JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Unveiling the Atomic-Level Determinants of Acylase−Ligand Complexes: An Experimental and Computational Study

Luca Mollica,[*,†] Gianluca Conti,[‡] Loredano Pollegioni,[‡,§] Andrea Cavalli,[†,∥] and Elena Rosini[*,‡,§]

[†]CompuNet, Istituto Italiano di Tecnologia, via Morego 30, 16163 Genova, Italy

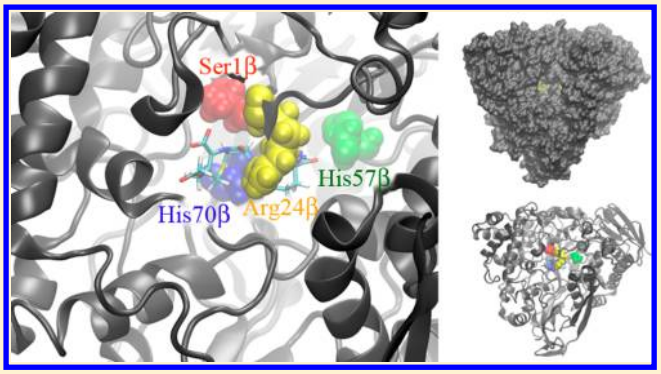
[‡]Dipartimento di Biotecnologie e Scienze della Vita, Università degli studi dell'Insubria, via J. H. Dunant 3, 21100 Varese, Italy

[§]The Protein Factory, Centro Interuniversitario di Biotecnologie Proteiche, Politecnico di Milano and Università degli studi dell'Insubria, via Mancinelli 7, 20131 Milano, Italy

[∥]Department of Pharmacy and Biotechnology, Alma Mater Studiorum, University of Bologna, via Belmeloro 6, 40126 Bologna, Italy

Supporting Information

ABSTRACT: The industrial production of higher-generation semisynthetic cephalosporins starts from 7-aminocephalosporanic acid (7-ACA), which is obtained by deacylation of the naturally occurring antibiotic cephalosporin C (CephC). The enzymatic process in which CephC is directly converted into 7-ACA by a cephalosporin C acylase has attracted industrial interest because of the prospects of simplifying the process and reducing costs. We recently enhanced the catalytic efficiency on CephC of a glutaryl acylase from *Pseudomonas* N176 (named VAC) by a protein engineering approach and solved the crystal structures of wild-type VAC and the H57*β*S−H70*β*S VAC double variant. In the present work, experimental measurements on several CephC derivatives and six VAC variants were carried out, and the binding of ligands into the VAC active site was investigated at an atomistic level by means of molecular docking and molecular dynamics simulations and analyzed on the basis of the molecular geometry of encounter complex formation and protein−ligand potential of mean force profiles. The observed significant correlation between the experimental data and estimated binding energies highlights the predictive power of our computational method to identify the ligand binding mode. The present experimental−computational study is well-suited both to provide deep insight into the reaction mechanism of cephalosporin C acylase and to improve the efficiency of the corresponding industrial process.



INTRODUCTION

Enzymes utilized in industrial processes as biocatalysts frequently show high activity under mild conditions and possess high substrate specificity, and these properties result in significant advantages of biocatalytic methods compared with chemical processes. On the other hand, the industrial application of these enzymes is frequently hampered by strict/narrow specificity, which can affect the recognition of structurally similar or related compounds. Broad substrate acceptance is of utmost importance to boost new biotechnological applications of known biocatalysts.[1] On this side, the focus is to unravel the relationships between protein structure and function and to shed light on the interactions between the macromolecular target and ligands.[2,3]

Glutaryl acylases (GAs, EC 3.5.1.93) are well-known industrial biocatalysts that belong to the N-terminal hydrolase class of hydrolytic enzymes and show wide substrate specificity. The most important application of GAs is the two-step process comprising removal of the D-*α*-amino adipoyl side chain of glutaryl-7-aminocephalosporanic acid (Gl-7-ACA), which is produced from cephalosporin C (CephC) by D-amino acid oxidase (DAAO, EC 1.4.3.3), to give 7-aminocephalosporanic acid (7-ACA).[4,5] Actually, more than 50 marketed cephalosporins are produced: two-thirds of semisynthetic cephalosporins are synthesized using 7-ACA as the starting scaffold.[6] There are several advantages to producing 7-ACA from CephC by using an enzymatic process rather than a chemical one, which are mainly related to the quality of the final product.[7]

We recently investigated the substrate preference[8] and solved the structure of a wild-type GA from *Pseudomonas* N176 (named VAC)[9] (Figure 1) as well as the structure of the H57*β*S−H70*β*S (HS-HS) double variant.[10] Structure-guided protein engineering studies allowed isolation of VAC variants with improved properties for the industrial application.[10,11]

In this work, the substrate preferences of different VAC variants for several CephC derivatives of industrial interest were investigated.[8] Binding of ligands into the active site was assessed by structural comparison and molecular dynamics (MD) simulations.[12,13] In particular, MD simulations were

Received: May 30, 2015
Published: September 30, 2015

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

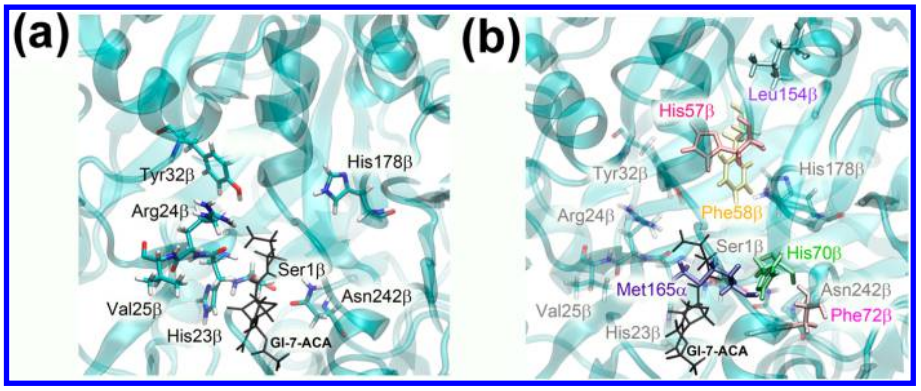


Figure 1. Structure of the wild-type VAC binding site. In (a), the most relevant residues surrounding the ligand (i.e., Gl-7-ACA, represented in black sticks) are represented in sticks. In (b), the residues that were mutated in order to generate the different VAC variants are represented in solid sticks and uniformly colored. The view in (b) has been rotated clockwise by 15° with respect to the one in (a).

carried out with nine VAC−ligand complexes formed with six different ligands. Our simulations showed a significant correlation between the experimental data and estimated binding energies, highlighting the rather good predictive power of MD simulations to identify the binding modes of different ligands. The mixed computational and experimental approach provides a powerful tool that can guide the design of novel enzymes for new ligands of industrial interest. In addition, the present study can provide fundamental knowledge for future investigations aimed at exploring the catalytic mechanism by means of quantum-mechanical calculations.

METHODS

Enzymes and Reagents. Recombinant VAC proteins (wild-type, HS-HS, M165αS-HS-HS, HS-HS-F58βD, HS-HS-F72βR, and HS-HS-L154βY) were obtained as described in ref 11. From 1 L of fermentation broth, 55, 77, 117, 24, 103, and 35 mg of wild-type, HS-HS, M165αS-HS-HS, HS-HS-F58βD, HS-HS-F72βR, and HS-HS-L154βY were produced with specific activities of 0.7, 3.0, 4.2, 1.3, 1.0, and 3.3 units (mg of protein)$^{-1}$ on CephC as the substrate, respectively. The final enzyme preparations were equilibrated in 20 mM potassium phosphate buffer (pH 8.0). The amount of protein was estimated by the absorbance at 280 nm using a molar extinction coefficient of 110 mM^{-1} cm^{-1}.

CephC, 7-ACA, and Gl-7-ACA were kindly supplied by Antibioticos S.p.A. (Rodano, MI, Italy). All of the other chemicals were of analytical grade and were used as received. CephC derivatives were prepared as described in ref 8.

Activity and Kinetic Measurements. The standard activity assay was based on hydrolysis of the substrate to give 7-ACA and the subsequent formation of a yellow Schiff base (with a maximum of absorbance at 415 nm, ε = 0.635 mM^{-1} cm^{-1}) by the reaction of *p*-4-dimethylaminobenzaldehyde (pDMAB) with the primary amino group of 7-ACA.[9,14] One unit of acylase is defined as the amount of enzyme that converts 1 μmol of substrate per minute under the assay conditions. In detail, 0.1 mL of enzyme was mixed with 0.1 mL of substrate at pH 8.0, and the mixture was incubated for 10 min at 25 °C. The reaction was stopped by addition of 0.6 mL of 20% acetic acid, and then 0.133 mL of 0.5% (w/v) pDMAB in methanol was added. The mixture was incubated for a further 10 min at 25 °C, and then the absorbance at 415 nm was measured. The kinetic parameters were determined similarly using a fixed amount of enzyme and various substrate concentrations (0−150 mM); activity versus substrate concentration data were analyzed according to the classical Michaelis−Menten equation or an equation modified to account for a substrate inhibition effect.[8]

Molecular Docking Studies. Automated ligand docking was performed using AutoDock Vina, a package based on an iterated local search global algorithm.[15] PyMOL software (www.pymol.org/) was used for analysis of the docking results and identification of non-covalent interactions between the protein and ligands. A script was employed to perform virtual screening experiments. The docking of different ligands (substrates) to VAC structures obtained by means of simple in silico mutagenesis or MD simulations (see the following section for details) with different ligands (substrates) was carried out. The script made it possible to automatically perform all dockings and to filter the results on the basis of the distance between the carbonyl carbon of the amide bond of the ligand and the nucleophile Oγ of Ser1β as set by the user (3−5 Å). The binding site was defined as centered on the residue Ser1β, and the search space was defined as the volume of the ligand augmented in each direction by 15 Å. The number of docking poses generated by each run was 10. A Lamarckian genetic algorithm was used.

The structures of wild-type VAC and the HS-HS VAC variant were retrieved from the Protein Data Bank (PDB) with the PDB IDs 4HSR and 4HST, respectively. The molecular models of the M165αS-HS-HS, HS-HS-F58βD, HS-HS-F72βR, and HS-HS-L154βY VAC variants were obtained starting from the HS-HS VAC variant using the mutagenesis wizard tool in PyMol.

MD Simulations. *a. MD Simulations of VAC Variants.* The structures of the ligand-free forms of the VAC variants were used as a starting point for MD simulation performed using the GROMACS package, version 4.6.7. Each was placed in the geometrical center of a parallelepiped-shaped box with a volume of ∼900 nm^3 (∼9 nm × 9 nm × 10 nm), minimized in vacuo via a steepest-descent method, and then solvated with almost 32 000 TIP3P water molecules[16] using TLEAP. To preserve the electroneutrality of the system, 11−13 water molecules (wild-type, 12; HS-HS, 13; M165αS-HS-HS, 12; HS-HS-F58βD, 13; HS-HS-F72βR, 11; HS-HS-L154βY, 12) were replaced with sodium ions. The system was minimized again with the steepest-descent method, followed by equilibration of the restrained protein (with an isotropic 1000 kJ mol^{-1} nm^{-1} force applied to each heavy atom of the protein backbone) in the *NVT* ensemble (up to 500 ps) at 300 K and then a 20 ns

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

Table 1. Kinetic Parameters of Wild-Type VAC and the HS-HS, M165αS-HS-HS, HS-HS-F58βD, HS-HS-F72βR, and HS-HS-L154βY VAC Variants on Different Substrates

substrate	parameter	wild-type[a]	HS-HS[a]	M165αS-HS-HS[b]	HS-HS-F58βD[b]	HS-HS-F72βR[b]	HS-HS-L154βY[b]
Gl-7-ACA	V_{max} (units mg^{-1})	36.4 ± 2.7	4.8 ± 0.1	6.2 ± 0.2	1.0 ± 0.1	52.9 ± 2.3	7.9 ± 0.8
	K_m (mM)	1.5 ± 0.2	6.9 ± 0.7	21.9 ± 2.4	0.4 ± 0.2	2.0 ± 0.2	2.0 ± 0.6
	V_{max}/K_m	24.3	0.7	0.3	2.5	26.4	3.9
	K_i (mM)	21.0 ± 3.4	–	–	–	6.3 ± 2.1	–
CephC	V_{max} (units mg^{-1})	0.7 ± 0.1	3.0 ± 0. 1	4.2 ± 0.1	1.3 ± 0.1	1.6 ± 0.06	3.3 ± 0. 2
	K_m (mM)	9.5 ± 0.3	12.2 ± 0.9	27.7 ± 1.4	1.5 ± 0.3	16.4 ± 2.3	4.5 ± 0.7
	V_{max}/K_m	0.08	0.24	0.15	0.87	0.1	0.73
	K_i (mM)	–	–	–	–	–	–
succinyl-7-ACA	V_{max} (units mg^{-1})	21.1 ± 1.9	4.9 ± 0.1	6.8 ± 0.2	0.2 ± 0.01	5.3 ± 1.9	4.4 ± 0.2
	K_m (mM)	2.8 ± 0.5	5.4 ± 0.5	4.1 ± 0.4	3.7 ± 0.5	22.4 ± 10.5	2.1 ± 0.3
	V_{max}/K_m	7.5	0.9	1.8	0.05	0.2	2.1
	K_i (mM)	23.9 ± 4.7	–	–	–	16.7 ± 8.4	–
adipyl-7-ACA	V_{max} (units mg^{-1})	11 ± 1.5	2.2 ± 0.05	4.8 ± 0.3	0.1 ± 0.01	2.4 ± 0.2	5.9 ± 0.3
	K_m (mM)	7.5 ± 0.5	4.3 ± 0.3	4.5 ± 0.8	2.3 ± 0.2	2.7 ± 0.7	5.4 ± 0.9
	V_{max}/K_m	1.5	0.5	1.1	0.06	0.9	1.1
	K_i (mM)	112 ± 15	–	–	–	–	–
pentanoyl-7-ACA	V_{max} (units mg^{-1})	2.2 ± 0.1	0.20 ± 0.05	3.8 ± 0.1	0.1 ± 0.003	2.6 ± 0.1	1.9 ± 0.1
	K_m (mM)	17.4 ± 3.1	16.1 ± 0.8	14.1 ± 1.6	2.7 ± 0.3	18.0 ± 2.0	3.8 ± 0.7
	V_{max}/K_m	0.1	0.01	0.3	0.05	0.1	0.5
	K_i (mM)	–	–	–	–	–	–

[a]Taken from ref 8. [b]Kinetic parameters on Gl-7-ACA and CephC were taken from ref 11.

unrestrained MD run. Electrostatics was treated with the cutoff method for short-range interactions and with the particle mesh Ewald method for the long-range ones (rlist = 1 nm, cutoff distance = 0.9 nm, vdW distance = 0.9 nm, PME order = 4).[17] The constant-temperature conditions were provided by means of the V-rescale thermostat,[18] which is a modification of Berendsen's coupling algorithm. The AMBER99SB force field was used for the simulations.

In order to sample the variations of the volume of the binding cavity and to extract from the simulations a set of structures to be docked with different ligands, we monitored the solvent-accessible surface (SAS) of the binding site, defined by all of the residues that are in direct contact with the ligand in PDB structure 4HSR along the trajectory. The binding-site SAS displayed fluctuations between 24 and 29 nm^2 for each system (see below). We first binned this distribution using 1 nm^2 steps and extracted a trajectory frame corresponding to the observed SAS. Then we selected only those to which a statistical relevance could be attributed, indicating that they can be considered representative of the conformational space explored by the binding site alone. To this aim, we performed a cluster analysis[19] based on a complete (i.e., including the backbone and side chains) geometric superimposition (within a 0.5 Å cutoff) of the binding-site residues along the trajectory, selecting only the structure of the pool that belonged to clusters with a population higher than 10%.

This procedure led to subensembles of structures for each VAC variant composed of a minimum of three to a maximum of six structures.

b. MD Simulations of VAC Variant–Ligand Complexes. Each compound used in the study was originally designed in PyMol, protonated with PyMol's internal utility assuming a neutral pH, and then geometry-optimized via a quantum-mechanical approach. Electron density calculations were performed in Gaussian 09[20] using density functional theory with the B3LYP functional and the 6-31G* basis set.[21] Partial charges were derived using the restrained electrostatic potential (RESP) method[22] as implemented in Antechamber[23] and included in an Amber-compatible (i.e., based on the General Amber Force Field (GAFF) parameter set) topological description suitable for GROMACS, the package used to perform all of the MD simulations of the protein–ligand complexes.

The structures of the complexes obtained by docking procedures as previously described were used as starting points for MD simulations performed using the GROMACS package (see above). Each was placed in the geometrical center of an almost-cubic-shaped box with a volume of ∼900 nm^3 (∼9 nm × 9 nm × 10 nm), minimized in vacuo via a steepest-descent method, and then solvated with almost 32 000 TIP3P water molecules[16] using TLEAP. To preserve the electroneutrality of the system, 13 (HS-HS-F72βR with Gl-7-ACA, HS-HS-F72βR with adipyl-7-ACA, and HS-HS-L154βY with pentanoyl-7-ACA) or 14 (wild-type with Gl-7-ACA, wild-type with succinyl-7-ACA, HS-HS with CephC, M165αS-HS-HS with adipyl-7-ACA, HS-HS-F58βD with CephC, and HS-HS-L154βY with CephC) water molecules were replaced with sodium ions. The systems were assembled and processed with the help of the recently developed BiKi software package (www.bikitech.com) as described previously,[24] and then the simulations were performed as described above.

c. PMF and Umbrella Sampling. In order to sample the potential of mean force (PMF) for the system, several configurations were generated by pulling each ligand away from the VAC binding cavity along the z axis at a constant speed and force (steered molecular dynamics (SMD)), and a harmonic potential was applied to selected configurations in order to sample the progressively increasing distance between the centers of mass of the ligand and its binding site (umbrella sampling (US)).[25] In order to facilitate the setup and to be able to explore the free energy surface by means of the SMD/US procedure, we built new systems starting from the average configurations obtained by sampling the last 10 ns of the whole standard simulations (see above). We put the complexes into

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227–2241

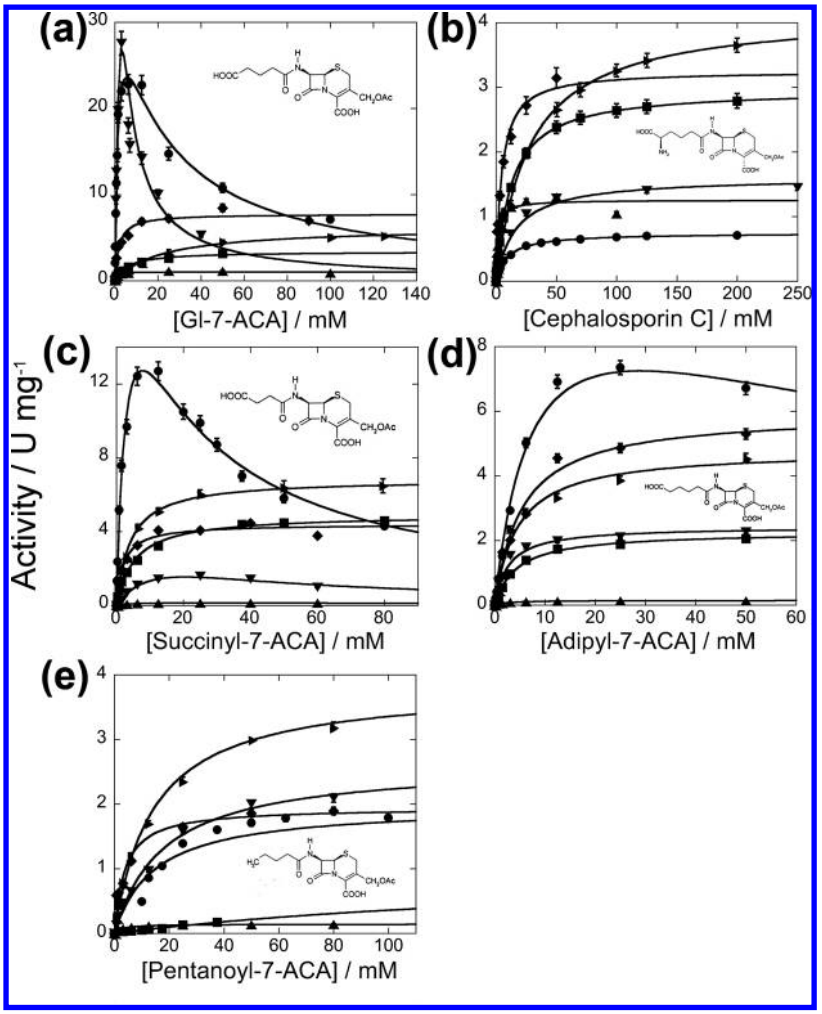


Figure 2. Dependence of the activity values for the reactions of wild-type VAC (●) and the HS-HS (■), M165αS-HS-HS (▶), HS-HS-F58βD (▲), HS-HS-F72βR (▼), and HS-HS-L154βY (◆) VAC variants on the concentrations of different cephalosporin derivatives: (a) Gl-7-ACA; (b) CephC; (c) succinyl-7-ACA; (d) adipyl-7-ACA; (e) pentanoyl-7-ACA. The data points were obtained as described in the text and were determined using different amounts of enzyme (wild-type VAC, 3.8−32 μg/assay; HS-HS, 57−445 μg/assay; M165αS-HS-HS, 10−20 μg/assay; HS-HS-F58βD, 7.6−264 μg/assay; HS-HS-F72βR, 26.6−35 μg/assay; HS-HS-L154βY, 9.6−21 μg/assay). The curves are from fits based on the classical Michaelis−Menten equation or its modification to account for a substrate inhibition effect. The reported values are averages of three experiments; where not shown, the error bars are smaller than the symbol used.

highly asymmetrical boxes with dimensions of 9 nm × 9 nm × 15 nm and the open side of the binding pocket oriented along the z axis, allowing easy unbinding of the ligand following the simplest pathway, and we brought the ligand from a bound state to a fully free form in solution. The new boxes were solvated with 45 000 molecules of TIP3P water, and the charge was neutralized as previously described. The system was first minimized in vacuo and in explicit solvent and then equilibrated with the solvent for 200 ps, restraining all of the solute atoms, including the ligand ones, in order to preserve the overall structure of the system and to equilibrate the water around the complex.

The pulling speed was set to 0.0005 nm ps^{-1} and the force to 1200 kJ mol^{-1} nm^{-1}. For each system, 36 or 37 configurations (depending on the protein−ligand distance reached at the end of the simulation) were generated with a maximum distance of up to 7 nm between the centers of mass of the VAC variant and the ligand. Each of them was used as a starting point for a short (150 ps) protein−ligand restrained equilibration followed by a protein-restrained production (1.5 ns) MD simulation with the pulling speed set to zero but keeping the system under the action of a force of the same entity used during the US procedure. The forces resulting from these simulations were analyzed using the weighted histogram analysis method (WHAM),[26] and its final value, when no interactions occur according to the PMF definition, was normalized to zero. The electrostatics, the temperature, and the pressure in the PMF calculations were controlled as in the rest of the MD simulations of the present work.

All of the simulations were performed on a couple of in-house machines equipped with two esacore Intel Xeon processors and 2 Nvidia GTX 780 GPUs for a total of ∼1000 CPU days.

■ RESULTS AND DISCUSSION

Kinetic Properties on CephC Derivatives. The evaluation of the kinetic parameters of different VAC variants on the CephC derivatives was carried out spectrophotometrically employing an assay mixture containing increasing substrate concentrations.

Gl-7-ACA. As shown in Table 1, the highest activity on Gl-7-ACA and kinetic efficiency (expressed by the V_{max}/K_m ratio)

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

were observed for HS-HS-F72βR and wild-type VAC, with all of the other enzyme variants showing lower activities on this substrate.[8,11] It is noteworthy that a substrate inhibition effect was apparent for both of the best VAC variants at substrate concentrations of >6 mM (Figure 2a).

CephC. The kinetic parameters on the most valuable industrial substrate, CephC, which is currently used as the starting compound to produce semisynthetic cephalosporins,[4,7] were determined previously.[8,11] The kinetic efficiency was 9- and 10.8-fold higher for the HS-HS-L154βY and HS-HS-F58βD variants respectively, compared with wild-type VAC. For the HS-HS-L154βY variant, this was largely the result of a higher V_{max} value, whereas for the HS-HS-F58βD variant this was mainly due to an ~6-fold lower K_m value (Table 1). When CephC was used as the substrate, no substrate inhibition effect was apparent for any of the VAC variants tested (Figure 2b).

Succinyl-7-ACA. The kinetic properties of all of the VAC variants under investigation were strongly affected by shortening of the side chain at C7 (see Figure S1b), as observed using succinyl-7-ACA as the substrate. The most significant changes were a 10-fold increase in K_m and a 100-fold decrease in maximal activity for the HS-HS-F72βR and HS-HS-F58βD VAC variants, respectively, compared with the wild-type enzyme (Table 1). Interestingly, because of the substrate inhibition effect observed for wild-type VAC, at substrate concentrations of ≥80 mM the HS-HS-L154βY and M165αS-HS-HS VAC variants showed activity values similar to that of the wild-type counterpart (Figure 2c).

Adipyl-7-ACA. On the substrate adipyl-7-ACA (showing an additional CH_2 group in the side chain at position 7 compared with Gl-7-ACA), all of the VAC variants possessed lower K_m values compared with wild-type VAC as well as lower enzymatic activities (Table 1). The best VAC variants for hydrolysis of adipyl-7-ACA were wild-type VAC and the M165αS-HS-HS and HS-HS-F72βR variants (with the engineered variants showing higher affinities for the substrate). Furthermore, none of the HS-HS VAC variants tested on this substrate showed a substrate inhibition effect, whereas for wild-type VAC inhibition was observed at substrate concentrations of ≥25 mM (Figure 2d).

Pentanoyl-7-ACA. The affinity of the HS-HS-L154βY VAC variant for pentanoyl-7-ACA (lacking the terminal carboxyl group at C7 respect to Gl-7-ACA) was up to 3.5-fold better than those of wild-type VAC and the HS-HS, M165αS-HS-HS, and HS-HS-F72βR variants (Table 1). This resulted in a higher kinetic efficiency value for this VAC variant (up to 5-fold). Furthermore, the M165αS-HS-HS variant showed a 1.7-fold higher V_{max} value compared with wild-type VAC. For all of the enzyme variants tested, no substrate inhibition effect occurred (Figure 2e).

VAC−Ligand Modeling I: Generation of Reliable Starting Poses. To investigate the mode of CephC derivative binding to the active site of wild-type VAC (Figure 1) or different VAC variants and being aware of the importance of the dynamics in shaping the binding (and, subsequently, of the impossibility of correctly performing standard docking calculations on such systems), we adopted a hybrid approach based upon MD simulations and docking calculations in order to generate reliable VAC complexes.

We then carried out 20 ns molecular dynamics (MD) simulations of all of the VAC variants, aimed at identifying reliable structures that could accommodate the various ligands examined in the present work, by also taking into account the mobility and rearrangement of the protein structure. Such simulations were performed starting from the tridimensional structures of wild-type VAC (Figure 1) and the HS-HS variant[10] or the molecular models produced for the M165αS-HS-HS, HS-HS-F58βD, HS-HS-F72βR, and HS-HS-L154βY variants. In all cases, the backbone root-mean-square deviation (RMSD) was equal to or lower than 1.5 Å (see Figure S2). This finding indicates that the resulting structures could be considered stable and generally compatible through all of the simulation with the experimental crystal structures (within the approximation of the force field used and of the fairly good original resolution, 2.13 Å) and that the investigated substitutions are structure-conserving.

We selected a set of structures along the trajectory that could be representative of the binding-site dynamics according to the variation of the overall SAS of the binding pocket, considering their statistical relevance according to a cluster analysis applied to the structure of the region of interest (for details, see Methods). SAS variations along the simulation time (see Figure S3) show the specific effect of single-point mutations on the dynamics of the binding site even without heavy effects on the overall structure (see the RMSDs in Figure S2), ranging from a very stable binding site (M165αS-HS-HS; Figure S3b) to ones whose accessible surface is quite variable (HS-HS-F58βD; Figure S3c). This highlights the importance of dynamics in the selection process of ligand binding by the free protein and the necessity to include this information in the modeling of complexes, focusing on a statistical representation of structures like the one we adopted.

In this respect, we then docked each selected ligand (Gl-7-ACA, CephC, succinyl-7-ACA, adipyl-7-ACA, or pentanoyl-7-ACA) within all of the selected conformations of all of the VAC variants, leading to a set of complexes for each enzyme variant that were ranked according to the AutoDock Vina scoring function (based on the estimated free energy of binding), restraining the distance between atoms involved in the nucleophilic substitution during the formation of the enzyme−substrate complex. On the basis of the AutoDock Vina rankings and the agreement between the computed and experimental apparent binding constants (Figure 3), we generated a family of structures containing one complex for each ligand−VAC variant. We then ran 100 ns MD simulations of some of the complexes generated as described above. In

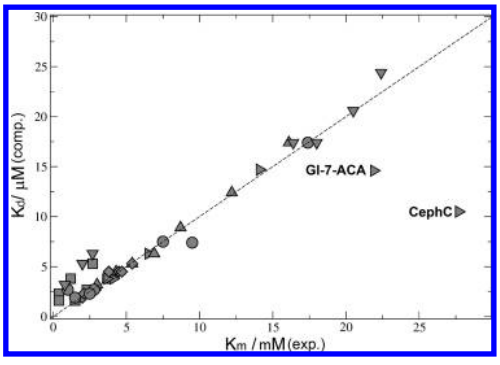


Figure 3. Analysis of the correlation between the experimental (K_m) and theoretical (K_d) dissociation constants for complexes of different VAC variants (wild-type, ●; HS-HS, ▲; M165αS-HS-HS, ▶; HS-HS-F58βD, ■; HS-HS-F72βR, ▼; HS-HS-L154βY, ◆) with all of the available ligands. The dashed line represents the 1:1 correlation between the two sets of values.

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

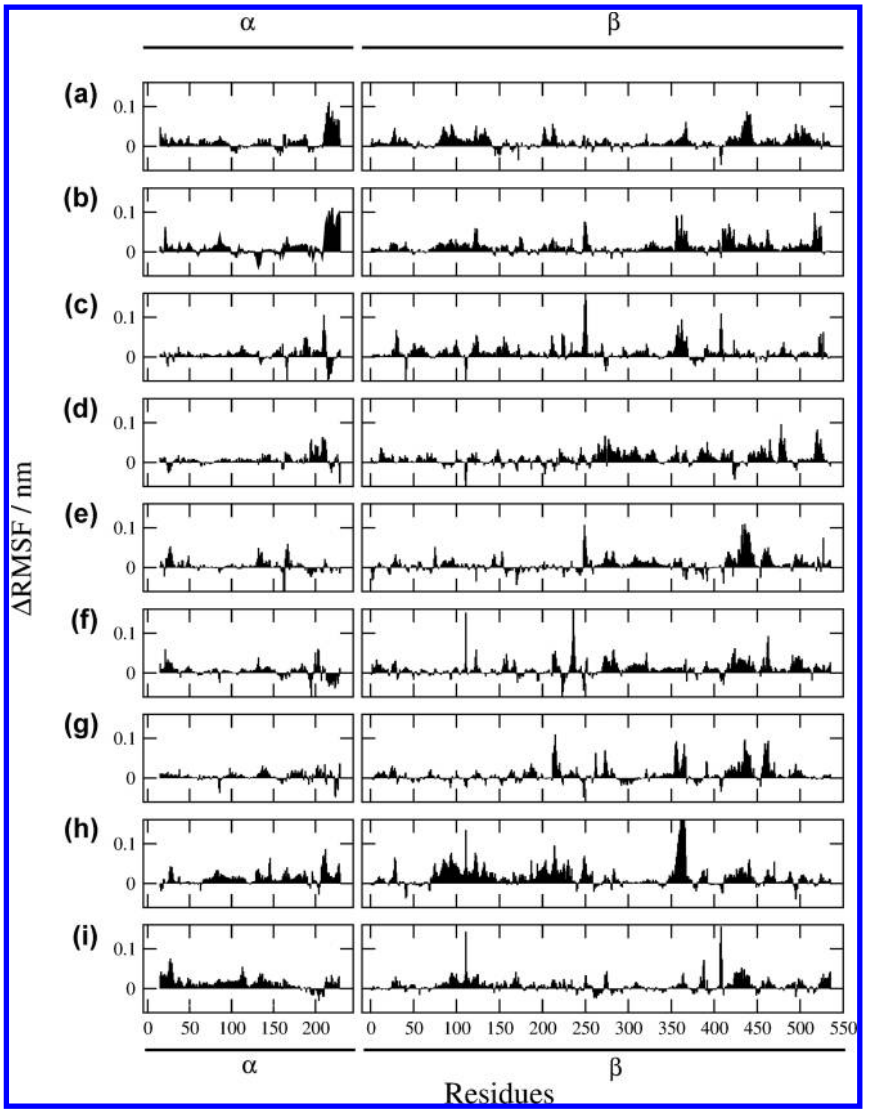


Figure 4. Difference in the RMSFs for free and ligand-bound forms of different VAC variants (ΔRMSF): (a) wild-type VAC with Gl-7-ACA; (b) wild-type VAC with succinyl-7-ACA; (c) HS-HS VAC with CephC; (d) M165αS-HS-HS VAC with adipyl-7-ACA; (e) HS-HS-F58βD VAC with CephC; (f) HS-HS-F72βR VAC with Gl-7-ACA; (g) HS-HS-F72βR VAC with adipyl-7-ACA; (h) HS-HS-L154βY VAC with CephC; (i) HS-HS-L154βY VAC with pentanoyl-7-ACA.

particular, we focused on those systems that coupled reproducibility of the experimental data with higher ligand affinity (Table 1), i.e., systems corresponding to symbols that lie on the dashed line in Figure 3 (representing the ideal theoretical−experimental match) or in its close proximity were selected. Moreover, we selected complexes that experimentally displayed high V_{max}/K_m values, thus representing the ones of major interest from an industrial point of view, as the conversion rate is directly proportional to this ratio. On the basis of these assumptions, the focus on the complexes of the M165αS-HS-HS and HS-HS-F72βR VAC variants with adipyl-7-ACA is mainly due to the higher affinities of those variants for this substrate in comparison with that of wild-type enzyme (see Table 1 and Figure 2).

VAC−Ligand Modeling II: MD Simulations of Enzyme−Substrate Complexes. The RMSD of the backbone in the MD simulations with respect to the crystal structures never displayed values greater than 2 Å (see Figure S4), hence assuring that the overall protein fold was well-conserved along the simulations.

The difference between the free and bound VAC root-mean-square fluctuations (RMSFs) of each residue's α-carbon (ΔRMSF) is reported in Figure 4. In the majority of the reported cases, the ΔRMSF is spread everywhere and its magnitude is small, indicating that the system is not perturbed to a significant extent by the presence of the ligand. In some cases, the structure seems to give a localized response to the binding, with enhanced dynamics of specific regions that act as a dynamic signature of the bound state. Such an evident alteration of the protein dynamics is clearly visible for wild-type VAC in the C-terminal region of the α-chain (Figure 4a,b), where the helical structure dynamics is increased as a result of the presence of different ligands, without however altering the global fold (and stability) of the protein. Also, upon CephC binding the HS-HS-L154βY VAC variant displays an extended augmented mobility in the central region of the β-chain centered around residue 360β (Figure 4h), and a common dynamics signature (435β−445β) is shared by two distinct complexes of the HS-HS-F72βR VAC variant (Figure 4f,g).

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

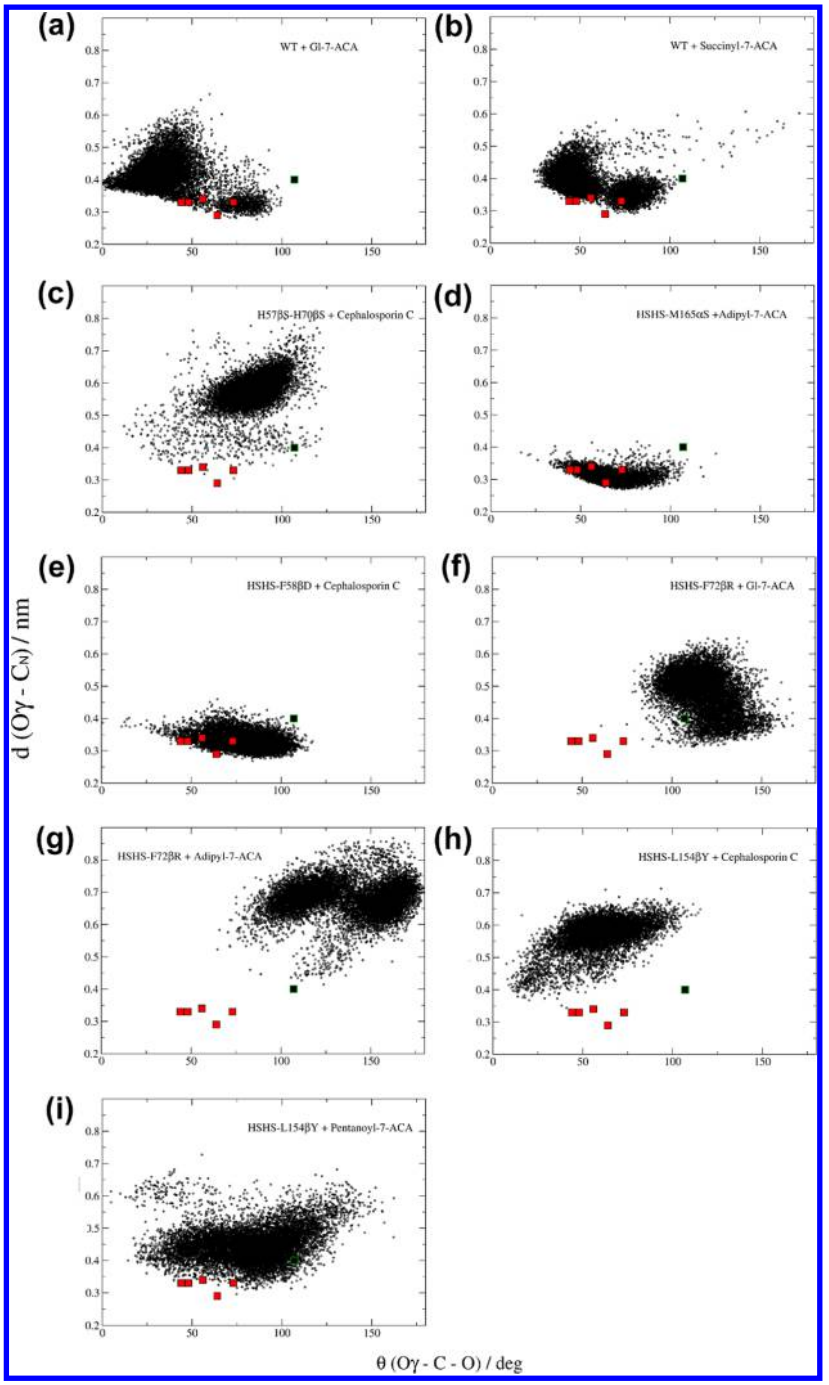


Figure 5. Distributions of distances between the hydroxyl oxygen atom of Ser1β (Oγ) and the acceptor carbonyl carbon (C_N) (see Figure S1) vs the Oγ−C_N−O_N angle. The red squares correspond to values computed for the PDB-deposited structures of complexes of penicillin (PDB IDs 1GM8, 1GM9,[28] and 1FXV[29]) and cephalosporin (PDB ID 1JVZ[30]) acylases. In each panel, the black square surrounded by a green border refers to the couple of values corresponding to the standard Bürgi−Dunitz geometry (see the text for details). The BD angle and distance variations during the simulation time are reported in Figure S5.

These results are in agreement with a number of experimental data showing that amino acid substitutions far from the active site (see Figure 1b) can synergistically generate small structural changes affecting the dynamics (and/or stability) of the protein in a way that enhances substrate binding or subsequently catalytic turnover.[9,27]

VAC−Ligand Modeling III: Interactions within the Binding Site. Despite the interest of the scientific community in acylases in terms of both kinetics and structure, few studies investigating the features of the encounter complex (EC) and/or transition state (TS) in such systems at the atomic level have been reported.[28−30] Even more interestingly, one almost unique feature of acylases in terms of stereochemistry of the formation of the EC is the absence (for the system whose enzyme−substrate structure is known) of a geometry fully compatible with the nucleophilic attack mechanism, which has been widely recognized for this enzyme family.[28−30] Such a mechanism requires a distance of 2−4 Å between the attacking hydroxyl oxygen atom (Oγ) of Ser1β of the enzyme and the acceptor carbonyl carbon (C_N) of the substrate (Figure S1) and a value

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

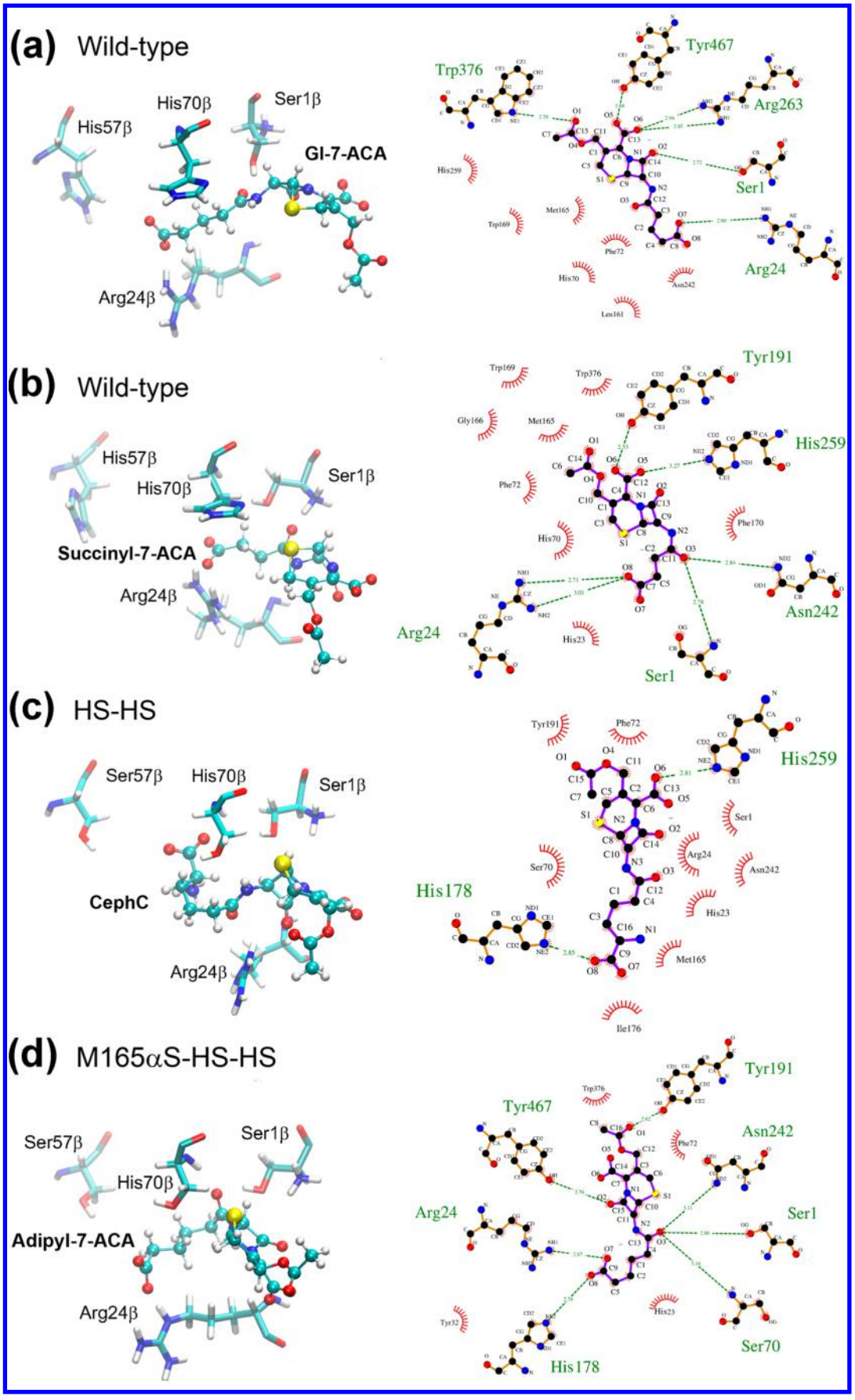


Figure 6. continued

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

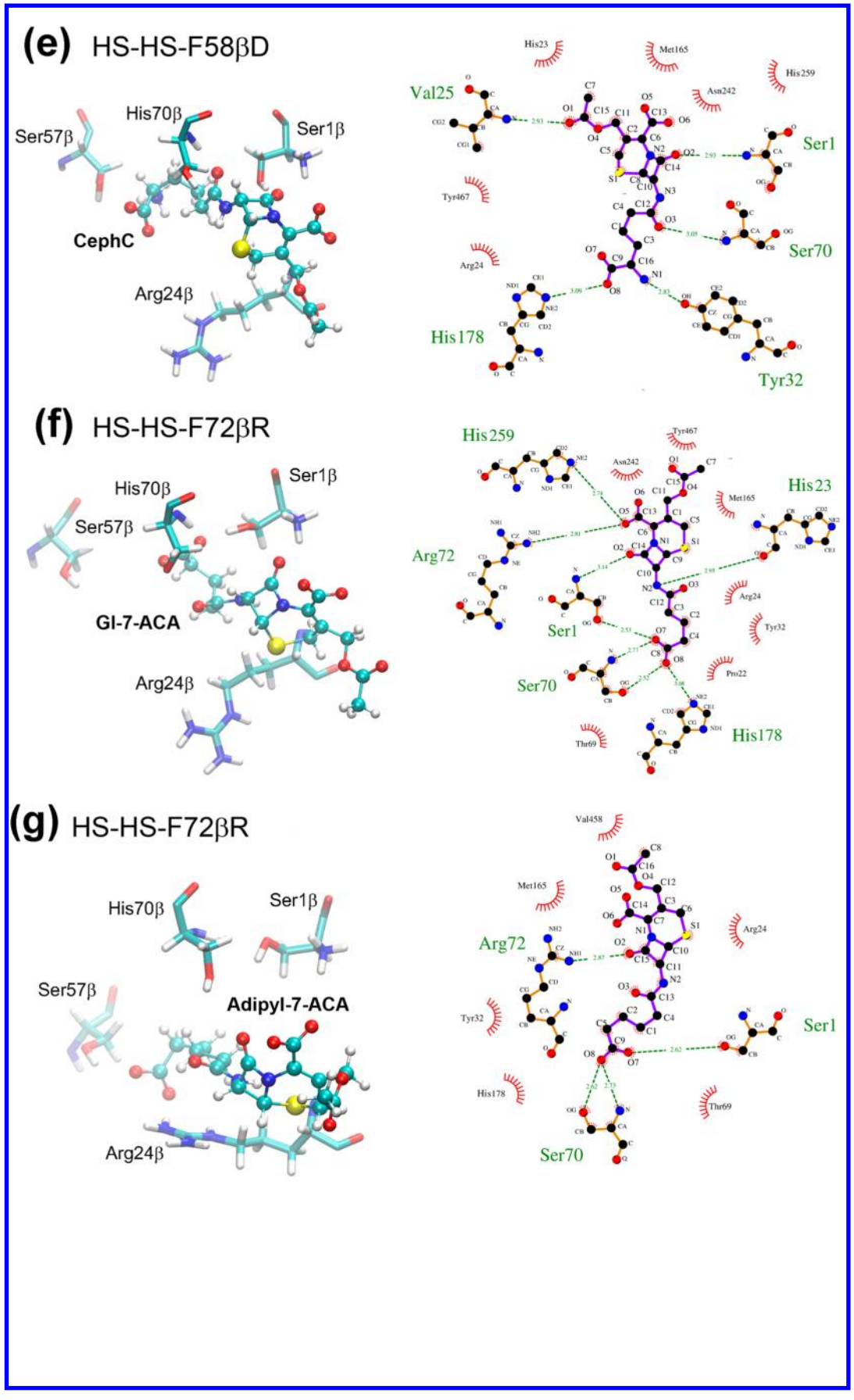


Figure 6. continued

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

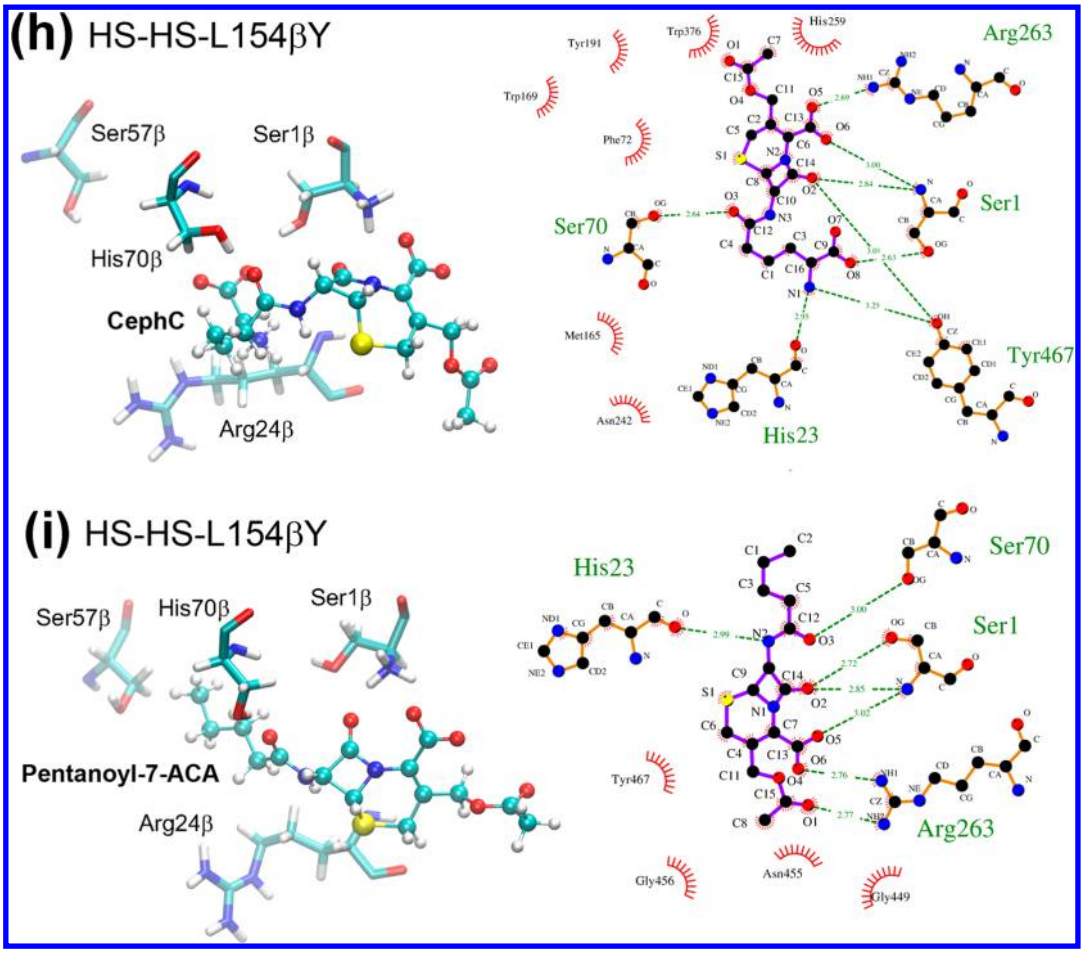


Figure 6. (left) Binding sites of different VAC complexes and (right) schematic representations of the interactions between ligands and active-site residues (potential hydrogen bonds are represented as dotted lines; potential hydrophobic interactions are represented as red arcs with spikes; hydrogen and oxygen atoms are represented in blue and red, respectively). (a, b) Wild-type VAC with Gl-7-ACA and succinyl-7-ACA. (c) HS-HS VAC with CephC. (d) M165αS-HS-HS VAC with adipyl-7-ACA. (e) HS-HS-F58βD VAC with CephC. (f, g) HS-HS-F72βR VAC with Gl-7-ACA and adipyl-7-ACA. (h, i) HS-HS-L154βY VAC with CephC and pentanoyl-7-ACA.

of around 107° for the angle between the carbonyl group and the attacking atom (Oγ−C$_N$−O$_N$, also known as the Bürgi−Dunitz (BD) angle; see Figure S1).[31] Although a few structures of ECs of acylases with the substrates have been solved, it seems quite clear that for both penicillin (PDB IDs 1GM8, 1GM9,[28] and 1FXV[29]) and cephalosporin (PDB ID 1JVZ[30]) acylases, the serine residue responsible for the nucleophilic attack is oriented with its side chain pointing in the opposite direction with respect to the BD geometry, with an angle of 44−73° and an Oγ−C$_N$ distance of 2.9−3.8 Å. To date, no experimental evidence has been reported that could explain such a geometry (to which we will refer in the following as the anti-Bürgi−Dunitz (ABD) geometry) and how it could reach a favorable geometry for the nucleophilic attack on the carbonyl moiety.

i. Wild-Type VAC. The wild-type complexes with Gl-7-ACA and succinyl-7-ACA display the ABD geometry and a favorable distance (~3.5 Å) (Figure 5a,b and Figure 6a,b) and an experimental ratio of their K_m values of ~2 and of their V_{max}/K_m ratios of ~3 (Table 1). This suggests that despite the ideality of the EC geometry during a nucleophilic attack, it does not fully explain the kinetic efficiency of an enzyme, which strongly depends on the nature of the binding forces, i.e., the ones that determine the value of K_S (the ratio of the substrate dissociation and association rate constants, often resembling K_m) and that depend on the nature of the protein−ligand interactions.

The Arg24β side chain stabilizes the interactions with both ligands, i.e., with the carboxylic terminal groups of Gl-7-ACA and succinyl-7-ACA (Figure 6a,b). Interestingly, the Ser1β and Asn242β residues appear to be together directed toward the carbonyl group of succinyl-7-ACA that acts as an acceptor during the reaction. This suggests a potential competition between Oγ of Ser1β and Oδ of Asn242β in the nucleophilic attack, together playing a synergistic role in the electrostatic stabilization of the EC, thus partially explaining the lower catalytic efficiency of wild-type VAC in the hydrolysis of succinyl-7-ACA than that of Gl-7-ACA. In the latter, indeed, Ser1β seems to be harbored close to the configuration able to promote the nucleophilic attack and in the ABD configuration, with the Trp376β, Tyr467β, and Arg263β residues stabilizing the orientation of the lactam scaffold before the reaction occurs (Figure 6a,b). It is noteworthy that in this case Ser1β explores less conformational space than in the complex with succinyl-7-ACA (Figure 5a,b). The same kind of stabilization is also performed for succinyl-7-ACA by aromatic residues by means of their polar/charged portions: as a general rule, aromatic residues in wild-type VAC optimize the active-site pocket to accommodate the ligands by balancing the apolar and Coulombic interactions.

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

Table 2. Hydrogen Bonds Formed between the Ligands and Different VAC Variants Averaged over the Last 50 ns for the Most Relevant Residues within the Binding Site. (a) Minimum and Maximum Numbers of Hydrogen Bonds for Every Relevant Amino Acid in the Binding Site; (b) Average Numbers of Hydrogen Bonds Formed between the Ligands and All of the Residues That Constitute the Binding Site (i.e., Residues within 6 Å of the Surface of the Ligand), along with the Standard Deviations and Maximum and Minimum Values (The Variation of the Number of Ligand-Binding-Site Hydrogen Bonds along the Simulations Is Reported in Figure S6)

	wild-type		HS-HS	M165aS-HS-HS	HS-HS-F58βD	HS-HS-F72βR		HS-HS-L154βY	
residue	Gl-7-ACA	succinyl-7-ACA	CephC	adipyl-7-ACA	CephC	adipyl-7-ACA	Gl-7-ACA	CephC	pentanoyl-7-ACA
				(a) Minimum and Maximum Numbers of Hydrogen Bonds					
191α		1	1−2	1−2	0−1			1−2	1−2
1β	1	1−2		1	1−2	1−2	1−2	1	1
23β				1−2			1		
24β	1−2	1		1		0−1			
32β			0−1		0−1				
57β				1	0−1				
70β		1	0−1		0−1	0−2	1−2		
72β			1−2	1−2		1−2	1−2		
178β					1				
242β			0−1		0−1		1		
259β							1−2		
263β	0−1	0−1						1−2	1−2
376β	0−1								
467β	1	1		1−2				1−2	
	(b) Average Numbers of Hydrogen Bonds Formed between the Ligands and All of the Residues That Constitute the Binding Site								
Avg	4.7	5.5	2.0	6.4	4.5	6.7	6.5	4.0	3.8
SD	0.8	1.2	1.0	1.0	1.1	1.2	1.1	1.5	1.0
Max	8	9	6	10	8	12	11	9	7
Min	2	1	0	2	1	3	3	0	0

ii. HS-HS-F72βR VAC. The HS-HS-F72βR VAC variant is the only one that adopts a proper BD angle configuration (Figure 5f,g), which is centered around 117° for Gl-7-ACA and around 134° for adipyl-7-ACA but with a rather large spread around the average value, displaying a rather bimodal distribution and lying for a large part of the time in line with the C−O axis. The distance is a bit too large with respect to that compatible with a nucleophilic attack (4.5 Å); however, the interaction of the Ser1β side chain with the carboxylic moiety of adipyl-7-ACA or Gl-7-ACA favors the orientation of the Oγ atom for a good nucleophilic attack (Figure 6f,g). The distance between the atomic species involved in this complex can in principle be correlated to the probability of observing the reaction; in this way, the difference observed for the two ligands partially explains the lower V_{max}/K_m ratio observed experimentally for adipyl-7-ACA as the substrate (see Table 1). Indeed, Ser1β also explores less conformational space with adipyl-7-ACA than in the complex with Gl-7-ACA (Figure 5f,g), narrowing the geometric distribution of Oγ of Ser1β with respect to the ligand carbonyl group that constitutes the target of the nucleophilic attack. This increases the probability of placing the reacting chemical species in a favorable position, hence potentially influencing the reaction kinetics by favoring the formation of the Michaelis complex (see Figure S5f,g).

The guanidinium moiety of the arginine side chain at position 72β plays an important role in the ligand stabilization by means of the electrostatic interaction with both the charged carboxylic group at C4 and the dipolar carbonylic group at C3 of the lactam moiety of the ligand (see Figure 6f,g and Figure S1b), which together form a strong electrostatic signature of the family of compounds studied here. In both cases, a Coulombic interaction also drives the stabilization of the glutaryl and adipyl functional groups by means of Ser1β and Ser70β. The substrate binding site is rich in His residues (23β, 178β, 259β) that are prone to stabilize Gl-7-ACA by polar interactions, thus explaining the slightly higher preference (i.e., in terms of K_m; see Table 1) of HS-HS-F72βR for this compound (Figure 6f). Interestingly, in a one-pot system to directly convert CephC into 7-ACA, the HS-HS-F72βR VAC variant was the best choice when employed in combination with DAAO; this is in agreement with the high enzymatic activity on Gl-7-ACA (see Table 1) and the absence of product inhibition effects.[5] Under optimized conditions and with addition of further aliquots of the biocatalysts, >98% of CephC was converted, pointing to 7-ACA as the main reaction product. It is noteworthy that HS-HS-F72βR is the VAC variant that has the ability on average to maximize the number of hydrogen bonds formed with all of the considered substrates (see Table 2 and Figure S6f,g).

iii. HS-HS-L154βY VAC. To date, the HS-HS-L154βY VAC variant is the most promising enzyme variant for CephC bioconversion because it has the highest V_{max} (Table 1). Thus, great interest has been focused on understanding how it works and harbors alternative ligands. This is also of main interest because of the distance of residue 154β from the first shell of residues that form the binding site (see Figure 1b), suggesting that despite the fact that the structure of the protein remains unaltered upon substitution of L154β, the larger side chain alters the transmission of motion through the protein structure, therefore influencing the enzymatic kinetics.

The distance between the Ser1β Oγ atom and the C_N atom in HS-HS-L154βY complexes is always between 4.5 and 5.5 Å, which is more disfavorable for the nucleophilic attack than in the other complexes, thus justifying, at least partially, the lower V_{max}/K_m ratio (Table 1). One peculiarity is that the observed ability of the ligand to explore BD- and ABD-type orientations is ligand-dependent, with CephC experiencing rearrangements

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

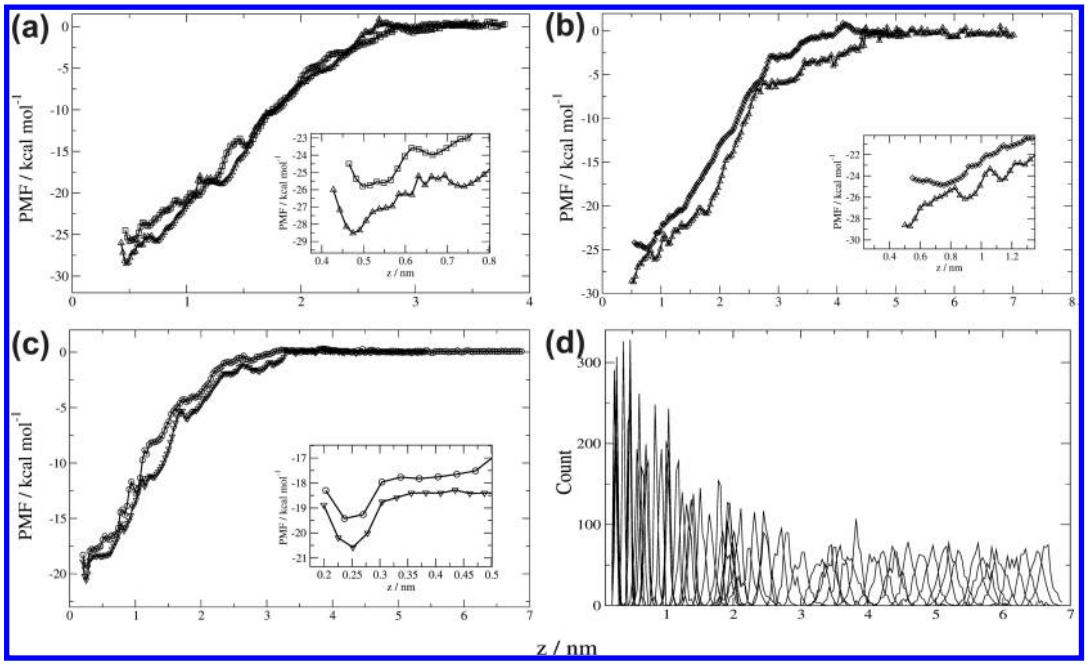


Figure 7. (a–c) Potential of mean force curves for different VAC complexes with Gl-7-ACA (△), CephC (○), succinyl-7-ACA (□), adipyl-7-ACA (◇), and pentanoyl-7-ACA (▽): (a) wild-type VAC; (b) HS-HS-F72βR VAC variant; (c) HS-HS-L154βY VAC variant. The same curves for the HS-HS, M165αS-HS-HS, and HS-HS-F58βD VAC variants are reported in Figure S8. (d) WHAM histograms along the dissociation coordinate z for the HS-HS-L154βY VAC variant in complex with CephC. The histograms for all of the PMF profiles are reported in Figure S10. Figure S11 reports the convergence test for each calculation.

in the binding site that never orient the Oγ atom in a BD-type orientation, whereas in the case of pentanoyl-7-ACA both orientations are visited and even with a shorter distance (Figure 5h,i). Interestingly, during the simulation time the distance between the Oγ and C_N atoms in the HS-HS-L154βY complex with pentanoyl-7-ACA (see Figure S5i) remains almost constant, while the Oγ–C_N–O_N angle explores a wide conformational space, making it actually responsible for different behavior with respect to the other complex. For both ligands, the triad formed by Ser1β, His23β, and Ser70β is positioned in almost identically, favoring the ABD geometry and conferring on Ser1β an important catalytic role, with the His23β making a hydrogen bond with the amide bond of the cephalosporanic moiety (Figure 6h,i). A peculiar role is played by Tyr467β, whose interactions with both the cephalosporanic and aminoadipic moieties of CephC are able to keep the ligand in a closer configuration than usual.

iv. HS-HS VAC. In terms of EC geometry and kinetic parameters, the worst and clearest scenario among the systems investigated is represented by the HS-HS VAC variant with CephC. This displays a rather low substrate affinity ($K_m \approx 12$ mM; see Table 1), the ability of residue Ser1β to explore all of the possible angles, thus allowing both the BD- and ABD-type geometry, and a long distance of the Oγ atom of the Ser1β residue from the stereocenter (Figure 5c). Altogether, these issues dramatically decrease the probability of reaction and of TS complex formation, confirming the observed lowest kinetic efficiency among all of the VAC–ligand complexes investigated (up to 10-fold lower compared with VAC complexes with Gl-7-ACA; Table 1).

Tyr32β, His178β, and His259β are the residues responsible for anchoring the ligand to the binding site harboring the two extreme carboxylic units that are present in the lactam and the carboxypentanoyl moieties (Figure 6c and Figure S1). Moreover, the interactions of Arg24β with the ligand are not constant, with an average lifetime of ∼5 ns and a continuous formation and disruption of a hydrogen bond between the side-chain imino group and the carbonyl target (C_N) of the nucleophilic attack. Interestingly, the acylase evolutionary analysis suggested that this residue is the most variable one among active-site residues, thus highlighting its importance for the evolution of VAC substrate specificity; this conclusion is also confirmed by the high number of variants at this position isolated by protein engineering studies.[11]

v. M165αS-HS-HS and HS-HS-F58βD VAC Variants. The HS-HS VAC variants containing the M165αS or F58βD substitutions display very clear statistics in terms of distances and angular geometry that perfectly match the known experimentally determined orientation of Ser1β with respect to the carbonyl carbon in similar species (Figure 5d,e). In the substrate-bound forms of these two VAC variants, His178β has the same anchoring role played in the HS-HS and HS-HS-F72βR variants in the interactions between the ligand and the protein, i.e., stabilizing with its aromatic ring the free charge of the terminal carboxylic moiety shared by CephC and Gl-7-ACA, thus constituting a typical signature of the system in the formation of the EC (Figure 6d). In both cases, one residue responsible for anchoring the aminocephalosporanic moiety of the ligand forms hydrogen bonds, with the notable difference of involving the aromatic side chain of Tyr191α in M165αS-HS-HS and the backbone amide group of Val25β in HS-HS-F58βD (Figure 6d,e). In both VAC complexes, Ser1β and Ser70β are in favorable positions for the nucleophilic attack, with a slightly more favorable position in the M165αS-HS-HS VAC variant. One major difference is represented by the interactions of Tyr32β side chains with the terminal groups of adipyl-7-ACA and CephC, respectively, that at least partially explains the difference in the substrate affinities of the two VAC variants (a

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227–2241

3-fold difference in K_m values was observed; see Table 1 and Figure 6d,e).

Interestingly, despite the moderate experimentally measured affinity of M165αS-HS-HS for adipyl-7-ACA as a substrate (K_m = 4.5 mM), this complex is the one that, together with HS-HS-F72βR, displays the highest number of protein−ligand hydrogen bonds (see Table 2 and Figure S6).

Potentials of Mean Force for the VAC Variant Complexes. MD simulations can describe the atomic-level characteristics of the ligand interactions in the binding pocket of the biological counterpart, thus providing information that may be useful for different purposes. However, this kind of information can hardly be used for quantitative comparison with experimental data, with a few exceptions, including NMR observables, B factors, SAXS curves, etc.

To provide a description of the energetic profile of the interactions, we performed a calculation of the PMF for unbinding from the Michaelis complex as obtained from MD simulations (i.e., on the structure obtained as an average of the frames recorded in the last portion of the simulations). The calculations were performed on all of the systems by means of umbrella sampling, extracting the ligand from the binding site and applying a force constant to the ligand center of mass along an axis (z) that starts from the center of mass of Ser1β and crosses the entrance of the VAC binding cavity in order to avoid unnatural clashes (see Figure S7).

Bearing in mind that the depth of the PMF profile (ΔPMF = $PMF_{max} - PMF_{min}$) constitutes the major contributor to the free energy of binding,[32] we can assume from all of the PMF profiles that in the Michaelis complexes all of the ligands were in a free energy minimum (Figure 7 and Figure S8), showing that the structures averaged over the final parts of the MD trajectories could actually be considered reliable minima and were good starting points for this kind of simulation.

Ideally, it would be desirable to perform a comparison of all of the free energy profiles to rank the systems and to compare them with the experimental K_m values, thus generating an in silico method for screening these systems and getting the best conditions for Michaelis complex formation.

However, some relevant aspects discourage a free-energy-based treatment of the binding and of its comparison for all of the systems considered in the present work: (i) Bearing in mind that we are discussing enzymatic systems, a comparison of the dissociation constant K_d, the quantity directly obtainable from ΔG, with K_m would be difficult because $K_m \approx K_d$ if $k_{cat} \ll k_{-1}$, but otherwise K_m corresponds to the upper limit of the value that K_d can assume. (ii) It is rather difficult to handle the definition of the standard state of the protein, the nonuniqueness of the unbinding path, and the weight of conformational flexibility of ligands in the calculation.[32,33] (iii) Despite the low speed and low force applied to the ligand to simulate its unbinding, an unquantifiable amount of irreversible physical work (mainly due to frictional forces and to the limited sampling of the unbinding path) is always associated with this kind of simulation.[34] (iv) The systems are very similar, but obvious differences are present in terms of chemical reactivity, geometry, and dynamics.

Nevertheless, if we consider a comparison between the complexes formed by the same VAC variant with two ligands, the PMF curve difference between the two systems ($\Delta\Delta$PMF) could be considered reliable within the error associated with energy numerical evaluation (~5%; bars not shown in Figure 7 and Figure S8). This allowed us to evaluate systems where all of the aforementioned effects/issues are comparable and hence negligible when differences are considered.

In this context, the relative ranking within couples of ligands binding to wild-type VAC or the HS-HS-F72βR or HS-HS-L154βY VAC variant (Figure 7) is in perfect agreement with the one based on the experimentally obtained K_m values, i.e., the lowest K_m always corresponds to the biggest ΔPMF. All of the $\Delta\Delta$PMF values are in the 1−4 kcal mol^{-1} range (see Table 3).

Table 3. Comparisons of the Experimental K_m Values and Calculated ΔPMF Values[a] for All of the Complexes Analyzed in the Present Work

system	K_m/mM	ΔPMF/kcal mol^{-1}
wild-type−Gl-7-ACA	1.5	28.6
wild-type−succinyl-7-ACA	2.8	25.9
HS-HS-F72βR−Gl-7-ACA	2.0	28.8
HS-HS-F72βR−adipyl-7-ACA	2.7	24.9
HS-HS-L154βY−pentanoyl-7-ACA	3.8	20.6
HS-HS-L154βY−CephC	4.5	19.5

[a] $\Delta PMF = PMF_{max} - PMF_{min}$.

Interestingly, the difference displayed by the energetic content of the interaction between pentanoyl-7-ACA or CephC and HS-HS-L154βY (Figure 7c) can be partially ascribed to the difference in water contents of the binding sites: both the radial distribution function (RDF) of water around the ligand (see Figure S9a) and the number of hydrogen bonds formed by the solute and solvent molecules (Table 2 and Figure S9b) indicate that the solvent in HS-HS-L154βY could have a role in influencing the formation/stabilization of the encounter complex. In this case, CephC in the binding cavity forms twice the number of hydrogen bonds with water molecules as formed by pentanoyl-7-ACA with the solvent (Table 2 and Figure S9b).

It is noteworthy that the ΔPMF value for the complex formed by M165αS-HS-HS with adipyl-7-ACA (Figure S8b) is rather large compared with those for the other complexes. In light of the large number of hydrogen bonds formed by the ligand with the binding site (Table 2 and Figure S6d) and the nature of the K_m value explained above, this suggests for this specific case the presence of a strong kinetic bias in the conversion of the reagent to product.

CONCLUSIONS

We have demonstrated the useful combination of protein dynamics and investigation of substrate specificity of some VAC variants toward different CephC derivatives to provide a detailed atomic- and residue-level description of the encounter complexes, the necessary step that needs to be controlled and finely tuned in order to get the highest efficiency in terms of enzymatic bioconversion. In particular, we coupled a starting docking procedure with a dynamic exploration of the conformational space of the system: the former step matched the experimental values, and the latter step allowed the adaptation of ligand poses into the protein binding cavity on the basis of intermolecular and intramolecular forces computed over time. Moreover, umbrella sampling simulations allowed the evaluation of the relative energetic stability of the average structures obtained by means of MD simulations, thus further demonstrating that the overall procedure provides results in good agreement with the experimental observations.

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

A deeper understanding of the binding mode is desirable with respect to both basic science and industrial applications. In the future, this could allow the identification of further putative amino acid substitutions to evolve new VAC variants with improved properties in the bioconversion of additional substrates of industrial interest.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00535.

Schematic representation of nucleophilic attack, backbone RMSDs along the simulations for the free VAC variants, SAS variations of the binding site in the free protein along the simulations for the free VAC variants, backbone RMSDs along the simulations for the VAC variants in complex with various substrates, variation along the simulations of the geometric parameters that define the BD and ABD geometries, variation of the number of hydrogen bonds between the ligand and the binding site for each complex along the simulations, top view of the VAC binding site, additional PMF curves for different VAC−ligand complexes, RDF of water in the binding cavity of HS-HS-L154βY VAC and the corresponding formation of hydrogen bonds between CephC or pentanoyl-7-ACA and the water over the whole simulation, WHAM histograms of the PMF calculations, and convergence of the PMF profiles (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E.R.: Fax: +39 0332421500. Tel.: +39 0223993062. E-mail: elena.rosini@uninsubria.it.

*L.M.: Fax: +39 0107170187. Tel.: +39 01071781960. E-mail: luca.mollica@iit.it.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grant from Fondo di Ateneo per la Ricerca to L.P. (Università degli studi dell'Insubria). We are grateful to Gianluca Tomasello for help in docking analyses.

■ REFERENCES

(1) Pollegioni, L.; Molla, G. New Biotech Applications from Evolved D-Amino Acid Oxidases. *Trends Biotechnol.* **2011**, *29*, 276−283.

(2) Niu, M. M.; Dong, F. G.; Tang, S.; Fida, G.; Qin, J. Y.; Qiu, J. D.; Liu, K. B.; Gao, W. D.; Gu, Y. Q. Pharmacophore Modeling and Virtual Screening for the Discovery of New Type 4 cAMP Phosphodiesterase (PDE4) Inhibitors. *PLoS One* **2013**, *8*, e82360.

(3) Yang, S. Y. Pharmacophore Modeling and Applications in Drug Discovery: Challenges and Recent Advances. *Drug Discovery Today* **2010**, *15*, 444−450.

(4) Pilone, M. S.; Pollegioni, L. D-Amino Acid Oxidase as an Industrial Biocatalyst. *Biocatal. Biotransform.* **2002**, *20*, 145−159.

(5) Conti, G.; Pollegioni, L.; Rosini, E. One-Pot Conversion of Cephalosporin C by Using an Optimized Two-Enzyme Process. *Catal. Sci. Technol.* **2015**, *5*, 1854−1863.

(6) Elander, R. P. Industrial Production of Beta-Lactam Antibiotics. *Appl. Microbiol. Biotechnol.* **2003**, *61*, 385−392.

(7) Pollegioni, L.; Rosini, E.; Molla, G. Cephalosporin C Acylase: Dream and(/or) Reality. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 2341−2355.

(8) Rosini, E.; Monelli, C. S.; Pollegioni, L.; Riva, S.; Monti, D. On the Substrate Preference of Glutaryl Acylases. *J. Mol. Catal. B: Enzym.* **2012**, *76*, 52−58.

(9) Pollegioni, L.; Lorenzi, S.; Rosini, E.; Marcone, G. L.; Molla, G.; Verga, R.; Cabri, W.; Pilone, M. S. Evolution of an Acylase Active on Cephalosporin C. *Protein Sci.* **2005**, *14*, 3064−3076.

(10) Golden, E.; Paterson, R.; Tie, W. J.; Anandan, A.; Flematti, G.; Molla, G.; Rosini, E.; Pollegioni, L.; Vrielink, A. Structure of a Class III Engineered Cephalosporin Acylase: Comparisons with Class I Acylase and Implications for Differences in Substrate Specificity and Catalytic Activity. *Biochem. J.* **2013**, *451*, 217−226.

(11) Conti, G.; Pollegioni, L.; Molla, G.; Rosini, E. Strategic Manipulation of an Industrial Biocatalyst-Evolution of a Cephalosporin C Acylase. *FEBS J.* **2014**, *281*, 2443−2455.

(12) Orozco, M. A Theoretical View of Protein Dynamics. *Chem. Soc. Rev.* **2014**, *43*, 5051−5066.

(13) Widderich, N.; Pittelkow, M.; Hoppner, A.; Mulnaes, D.; Buckel, W.; Gohlke, H.; Smits, S. H.; Bremer, E. Molecular Dynamics Simulations and Structure-Guided Mutagenesis Provide Insight into the Architecture of the Catalytic Core of the Ectoine Hydroxylase. *J. Mol. Biol.* **2014**, *426*, 586−600.

(14) Balasingham, K.; Warburton, D.; Dunnill, P.; Lilly, M. D. The Isolation and Kinetics of Penicillin Amidase from *Escherichia coli*. *Biochim. Biophys. Acta* **1972**, *276*, 250−256.

(15) Trott, O.; Olson, A. J. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455−461.

(16) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(17) Darden, T.; Perera, L.; Li, L.; Pedersen, L. New Tricks for Modellers from the Crystallography Toolkit: the Particle Mesh Ewald Algorithm and its Use in Nucleic Acid Simulations. *Structure* **1999**, *7*, R55−R60.

(18) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(19) Jarvis, R. A.; Patrick, E. A. Clustering Using a Similarity Measure Based on Shared Near Neighbors. *IEEE Trans. Comput.* **1973**, *C22*, 1025−1034.

(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.

(21) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *Ab initio* Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

(22) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: the RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(23) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.;

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241

Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. *AMBER 14*; University of California: San Francisco, 2014.

(24) Mollica, L.; Decherchi, S.; Zia, S. R.; Gaspari, R.; Cavalli, A.; Rocchia, W. Kinetics of Protein-Ligand Unbinding Via Smoothed Potential Molecular Dynamics Simulations. *Sci. Rep.* **2015**, *5*, 11539.

(25) Kästner, J. Umbrella Sampling. *WIREs Comput. Mol. Sci.* **2011**, *1*, 932−942.

(26) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(27) Sacchi, S.; Rosini, E.; Molla, G.; Pilone, M. S.; Pollegioni, L. Modulating D-Amino Acid Oxidase Substrate Specifity: Production of an Enzyme for Analytical Determination of All D-Amino Acids by Directed Evolution. *Protein Eng., Des. Sel.* **2004**, *17*, 517−525.

(28) Alkema, W. B.; Hensgens, C. M.; Kroezinga, E. H.; de Vries, E.; Floris, R.; van der Laan, J. M.; Dijkstra, B. W.; Janssen, D. B. Characterization of the Beta-Lactam Binding Site of Penicillin Acylase of *Escherichia coli* by Structural and Site-Directed Mutagenesis Studies. *Protein Eng., Des. Sel.* **2000**, *13*, 857−863.

(29) McVey, C. E.; Walsh, M. A.; Dodson, G. G.; Wilson, K. S.; Brannigan, J. A. Crystal Structures of Penicillin Acylase Enzyme-Substrate Complexes: Structural Insights into the Catalytic Mechanism. *J. Mol. Biol.* **2001**, *313*, 139−150.

(30) Kim, Y.; Hol, W. G. Structure of Cephalosporin Acylase in Complex with Glutaryl-7-aminocephalosporanic Acid and Glutarate: Insight into the Basis of its Substrate Specificity. *Chem. Biol.* **2001**, *8*, 1253−1264.

(31) Bürgi, H. B.; Dunitz, J. D.; Lehn, J. M.; Wipff, G. Stereochemistry of Reaction Paths at Carbonyl Centres. *Tetrahedron* **1974**, *30*, 1563−1572.

(32) Woo, H. J.; Roux, B. Calculation of Absolute Protein-Ligand Binding Free Energy from Computer Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6825−6830.

(33) Deng, Y.; Roux, B. Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. *J. Phys. Chem. B* **2009**, *113*, 2234−2246.

(34) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. Free Energy Calculation from Steered Molecular Dynamics Simulations Using Jarzynski's Equality. *J. Chem. Phys.* **2003**, *119*, 3559−3566.

DOI: 10.1021/acs.jcim.5b00535
J. Chem. Inf. Model. 2015, 55, 2227−2241