



DOI: 10.1002/ijch.202000012

Molecular Dynamics Simulations of Antibiotic Ceftaroline at the Allosteric Site of Penicillin-Binding Protein 2a (PBP2a)

Ying-Chih Chiang,^{*[a]} Mabel T. Y. Wong,^[a] and Jonathan W. Essex^[a]

Abstract: Methicillin-resistant *Staphylococcus aureus* (MRSA) tolerates β -lactam antibiotics by carrying out cell wall synthesis with the transpeptidase Penicillin-binding protein 2a (PBP2a), which cannot be inhibited by β -lactams. It has been proposed that PBP2a's active site is protected by two loops to reduce the probability of it binding with β -lactams. Previous crystallographic studies suggested that this protected active site opens for reaction once a native substrate binds at an allosteric domain of PBP2a. This opening was proposed for the new β -lactam cefataroline's mechanism in

successfully treating MRSA infections, i.e. by it binding to the allosteric site, thereby opening the active site to inhibition. In this work, we investigate the binding of cefataroline at this proposed allosteric site using molecular dynamics simulations. Unstable binding was observed using the major force fields CHARMM36 and Amber ff14SB, and free energy calculations were unable to confirm a strong allosteric effect. Our study suggests that the allosteric effect induced by cefataroline is weak at best.

Keywords: molecular dynamics simulations · β -lactams · allosteric mechanism · PBP2a

Introduction

Ever since Sir Alexander Fleming discovered penicillin in 1928, β -lactams – the core structure for all penicillin-like antibiotics (Figure 1A) – have saved numerous lives and are still regularly used against bacteria. β -lactams target the transpeptidase enzyme involved in bacterial cell wall synthesis. This enzyme catalyses the cross-linkage between peptidoglycans, recognising the acyl-D-Ala-D-Ala moiety as its native substrate. Owing to the structural similarity between this moiety^[1] and β -lactams, the latter can bind to transpeptidase and covalently inhibit the active site's catalytic serine.^[2] Thus, penicillin and other β -lactams kill bacteria by sabotaging their cell wall synthesis. Owing to the favoured binding between transpeptidase and penicillin, this enzyme is also termed Penicillin-binding protein (PBP).

Unfortunately, following the introduction of β -lactams, bacteria soon evolved to overcome them. The Gram-positive bacteria *Staphylococcus aureus*, which causes skin infections and is a risk to those recovering after surgery, has been an important subject in a number of studies. As early as 1942, penicillin-resistant staphylococci were already appearing in hospitals, roughly the same time when penicillin was first used to treat disease.^[3] Soon it was identified that those strains produced an enzyme called penicillinase to break down penicillin.^[4] Penicillinase-type enzymes (so called β -lactamases) then quickly spread across various species, leading to a wave of penicillin resistance. Following this resistance, new drugs were developed. The resultant β -lactam, methicillin, was effective toward PBPs and could not be metabolised by penicillinase. Again, resistance to methicillin was reported shortly after its introduction,^[3] *Staphylococcus aureus* brought in a new transpeptidase – PBP2a – that did not interact with

methicillin and thus had unhindered cell wall synthesis. This is the now-infamous methicillin-resistant *Staphylococcus aureus* (MRSA). Not limited to methicillin, MRSA is in fact resistant to almost all β -lactams, leaving vancomycin as the major effective antibiotic for treatment.^[5] This pressure was partially relieved in 2010, when the new drug cefataroline (a fifth-generation β -lactam) was approved.

MRSA's PBP2a differs from the methicillin-susceptible PBP in that it has a slightly distorted active site (residues 402–408 and 594–603), such that the catalytic residue Ser403 is occluded.^[6] Additionally, the α 2- α 3 loop (residues 417–457) and β 3- β 4 loop (residues 602–612) cover the entrance of the active site. Consequently, the binding of a β -lactam to PBP2a is hindered. This mechanism accounted for resistance to methicillin, but did not justify why native activity remained unaffected. Subsequently, Mobashery *et al.* discovered an unexpected allosteric domain on PBP2a by crystallising it with various ligands, including different β -lactams and a synthetic peptidoglycan.^[7] The reported crystal structures disclosed three allosteric ligands: a synthetic peptidoglycan (pdb code 3zg5), a muramic acid (3zfz), and a cefataroline (3zfx). The muramic acid binds at the same location as the saccharide component of peptidoglycan, while the cefataroline binds roughly at the

[a] Y.-C. Chiang, M. T. Y. Wong, J. W. Essex

School of Chemistry, University of Southampton, Southampton, SO17 1BJ, United Kingdom
E-mail: Y-C.Chiang@soton.ac.uk

© 2020 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

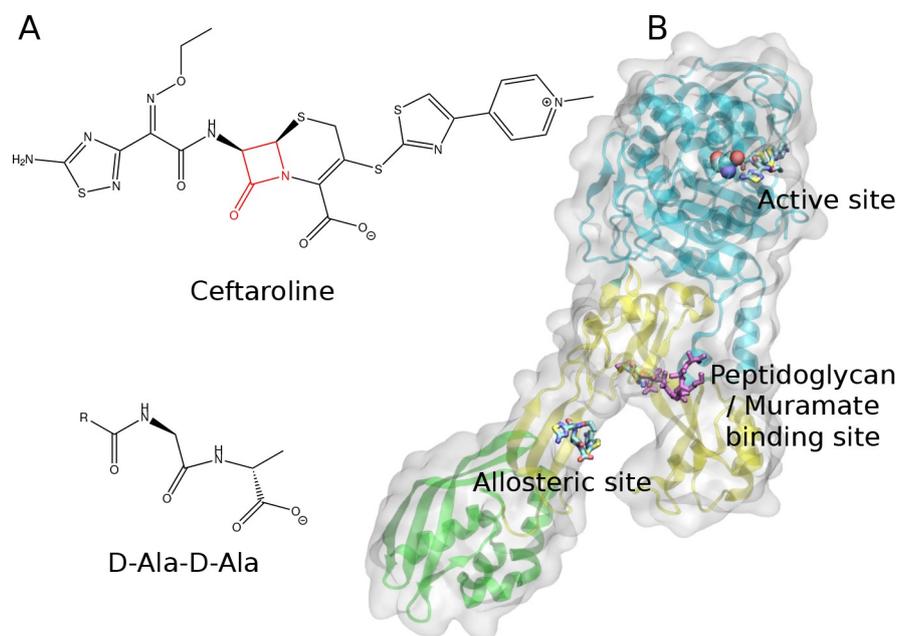


Figure 1. (A) Chemical structures of ceftaroline and the acyl-D-Ala-D-Ala moiety. The core structure of ceftaroline is a β -lactam (in red), which mimics PBP's native substrate D-Ala-D-Ala on a peptidoglycan. (B) PBP2a (pdb code 3zfv) binding with allosteric ligands muramate and ceftaroline. The active site is acylated by ceftaroline, and the catalytic residue Ser403 is depicted in spheres. The synthetic peptidoglycan (3zg5, in purple) also binds at the allosteric domain, where its glucoside moiety superimposes a muramate. The allosteric site of ceftaroline (in sticks) is identified by the ligand's binding location. The protein is coloured by domain: the N-terminal extension (residues 27–137) in green; the rest of allosteric domain (residues 138–326) in yellow; and the transpeptidase domain (residues 327–668) in blue.

predicted position of the nascent peptidoglycan's acyl-D-Ala-D-Ala moiety. The former (allosteric site for muramic acid) is about 46 Å away from the active site, and the latter (allosteric site for ceftaroline) is about 65 Å away, see Figure 1B for their locations on PBP2a. Furthermore, the active site in structure 3zfv is covalently inhibited by another ceftaroline. Based on this observation, an allosteric mechanism was proposed: once the allosteric site is occupied by a native substrate, PBP2a undergoes a conformational change, exposing the active site for reaction. Most importantly, the α_2 - α_3 loop moves away, allowing the substrate to enter the active site. Thus, only the native substrate (or ligands that closely resemble it) can trigger this allosteric mechanism for active site binding. The mechanism for conformational change was then further illustrated by the protein surface's reorganisation of its hydrogen bonding network upon ligand binding.^[8] While the allosteric mechanism was proposed for the native substrate, it was argued that ceftaroline could trigger the same mechanism. This was further supported by three key experiments: (1) the crystal structure revealed the allosteric site for ceftaroline (pdb codes 3zfv and 3zg0 for soaking and co-crystallising experiments respectively); (2) the dissociation constant of ceftaroline from the allosteric site can be measured by fluorescence quenching ($K_d = 20 \pm 4 \mu\text{M}$);^[9] and (3) mass spectroscopy revealed a synergy effect when mixing ceftaroline with an inactive β -lactam.^[10]

Since there are numerous experiments suggesting ceftaroline binding at PBP2a's allosteric domain, it may also be interesting to study this binding using molecular dynamics (MD) simulations. In this article, we report the results from our MD simulations of PBP2a with ceftaroline at both the allosteric and active sites, using both CHARMM36^[11,12] and Amber ff14SB^[13] force fields. Our results and discussion are given in the next section, followed by conclusions drawn from these simulations. Details of system construction, simulation protocols, and analysis methods employed, are all given in the methods section, and will not be mentioned elsewhere unless necessary. Finally, we briefly define the terminology used in this article. The term "allosteric site" will only refer to ceftaroline's allosteric binding site, while the binding site for synthesised peptidoglycan and muramic acid will simply be termed the muramate binding site. In situations when both allosteric and active sites are occupied by ceftaroline, the former will be addressed as the allosteric ceftaroline and the latter the orthosteric ceftaroline.

Results and Discussion

Since protein function normally depends on a few residues conserved across a protein family, the first step towards understanding the allosteric binding mechanism was sequence alignments. Here we present two results, with the first from

ConSurf.^[14,15] ConSurf takes a PDB's amino acid sequence and searches for homologues, identifying positions where the same (i. e. conserved) residues appear across a large number of these sequences. These conserved residues are usually crucial for the protein's function. The PDB 3zfz was uploaded to the web server as a template, which identified 544 non-redundant homologues, each with a minimal sequence identity of 35%, on the UniRef90 database using the HMMER algorithm.^[16,17] A Multiple Sequence Alignment (MSA) was then generated from 150 representative sequences using MAFFT.^[18] The Bayesian method was used to calculate the evolution rate of each residue. Figure 2 shows that ceftaroline's allosteric site lacks the clusters of highly conserved residues seen at the active and muramate binding sites. Furthermore, the two residues categorised as being most conserved in the allosteric site are Glycines, making them unlikely candidates for forming specific interactions with ceftaroline.

The second alignment was performed using the Clustal Omega MSA web server^[19] with nine different PDBs. Clustal Omega takes a set of sequences and carries out pairwise alignments, starting from the pair with the highest sequence identity. The employed PDB codes for the nine proteins (including PBP2a) are 1qme, 2wad, 3equ, 3oc2, 3vsk, 3zfz, 5e31, 5lp4, and 5u47. These are the non-repeating PBP isoforms from different species that were identified by searching PBP2a sequences in the Pfam database.^[20] The result allows identification of residues that remained unchanged across the set of provided sequences. Three groups of conserved residues were found among the nine proteins, as shown in Figure 3A. The first group of conserved residues are located around the active site (G402, S403, K406, S462,

G520, K597, G599, T600), suggesting that they are crucial for PBP2a function. Indeed, S403 is the catalytic residue for peptidoglycan cross-linkage, while the others could participate in anchoring the substrate. The second group of residues (R151, G152, R241, G279, E284, L291) are around the allosteric site for muramate and synthetic peptidoglycan, corresponding to the sugar tongue region within PBP's dimerisation domain (residues 146–313) as identified by Pfam.^[20] Identifying these conserved residues is no coincidence: experimental data state that muramate is a native allosteric ligand of PBP2a.^[7] Thus, those residues must be responsible for some biological function, which strongly supports the mechanism of allosteric control by the native substrate.^[7] As a remark, we mention that Pfam assigns the allosteric domain (residues 146–313) to a dimerisation domain. This is reasonable since PBP2a was frequently observed as a dimer in crystal structures, although *in vivo* it is often modelled as a monomer.^[8] The last group of conserved residues (D156, F211, G265, Q333, G355, L358, G489, Q530, G542) are scattered across the protein as shown in Figure 3A. While their involvement in allosteric signalling is unclear, these residues certainly do not support ceftaroline's binding at the allosteric site, as none of them are within 10 Å of the ligand.

The second step was to double check the binding sites using a quick prediction method; here we employed FTMAP.^[21,22] FTMAP performs a docking-like procedure using sixteen small ligands (probes) and the CHARMM force field with an implicit water model. Binding hot spots are elucidated through the consensus sites where multiple probes can dock preferably. Although FTMAP does not account for

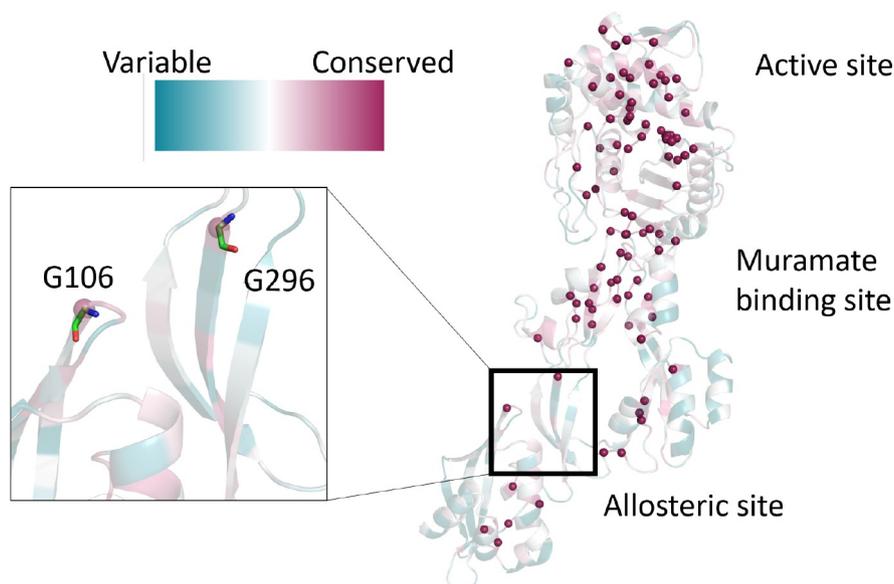


Figure 2. Conservation of *S. aureus* PBP2a residues, calculated using the ConSurf web server, accessed on 08/01/2020. The most highly conserved residues are shown as purple spheres. A close-up view of the reported allosteric site for ceftaroline is provided, with the most conserved residues labelled and shown in sticks.

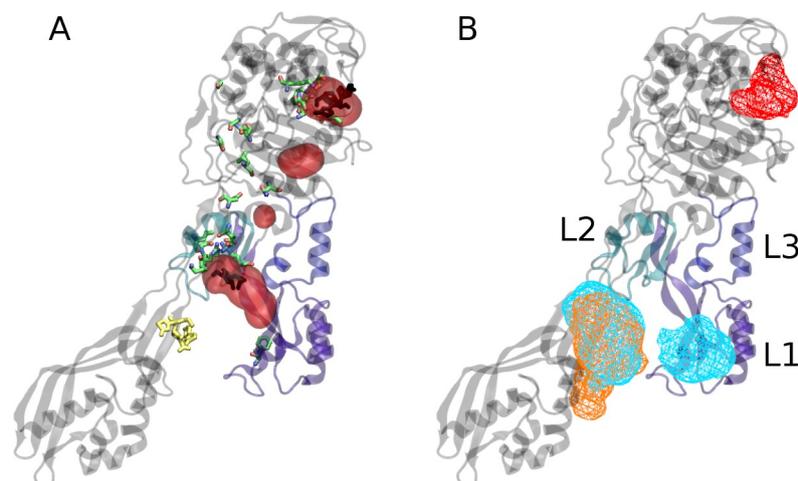


Figure 3. (A) Predicted binding hot spots and conserved residues on PBP2a. The binding hot spots (transparent red volume) were predicted by FTMAP.^[21,22] The two major sites coincide well with the active and muramate binding sites, where the ligands are coloured in black. Three groups of conserved residues are shown in sticks (green): two are around the binding sites, and one is scattered across the protein. Notably, the last group has no residues around the allosteric ceftaroline (yellow). (B) Ligand occupancy observed in three 100 ns MD simulations. The red, blue, and orange meshes are occupancies of ceftaroline in the active site (CHARMM36), the allosteric site (CHARMM36), and the allosteric site (Amber ff14SB) simulations respectively. The allosteric site ceftaroline effectively explored the whole allosteric domain (occupancy ≥ 0.01). Notably, ceftaroline wandered to lobe 1 (L1) and formed a stable complex, as suggested by the blue mesh. In contrast, binding at the active site seems to be more stable (occupancy ≥ 0.1) than the binding at the allosteric site.

protein flexibility, it can reliably predict openly accessible binding sites, including those of promiscuous proteins.^[23,24] Echoing observations from our sequence alignments, where no conserved residues were found near the allosteric ceftaroline, FTMAP predicts no binding sites in this particular area. In contrast, both muramate's binding site and the active site are clearly located by FTMAP – the predicted sites (red volume) coincide well with the associated ligands as shown in Figure 3A. Since neither of these two analyses suggest any specific residues for ceftaroline binding at the allosteric site, it becomes yet more interesting to understand how the drug triggers the allosteric regulation mechanism for active site binding.

Several MD simulations were conducted to further reveal ceftaroline's binding pattern at the allosteric site, using both CHARMM36 and Amber ff14SB force fields. Different force fields were used to test whether a consistent binding pattern could be found. In good agreement with FTMAP's prediction, the ligand's occupancy at the active site (see red mesh in Figure 3B) over three 100 ns CHARMM trajectories shows a localised binding. The red mesh encloses the volume that was occupied by ceftaroline for more than 10% of the simulation (occupancy ≥ 0.1). Also shown in Figure 3B are the occupancies of ceftaroline at the allosteric site, calculated from 3×100 ns trajectories simulated using CHARMM (blue mesh) and Amber (orange mesh) force fields. Both meshes label the volume with occupancy larger than 0.01. Unlike for the active site, the ligand at the allosteric site could move around the entire allosteric domain. In particular, the blue mesh demonstrates that ceftaroline fluctuated a large amount and did not

remain in one place for long. This observation is in line with no conserved residues being located around the allosteric ceftaroline. Consequently, the ligand may be loosely attached to PBP2a, forming an unstable complex through non-specific interactions such as van der Waals forces. This unstable binding of allosteric ceftaroline was generally observed across several sets of MD simulations. By measuring the end-to-end ligand distance (from methylpyridinium to the amine group, see Figure 1A), and the β -lactam's centre of mass to Glu145's backbone (C_{α}), one can characterise the “flexibility” of the allosteric binding. Simulations were performed using: a monomer complex with CHARMM force field (Figure 4A) where the allosteric site was occupied; a dimer complex with CHARMM force field (panel B) where the allosteric site was occupied; a monomer complex with Amber force field where both active and allosteric sites were occupied (panel C); and a dimer complex with Amber force field where the allosteric site was occupied (panel D). The original crystal structure conformation is marked by a yellow cross: the ligand is about 18 Å long and 7 Å away from Glu145's backbone. If one considers this an extended ligand conformation, most MD simulations explore a “folded” conformation where a smaller end-to-end distance is found. The difference between the two force fields is clear: the ligand described by CHARMM's associated CGenFF force field (with further optimisation on parameters with high penalty scores) finds more folded conformations, whereas the ligand described by Amber's counterpart GAFF (without optimisation) can explore extended conformations. Although both force fields lead to moderately different ceftaroline conformations, their MD simulations all

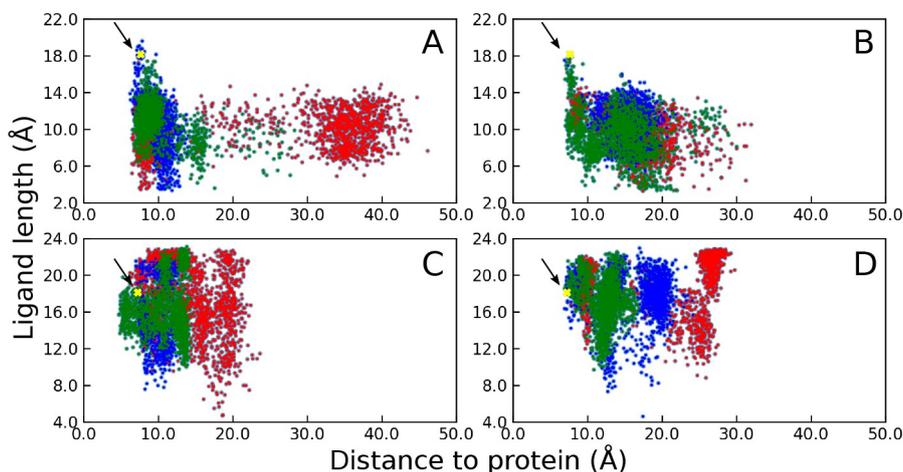


Figure 4. Distribution of allosteric ceftaroline conformations from three 100 ns simulations of PBP2a. (A) Monomer simulations using CHARMM36 force field. (B) Dimer simulations using CHARMM36 force field. (C) Monomer simulations with the active site occupied, using Amber ff14SB. (D) Dimer simulations with an empty active site using Amber ff14SB. Ligand length is the end-to-end distance between the methyl carbon at ceftaroline's methylpyridinium end and the nitrogen on the amine from its other end. The distance to protein refers to the distance between the β -lactam's centre of mass and Glu145's C α . The crystal structure conformation of ceftaroline is marked with a yellow cross that is highlighted by the black arrow. The widespread distributions show a general trend of unstable binding at the allosteric site.

reveal that the ligand can easily wander from its crystal structure allosteric binding site. For instance, the red distribution shown in Figure 4A corresponds to ceftaroline binding at a different area within the allosteric domain, viz. the blue mesh next to PBP2a's lobe L1 as depicted in Figure 3B.

We observed a large protein backbone motion in the MD simulations. These motions were captured by applying

principal component analysis (PCA) to the protein backbone conformations sampled in these trajectories. The three major backbone motions observed in the PBP2a monomer using the CHARMM36 force field are identified as in Figure 5A: (1) a bending mode, where the transpeptidase domain bends towards the N-terminus; (2) a wiggling mode, where the relative domain motion is perpendicular to the bending mode;

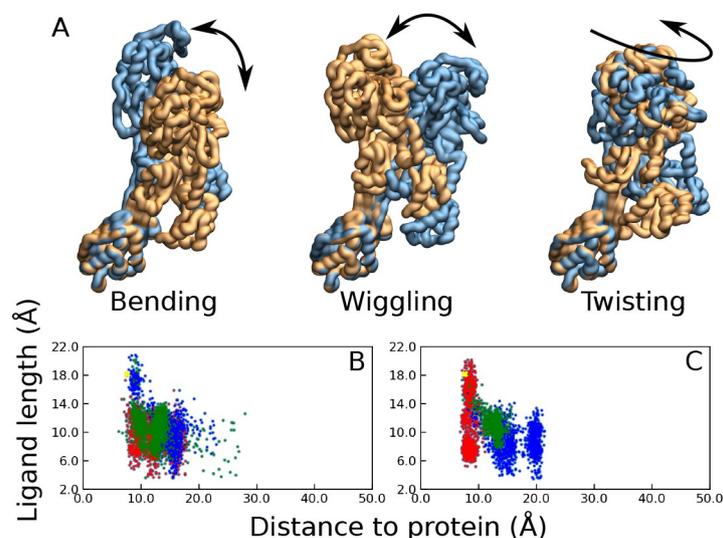


Figure 5. (A) PBP2a backbone motions revealed by principal component analysis of CHARMM36 force field monomer simulations. The colour decodes the extremes of the first three principal components, where the maximum is in blue and the minimum in orange. The bending mode, wiggling mode, and twisting mode account for 53.3%, 26.8%, and 14.1% of the variance in Cartesian coordinates respectively. (B) and (C) Distributions of allosteric ceftaroline conformations obtained from MD simulations using the CHARMM36 force field, where the protein backbone was restrained to the crystal structure conformation. The system was setup as a monomer (B) and dimer (C). Compared with the distributions shown in Figure 4, the ligand drifts less from its original position.

and (3) a twisting mode, where the transpeptidase domain rotates along the principal axis of PBP2a. Similar modes were found in other MD simulations. These modes, in particular the bending one, could disrupt ceftaroline's binding at the solvent-exposed allosteric site. This is congruent with later results (Figure 6A), where the free energy landscape for ceftaroline at the allosteric domain is rather flat. Thus the protein's backbone motion can easily displace the drug.

Restraining protein motion slightly improves the stability of ceftaroline's binding. By performing another set of simulations with the same setup as in previous MD with CHARMM, but restraining the protein backbone, one can see that the ligand distribution is mostly centered below 20 Å in Figure 5B and C. In contrast, the distribution shown in Figure 4A and B have a visible population between 30 to 40 Å, indicating that the ligand explores other parts of PBP2a and forms a loosely-bound complex with the protein. Additionally, the allosteric ceftaroline fluctuated less when PBP2a was modelled as a dimer instead of a monomer, c.f. Figure 4A and B as well as Figure 5B and C.

During MD simulations, the allosteric ceftaroline regularly drifted from its crystal structure-revealed binding site. This implies a flat free energy landscape around the allosteric domain. However, the dissociation constant K_d for covalently inhibited PBP2a was measured via fluorescence quenching,^[9] suggesting that a complex does form between the covalently

inhibited PBP2a and ceftaroline. Following this contradiction we further investigated the free energy profile of pulling a ceftaroline away from the allosteric site. This profile, i.e. the potential of mean force (PMF), was calculated using umbrella sampling^[25] using 120 windows, each simulated for 10 ns. Results that violated periodic boundary conditions were discarded prior to analysis. Owing to the collected samples being correlated, the autocorrelation time was calculated to account for said correlation, and 200 bootstraps were used to produce an average PMF.^[26] The standard deviation from this procedure is depicted by the error bar. The convergence check and the correlation time of each window are provided in the supporting information. Calculated PMFs and their associated histograms are shown in Figure 6, where a free energy curve with a shallow potential well is observed. The attraction towards the ligand spreads quite far, up to 3 nm, but is rather weak and any thermal fluctuation easily perturbs the complex, leading to an unbinding event as observed in MD simulations. The binding free energy estimated from the PMF is merely -2.44 kcal/mol, whilst experimental work suggests a value of -6.45 kcal/mol based on a K_d of $20 \mu\text{M}$.^[9] Since the histograms show a overlap relatively well with one another, this discrepancy could come from three sources: (1) insufficient sampling within each window, (2) the accuracy of the force field, and (3) an implicit assumption in the fluorescence quenching protocol. Briefly, the quenching technique measures

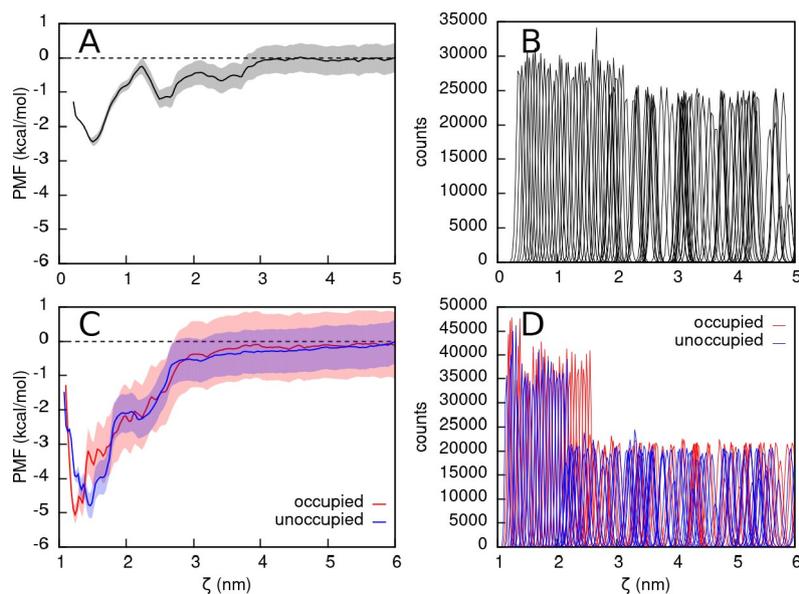


Figure 6. (A) PMF of ceftaroline being pulled away from the allosteric domain. The x axis, ζ , denotes the distance between centres of mass for the ceftaroline fused ring and backbone atoms of residues near the crystal structure ligand (see method section for details). The difference between the minimum and final flat potential gives the estimated binding free energy of ceftaroline at the allosteric site, which is about -2.4 kcal/mol. (B) Associated histograms obtained from PMF calculations using umbrella sampling. Most of the histograms overlap each other very well. For windows with $\zeta < 2$ nm, a slightly stronger force constant was employed, and hence taller histograms were obtained. (C) PMFs of ceftaroline being pulled away from the active site. The ζ now represents the distance between ceftaroline and the centre of mass of backbone atoms from residues around the active site. The red curve represents the free energy profile of a complex whose allosteric site is occupied by another ceftaroline, and the blue curve for when it is not occupied. (D) Histograms obtained for PMF calculations shown in (C). Good overlaps suggest a well converged profile. Again, different force constants were employed, and details are given in the Methods section.

the change in fluorescence from tryptophan and tyrosine residues upon formation of the protein-ligand complex. By changing ligand concentration, the fluorescence is eventually saturated and the binding constant can be determined.^[27] However, this relies on *a priori* knowledge of the nature and number of possible ligand binding sites, e.g. ceftaroline only binding at the allosteric site on PBP2a that has covalently inhibited active sites. The insufficient sampling within each window leads to an error of approximately 0.5 kcal/mol. This error was obtained by comparing PMF values from bootstrapping and correlated data, and is of similar magnitude to the maximum error along the entire PMF. Thus, the error bar shown in Figure 6A for $\zeta < 2$ nm is clearly underestimated. For the second source, we admit that there is still room for force field improvement, but this too can be said for the Amber force field as ceftaroline also drifts in those MD simulations. Although the fluorescence quenching technique's implicit assumption seems almost natural based on the currently available crystal structure, MD simulations (using the same CHARMM force field) reveal that the ligand can form different kinds of complexes, which we will discuss shortly.

The next two PMFs are free energy profiles of pulling ceftaroline away from the active site, with PBP2a's allosteric site either occupied or unoccupied (Figure 6C's red and blue curves respectively). This comparison allows one to determine whether the allosteric mechanism exists. Initially, these results seem to be in good agreement with the mechanism proposed in the literature; the PMF with an occupied allosteric site (red curve) has a potential minimum at a smaller ζ , viz. the ligand is slightly closer to the catalytic residue and its binding is slightly stronger (binding free energy -5.1 kcal/mol for the red curve compared to -4.8 kcal/mol for the blue curve). However, as we have mentioned previously, one should take into account the error bar from bootstrapping, which may be underestimated for $\zeta < 2$ nm. A good estimation for the magnitude of error is about 1 kcal/mol for the red curve and 0.7 kcal/mol for the blue curve, based on estimation of the maximum error along the PMFs and deviation of the correlated PMF from the bootstrap one. Thus, the difference between the two PMF curves is not statistically significant and therefore negligible. Although these calculations cannot support the theory of ceftaroline inducing the allosteric mechanism, experiments have shown that when using allosteric drugs (e.g. quinazolinone) together with the originally ineffective antibiotic piperacillin and β -lactamase inhibitor tazobactam, a synergetic effect against MRSA can be achieved.^[28] For completeness, histograms corresponding to the two PMFs in Figure 6C are depicted in Figure 6D, showing good overlap between them.

In addition to the MD simulations previously mentioned, another set were conducted involving a PBP2a *apo* monomer with six ceftarolines moving freely in solvent. Simulations were performed with both CHARMM and Amber force fields, and each set again consisted of three independent 100 ns runs. One CHARMM trajectory shows three ceftarolines binding to the surface of PBP2a at different positions. The first sits

directly outside the entrance to the active site on the transpeptidase domain (Figure 7A), while residue Tyr446 constantly gates the entrance, preventing the ligand from entering the pocket (Figure 7B). Another two ceftarolines bind at unexpected locations: one between lobes L1 and L3, with the L3 helix rotated 90 degrees from its original crystal structure position (see the non-transparent white and cyan helix conformations in Figure 7A). The other binds at the N-terminal extension. All three ceftarolines exhibit more stable binding than when at the proposed allosteric site – none of them dissociated from PBP2a even when the simulation time was doubled to 200 ns. Notably, all three binding events were accompanied by the coordination of one Ca^{2+} ion, via the carboxylate on ceftaroline and a Glu/Asp residue. An example is given in Figure 7B where Ca^{2+} is depicted by an orange sphere as in panel A. Glu447 interacts with the ion from one side and the ligand's carboxylate from the other. The Ca^{2+} were introduced to replace any formerly present Cd^{2+} from the crystallisation conditions. These Cd^{2+} can be found at multiple sites in crystal structures including 3zfv and 1vqq. In the 3zfv structure for example, two Cd^{2+} ions were coordinated by residues from both chains at the dimer interface, near ceftaroline's allosteric binding site. Since ceftaroline is not shown interacting with Cd^{2+} in any crystal structures, the observation of ceftaroline binding to Ca^{2+} may be a simulation artefact. Nevertheless, in light of this binding mechanism, the MD simulations of an allosteric complex, whose results were shown previously in Figure 4, had ceftaroline attracting Ca^{2+} from time to time. However, no Asp/Glu residues were in the right position for ceftaroline at its crystal pose, making it unlikely that the ligand was anchored via this mechanism. Consequently, the ligand wanders off and explores different parts of the allosteric domain, until it encounters an Asp/Glu residue to bind it through Ca^{2+} coordination, as revealed in Figure 4A by the blue mesh around lobe L1. In reality, the coordination of Ca^{2+} could be weaker than what is reported here. A recent study^[29] noted again the problem of current non-polarisable force fields: they fail to adjust the dielectric constant (i.e. relative permittivity) for the core of a large protein. The dielectric constant is much smaller in proteins (~ 2) than in bulk water (~ 80), as proteins cannot rearrange their conformations to screen an ion's charge the way water molecules can. The screening effect of proteins predominantly comes from their electron cloud being polarised by an ion. A quick solution is to scale the ion's charge by a factor of 0.75, making the real effect of dication coordination weaker.^[29,30] As a counterexample, simulations with Amber ff14SB and NaCl ions (i.e. without Ca^{2+}) were also conducted. Only one of the three 100 ns simulations had ceftaroline binding outside the active site entrance, whilst another had a ceftaroline between the lobes. Major clusters of the ligands are shown in Figure 7C, with cluster sizes making up 56.4% and 58.1% of the two trajectories respectively. Although their position differs slightly from those revealed using the CHARMM force field, these simulations generally suggest that ceftaroline can bind at the active site entrance or the lobes, forming a complex

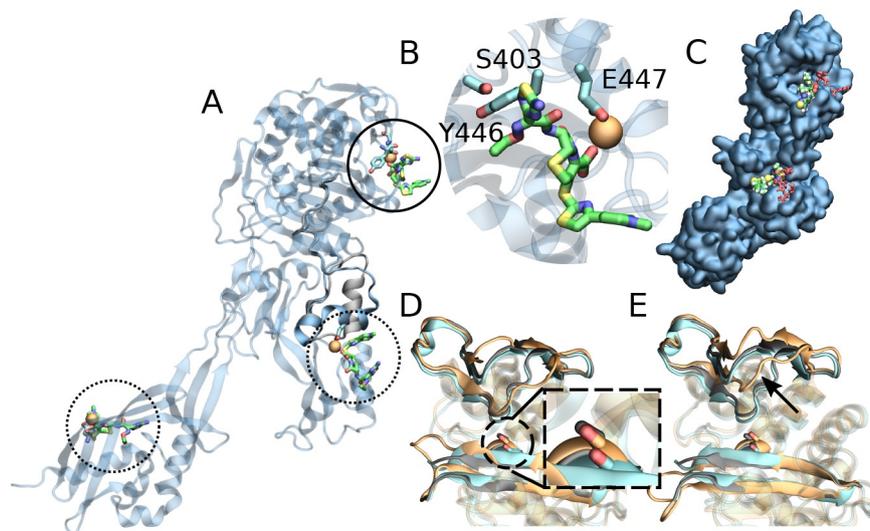


Figure 7. (A) Three binding spots revealed by a 200 ns MD simulation of *apo* protein with six free ceftarolines in solvent. Simulations were performed using the CHARMM36 force field. Ligands are depicted in green sticks. One is located by the entrance to the active site (circled), another at the allosteric domain (sandwiched between lobes L1 and L3), and the last one at the N-terminal extension. Shown here is the centroid of the major cluster on the backbone and three ligands, which occupies 66.2% of the trajectory. (B) Cefaroline at the active site from panel A. The binding is assisted by a coordinating Ca^{2+} (orange sphere) together with Glu447. Tyr446 acts as a gate, preventing the ligand from entering. The catalytic residue S403 is labelled to indicate the location of the active site. (C) Binding spots revealed by two independent 100 ns simulations with Amber ff14SB. The system was neutralised with just NaCl and no Ca^{2+} was introduced. Ligand conformations represent centroids of the major ligand clusters from two individual simulations. The major cluster of the occupied active site accounts for 56.4% of a 100 ns simulation, and that for a ligand binding between the lobes takes up 58.1% of another 100 ns trajectory. These are similar binding locations to those in panel A (depicted in red sticks). (D) The centroid of a cluster (accounting for 8.7% of the trajectories) for the active site conformation found in *apo* PBP2a simulations (orange). The active site conformation was identified by the backbone of residues 402–403, the α_2 - α_3 loop (residues 417–457), and the β_3 - β_4 loop (residues 602–612). An important difference between the *apo* and *holo* structures is the position of Ser403 (circled). The former is partially blocked by β -sheets, whilst the latter is accessible and can undergo reactions. This centroid conformation resembles the *holo* crystal structure (black) more than the *apo* one (cyan), proving that thermal fluctuations alone can prime the active site for reaction. (E) Similar to panel D but for a smaller cluster (size 1.9%). Note that the α_2 - α_3 loop (highlighted by the arrow) moves further away, opening up the entrance to the pocket.

with PBP2a there. Since there are Tyr residues at both locations, site-specific experiments would be required to determine the exact source of fluorescence quenching.

Finally, there are crystal structures whose active sites are inhibited without ligands occupying the allosteric domain, e. g. the methicillin-PBP2a complex (1mwu) and the penicillin G-PBP2a complex (1mw).^[6] This can occur due to thermal fluctuations causing conformational changes, inadvertently preparing the active site for reaction; MD simulations of the *apo* PBP2a monomer revealed these different active site conformations (Figure 7D and E). Noticeably large populations, such as one with a cluster size of 8.7%, adopt a *holo*-like conformation where the catalytic residue Ser403 is accessible. A significantly smaller population (cluster size 1.9%) has the active site opening up through movement of the α_2 - α_3 loop (residues 417–457). If a high concentration of β -lactam is present in solution, inhibition of the active site would then be possible.

Conclusions

In this work we employed MD simulations and free energy calculations to investigate the allosteric mechanism of PBP2a from MRSA. More specifically, we focussed on understanding the nature of ceftaroline's binding at its proposed allosteric site. Through sequence alignments via MAFFT and Clustal Omega, as well as FTMAP binding site prediction, the active site and muramate/peptidoglycan binding site were confirmed. These analyses support the allosteric mechanism that PBP2a binds to the saccharide component on peptidoglycan, which conditions the active site for reaction. Interestingly, MD simulations reveal ceftaroline's binding at the allosteric site to be unstable. The ligand can easily drift away due to the lack of conserved residues anchoring it, while a large protein backbone motion occurs naturally by thermal fluctuation. This unstable binding was observed in both monomer and dimer systems, as well as in simulations using CHARMM36 and Amber ff14SB force fields. Free energy calculations via umbrella sampling supported these observations from MD

simulations, with the presence of the allosteric ligand having no effect on orthosteric binding free energies. We thus suggest that ceftaroline's ability to trigger the allosteric mechanism is too weak to be observed using MD simulations. To conclude, our results are partially congruent with previous experiments of PBP2a. The native peptidoglycan substrate's allosteric regulation is supported by the presence of conserved residues and predicted binding sites, but ceftaroline's allosteric effect is not supported by the current study.

Methods

The initial structures for *apo* and *holo* simulations were taken from the RCSB Protein Data Bank, pdb codes 1vqq and 3zfv. The active site's acylated Ser403-ligand moiety was removed, and subsequently modelled as a normal serine in all simulations. Missing loops were constructed with MODELLER,^[31,32] and the protein's protonation states were determined using the PROPKA package.^[33,34] The protein (or complex) was solvated in a dodecahedron water box, with box edges 15 Å from the protein. This generated systems of 318,000 atoms (monomer) or 328,000 atoms (dimer) after salting. Crystal structure ions Cd^{2+} were replaced by Ca^{2+} for MD simulations. For CHARMM and Amber systems, 0.1 M CaCl_2 and NaCl were used for neutralising the systems respectively, except for PMF calculations where the CHARMM system was also neutralised with 0.1 M NaCl. Prior to production runs, systems were subjected to 50,000 steps steepest descent energy minimisation, followed by 500 ps NVT equilibration to equilibrate the temperature. 500 ps of NPT equilibration was then employed to equilibrate the pressure, and another 5 ns of NPT with protein backbone atoms restrained to equilibrate the side chains. For each system, three 100 ns production runs were performed using GROMACS 2018.6.^[35] The leapfrog stochastic dynamics integrator for Langevin dynamics, with a step size 2 fs, and a Verlet list with an update frequency of 20 fs, were employed for propagation. The van der Waals potential was switched off from 10 Å and cut off completely at 12 Å, while the short-range cutoff for electrostatic interactions was also set at 12 Å. Particle mesh Ewald summation (PME)^[36,37] was employed to account for electrostatic interactions involving periodic boundary conditions, and a grid space of 1.5 Å and spline order 6 were used. The system temperature was set to 300 K through Langevin dynamics, and the pressure to 1 bar using the Parrinello-Rahman barostat. Finally, all bonds were constrained via the LINCS algorithm.^[38,39] The CHARMM36^[11] and Amber ff14SB force fields^[13] were both employed. The associated ligand parameters were generated from CGenFF^[12,40, 41] and AmberTools (AM1-BCC).^[42,43] Ceftaroline's charges, bonds, angles, and dihedrals with high penalty scores were further optimised by using the force field toolkit (FFTK)^[44] with Gaussian.^[45] AM1-BCC also generated some dihedral parameters with high penalty scores, but no further optimisation was performed. Sequence alignments were

generated by Clustal Omega^[19] for the nine PDBs and MAFFT for the 150 representative homologues.^[18] Results were visualised using VMD^[46] and PyMOL.^[47] FTMAP^[21,22] was employed to predict possible binding sites, while MD simulations were analysed via VMD's VolMap plugin, clustering analysis,^[48] PCA,^[49] and self-implemented scripts for distance distributions. PMF calculations were conducted using GROMACS pulling code, following the protocol documented in literature.^[50] Briefly, the ligand was pulled away from its binding site. Frames from the resultant trajectory were used as initial conformations for sampling. The pull rate was chosen to be 0.05 Å/ps, while the allosteric and active ligands were pulled in the Y and X direction respectively. Pulling group A was chosen as the heavy atoms on ceftaroline's fused rings. Pulling group B was chosen as the protein backbone atoms in the vicinity of the pulled ligand, specifically those forming stable secondary structure elements in the simulation's starting structure. For the allosteric site, these were residues 71–73, 103, 106, 143–147, 294, 296–299, 309, and 316. For the active site, they were residues 404, 406, 461, 464, 591, 597–601, 642, and 643. The distance between the centers of mass of pulling group A and B is defined as ζ . To ensure a good overlap, windows were chosen based on increasing distance ζ for every 0.5 Å. This led to 120 windows for the allosteric site and 104 windows for the active site PMF calculations. For each window, 10 ns NPT simulations with restraint force constant $k = 1000 \text{ kJ/mol/nm}^2$, $k = 1500 \text{ kJ/mol/nm}^2$, or $k = 4000 \text{ kJ/mol/nm}^2$ were carried out for sampling. All data were then combined using g-wham^[25,26] and 200 bootstraps for error estimation.

Acknowledgement

YCC thanks Prof. Peter Roach and Prof. Gergely Katona for sharing their views on electron densities of the ligands, Yui Tik Pang for sharing his protocol on optimising ligand force fields with FFTK, The Royal Society for funding (NF171278), and iSolutions at the University of Southampton for computing time on the Iridis5 cluster. MTYW thanks the BBSRC (BB/P504713/1) and UCB for funding.

References

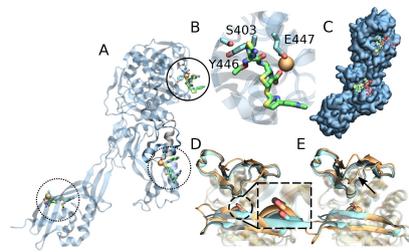
- [1] D. J. Tipper, J. L. Strominger, *Proc. Natl. Acad. Sci. USA* **1965**, *54*, 1133.
- [2] R. R. Yocum, D. J. Waxman, J. R. Rasmussen, J. L. Strominger, *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 2730.
- [3] F. D. Lowy, *J. Clin. Invest.* **2003**, *111*, 1265.
- [4] A. Bondi, Jr., C. C. Dietz, *Proc. Soc. Exp. Biol. Med.* **1945**, *60*, 55.
- [5] N. A. Turner, B. K. Sharma-Kuinkel, S. A. Maskarinec, E. M. Eichenberger, P. P. Shah, M. Carugati, T. L. Holland, V. G. Fowler, *Nat. Rev. Microbiol.* **2019**, *17*, 203.
- [6] D. Lim, N. C. J. Strynadka, *Nat. Struct. Biol.* **2002**, *9*, 870.

- [7] L. H. Otero, A. Rojas-Altuve, L. I. Llarrull, C. Carrasco-López, M. Kumarasiri, E. Lastochkin, J. Fishovitz, M. Dawley, D. Heseck, M. Lee, J. W. Johnson, J. F. Fisher, M. Chang, S. Mobashery, J. A. Hermoso, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16808.
- [8] K. V. Mahasenan, R. Molina, R. Bouley, M. T. Batuecas, J. F. Fisher, J. A. Hermoso, M. Chang, S. Mobashery, *J. Am. Chem. Soc.* **2017**, *139*, 2102.
- [9] J. Fishovitz, A. Rojas-Altuve, L. H. Otero, M. Dawley, C. Carrasco-López, M. Chang, J. A. Hermoso, S. Mobashery, *J. Am. Chem. Soc.* **2014**, *136*, 9814.
- [10] J. Fishovitz, N. Taghizadeh, J. F. Fisher, M. Chang, S. Mobashery, *J. Am. Chem. Soc.* **2015**, *137*, 6500.
- [11] J. Huang, A. D. MacKerell Jr, *J. Comput. Chem.* **2013**, *34*, 2135..
- [12] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. Mackerell Jr., *J. Comput. Chem.* **2010**, *31*, 671.
- [13] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Comput. Chem.* **2015**, *11*, 3696.
- [14] M. Landau, I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko, N. Ben-Tal, *Nucleic Acids Res.* **2005**, *33*, W299.
- [15] H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, N. Ben-Tal, *Nucleic Acids Res.* **2016**, *44*, W344.
- [16] B. E. Suzek, Y. Wang, H. Huang, P. B. McGarvey, C. H. Wu, the UniProt Consortium, *Bioinformatics* **2014**, *31*, 926.
- [17] J. Mistry, R. D. Finn, S. R. Eddy, A. Bateman, M. Punta, *Nucleic Acids Res.* **2013**, *41*, e121.
- [18] K. Katoh, D. M. Standley, *Mol. Biol. Evol.* **2013**, *30*, 772.
- [19] F. Madeira, Y. M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. Tivey, S. C. Potter, R. D. Finn, R. Lopez, *Nucleic Acids Res.* **2019**, *47*, W636.
- [20] S. El-Gebali, J. Mistry, A. Bateman, S. R. Eddy, A. Luciani, S. C. Potter, M. Qureshi, L. J. Richardson, G. A. Salazar, A. Smart, E. L. L. Sonnhammer, L. Hirsh, L. Paladin, D. Piovesan, S. C. E. Tosatto, R. D. Finn, *Nucleic Acids Res.* **2018**, *47*, D427.
- [21] R. Brenke, D. Kozakov, G.-Y. Chuang, D. Beglov, D. Hall, M. R. Landon, C. Mattos, S. Vajda, *Bioinformatics* **2009**, *25*, 621.
- [22] D. Kozakov, L. E. Grove, D. R. Hall, T. Bohnuud, S. E. Mottarella, L. Luo, B. Xia, D. Beglov, S. Vajda, *Nat. Protoc.* **2015**, *10*, 733.
- [23] O. Levsh, Y.-C. Chiang, C. F. Tung, J. P. Noel, Y. Wang, J.-K. Weng, *Biochemistry* **2016**, *55*, 6314.
- [24] Y.-C. Chiang, O. Levsh, C. K. Lam, J.-K. Weng, Y. Wang, *PLoS Comput. Biol.* **2018**, *14*, 1.
- [25] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, *J. Comput. Chem.* **1992**, *13*, 1011.
- [26] J. S. Hub, B. L. de Groot, D. van der Spoel, *J. Chem. Theory Comput.* **2010**, *6*, 3713.
- [27] M. van de Weert, L. Stella, *J. Mol. Struct.* **2011**, *998*, 144.
- [28] J. Janardhanan, R. Bouley, S. Martínez-Caballero, Z. Peng, M. Batuecas-Mordillo, J. E. Meisel, D. Ding, V. A. Schroeder, W. R. Wolter, K. V. Mahasenan, J. A. Hermoso, S. Mobashery, M. Chang, *Antimicrob. Agents Chemother.* **2019**, *63*, e02637–18.
- [29] B. J. Kirby, P. Jungwirth, *J. Phys. Chem. Lett.* **2019**, *10*, 7531.
- [30] I. V. Leontyev, A. A. Stuchebrukhov, *J. Chem. Theory Comput.* **2012**, *8*, 3207.
- [31] A. Fiser, R. K. G. Do, A. Sa'li, *Prot. Sci.* **2000**, *9*, 1753.
- [32] B. Webb, A. Sa'li, *Curr. Protoc. Bioinformatics* **2016**, *54*, 5.6.1.
- [33] M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 525.
- [34] C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 2284.
- [35] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1–2*, 19.
- [36] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089.
- [37] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577.
- [38] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463.
- [39] B. Hess, *J. Chem. Theory Comput.* **2008**, *4*, 116.
- [40] W. Yu, X. He, K. Vanommeslaeghe, A. D. MacKerell, Jr., *J. Comput. Chem.* **2012**, *33*, 2451.
- [41] *Cgenff interface at paramchem.org*; <https://cgenff.umaryland.edu>.
- [42] A. Jakalian, B. L. Bush, D. B. Jack, C. I. Bayly, *J. Comput. Chem.* **2000**, *21*, 132.
- [43] A. Jakalian, D. B. Jack, C. I. Bayly, *J. Comput. Chem.* **2002**, *23*, 1623.
- [44] C. G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid, J. C. Gumbart, *J. Comput. Chem.* **2013**, *34*, 2757.
- [45] *Gaussian 09, revision d.01* (2013), gaussian Inc. Wallingford CT.
- [46] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33.
- [47] *The pymol molecular graphics system*, version 2.0 Schrödinger, LLC.
- [48] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 236.
- [49] D. J. David, C. Charles, Jacobs, *Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins*, Humana Press, Totowa, NJ, **2014**, pp. 193–226.
- [50] J. A. Lemkul, D. R. Bevan, *J. Phys. Chem. B* **2010**, *114*, 1652.

Manuscript received: February 12, 2020

Revised manuscript received: March 12, 2020

Version of record online: ■■, ■■



*Y.-C. Chiang**, *M. T. Y. Wong*, *J. W. Essex*

1 – 11

Molecular Dynamics Simulations of Antibiotic Ceftaroline at the Allosteric Site of Penicillin-Binding Protein 2a (PBP2a)
