. A. Exposing a β-Lactamase "Twist": the Mechanistic Basis for the High Level of Ceftazidime Resistance in the C69F Variant of the Burkholderia pseudomallei PenI β-Lactamase. *Antimicrob. Agents Chemother.* **2016**, *60* (2), 777−788.

(20) Yi, H.; Song, H.; Hwang, J.; Kim, K.; Nierman, W. C.; Kim, H. S. The Tandem Repeats Enabling Reversible Switching between the Two Phases of β-Lactamase Substrate Spectrum. *PLoS Genet.* **2014**, *10* (9), No. e1004640.

(21) Hwang, J.; Cho, K. H.; Song, H.; Yi, H.; Kim, H. S. Deletion mutations conferring substrate spectrum extension in the class A β-lactamase. *Antimicrob. Agents Chemother.* **2014**, *58* (10), 6265−6269.

(22) Yi, H.; Kim, K.; Cho, K. H.; Jung, O.; Kim, H. S. Substrate spectrum extension of PenA in Burkholderia thailandensis with a single amino acid deletion, Glu168del. *Antimicrob. Agents Chemother.* **2012**, *56* (7), 4005−4008.

(23) Yi, H.; Cho, K. H.; Cho, Y. S.; Kim, K.; Nierman, W. C.; Kim, H. S. Twelve positions in a β-lactamase that can expand its substrate spectrum with a single amino acid substitution. *PLoS One* **2012**, *7* (5), No. e37585.

(24) Yi, H.; Choi, J. M.; Hwang, J.; Prati, F.; Cao, T. P.; Lee, S. H.; Kim, H. S. High adaptability of the omega loop underlies the substrate-spectrum-extension evolution of a class A beta-lactamase, PenL. *Sci. Rep.* **2016**, *6*, No. 36527.

(25) Papp-Wallace, K. M.; Becka, S. A.; Taracila, M. A.; Zeiser, E. T.; Gatta, J. A.; LiPuma, J. J.; Bonomo, R. A. Exploring the Role of the Omega-Loop in the Evolution of Ceftazidime Resistance in the PenA beta-Lactamase from Burkholderia multivorans, an Important Cystic Fibrosis Pathogen. *Antimicrob. Agents Chemother.* **2017**, *61* (2), No. e01941-16.

(26) Wong, W. T.; Chan, K. C.; So, P. K.; Yap, H. K.; Chung, W. H.; Leung, Y. C.; Wong, K. Y.; Zhao, Y. Increased structural flexibility at the active site of a fluorophore-conjugated beta-lactamase distinctively impacts its binding toward diverse cephalosporin antibiotics. *J. Biol. Chem.* **2011**, *286* (36), 31771−31780.

(27) Au, H.-W.; Tsang, M.-W.; So, P.-K.; Wong, K.-Y.; Leung, Y.-C. Thermostable β-Lactamase Mutant with Its Active Site Conjugated with Fluorescein for Efficient β-Lactam Antibiotic Detection. *ACS Omega* **2019**, *4* (24), 20493−20502.

(28) Song, X.; Bao, L.; Feng, C.; Huang, Q.; Zhang, F.; Gao, X.; Han, R. Accurate Prediction of Protein Structural Flexibility by Deep Learning Integrating Intricate Atomic Structures and Cryo-EM Density Information. *Nat. Commun.* **2024**, *15* (1), No. 5538.

(29) Martínez, L. Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. *PLoS One* **2015**, *10* (3), No. e0119264.

(30) Yu, H.; Dalby, P. A. Chapter Two - A beginner's guide to molecular dynamics simulations and the identification of cross-correlation networks for enzyme engineering. In *Methods in Enzymology*; Tawfik, D. S., Ed.; Academic Press, 2020; Vol. *643*, pp 15−49.

(31) Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Pérez-Hernández, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J. H.; Noé, F. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **2015**, *11* (11), 5525−5542.

(32) Bhowmik, D.; Gao, S.; Young, M. T.; Ramanathan, A. Deep clustering of protein folding simulations. *BMC Bioinf.* **2018**, *19* (18), No. 484.

(33) Li, S. C.; Goto, N. K.; Williams, K. A.; Deber, C. M. Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (13), 6676−6681.

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098

(34) Sturenburg, E. A novel extended-spectrum -lactamase CTX-M-23 with a P167T substitution in the active-site omega loop associated with ceftazidime resistance. *J. Antimicrob. Chemother.* **2004**, *54* (2), 406−409.
(35) Stojanoski, V.; Chow, D.-C.; Hu, L.; Sankaran, B.; Gilbert, H. F.; Prasad, B. V. V.; Palzkill, T. A Triple Mutant in the Ω-loop of TEM-1 β-Lactamase Changes the Substrate Profile via a Large Conformational Change and an Altered General Base for Catalysis. *J. Biol. Chem.* **2015**, *290* (16), 10382−10394.
(36) Vakulenko, S.; Golemi, D. Mutant TEM β-Lactamase Producing Resistance to Ceftazidime, Ampicillins, and β-Lactamase Inhibitors. *Antimicrob. Agents Chemother.* **2002**, *46* (3), 646−653.
(37) Philippon, A.; Slama, P.; Dény, P.; Labia, R. A Structure-Based Classification of Class A β-Lactamases, a Broadly Diverse Family of Enzymes. *Clin. Microbiol. Rev.* **2016**, *29* (1), 29−57.
(38) Reynolds, K. A.; Thomson, J. M.; Corbett, K. D.; Bethel, C. R.; Berger, J. M.; Kirsch, J. F.; Bonomo, R. A.; Handel, T. M. Structural and Computational Characterization of the SHV-1 β-Lactamase-β-Lactamase Inhibitor Protein Interface. *J. Biol. Chem.* **2006**, *281* (36), 26745−26753.
(39) Papp-Wallace, K. M.; Taracila, M.; Wallace, C. J.; Hujer, K. M.; Bethel, C. R.; Hornick, J. M.; Bonomo, R. A. Elucidating the role of Trp105 in the KPC-2 β-lactamase. *Protein Sci.* **2010**, *19* (9), 1714−1727.
(40) Reichmann, D.; Cohen, M.; Abramovich, R.; Dym, O.; Lim, D.; Strynadka, N. C.; Schreiber, G. Binding hot spots in the TEM1-BLIP interface in light of its modular architecture. *J. Mol. Biol.* **2007**, *365* (3), 663−679.
(41) Zhu, Y.; Gu, J.; Zhao, Z.; Chan, A. W. E.; Mojica, M. F.; Hujer, A. M.; Bonomo, R. A.; Haider, S. Deciphering the Coevolutionary Dynamics of L2 β-Lactamases via Deep Learning. *J. Chem. Inf. Model* **2024**, *64* (9), 3706−3717.
(42) Scott, O. B.; Gu, J.; Chan, A. W. E. Classification of Protein-Binding Sites Using a Spherical Convolutional Neural Network. *J. Chem. Inf. Model.* **2022**, *62* (22), 5383−5396.
(43) Wang, Y.; Huang, H.; Rudin, C.; Shaposhnik, Y. Understanding how dimension reduction tools work: an empirical approach to deciphering t-SNE, UMAP, TriMap, and PaCMAP for data visualization. *J. Mach. Learn. Res.* **2021**, *22*, 1−73.
(44) Madeira, F.; Pearce, M.; Tivey, A. R. N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **2022**, *50* (W1), W276−W279.
(45) Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25* (9), 1189−1191.
(46) Martínez-Rosell, G.; Giorgino, T.; De Fabritiis, G. PlayMolecule ProteinPrepare: A Web Application for Protein Preparation for Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2017**, *57* (7), 1511−1516.
(47) Doerr, S.; Harvey, M. J.; Noe, F.; De Fabritiis, G. HTMD: High-Throughput Molecular Dynamics for Molecular Discovery. *J. Chem. Theory Comput.* **2016**, *12* (4), 1845−1852.
(48) Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V. W. D.; Forouzesh, N.; Giese, T. J.; Gotz, A. W.; Gohlke, H.; et al. AmberTools. *J. Chem. Inf. Model.* **2023**, *63* (20), 6183−6191.
(49) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696−3713.
(50) Mark, P.; Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **2001**, *105* (43), 9954−9960.
(51) Harvey, M. J.; Giupponi, G.; Fabritiis, G. D. ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* **2009**, *5* (6), 1632−1639.
(52) Cerutti, D. S.; Duke, R. E.; Darden, T. A.; Lybrand, T. P. Staggered Mesh Ewald: An extension of the Smooth Particle-Mesh Ewald method adding great versatility. *J. Chem. Theory Comput.* **2009**, *5* (9), 2322−2338.
(53) Wells, B. A.; Chaffee, A. L. Ewald Summation for Molecular Simulations. *J. Chem. Theory Comput.* **2015**, *11* (8), 3684−3695.
(54) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* **1999**, *20* (8), 786−798.
(55) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32* (10), 2319−2327.
(56) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernandez, C. X.; Schwantes, C. R.; Wang, L. P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109* (8), 1528−1532.
(57) Wang, Y.; Huang, H.; Rudin, C.; Shaposhnik, Y. Understanding how dimension reduction tools work: an empirical approach to deciphering t-SNE, UMAP, TriMAP, and PaCMAP for data visualization. *J. Mach. Learn. Res.* **2021**, *22* (201), 1−73.
(58) Tomasello, G.; Armenia, I.; Molla, G. The Protein Imager: a full-featured online molecular viewer interface with server-side HQ-rendering capabilities. *Bioinformatics* **2020**, *36* (9), 2909−2911.
(59) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90−95.

https://doi.org/10.1021/acs.jcim.5c00271
J. Chem. Inf. Model. 2025, 65, 5086−5098