Modeling and docking of cephalosporin acylase for evaluation of the computational methods

Yasuhiro ISOGAI, Yurie SUNAKAWA, Masato YAMADA, and Seiji HASHIMOTO

Department of Biotechnology, Faculty of Engineering

Summary

The antibiotic acylases belonging to the N-terminal nucleophile hydrolase superfamily are key enzymes for the industrial production of antibiotic drugs, and are targets for engineering to improve their stabilities and catalytic activities. Cephalosporin acylase (CA) are one of the most intensively studied enzyme catalyzing the deacylation of β-lactam antibiotics. In this study, the substrate binding by native CA and its mutants was simulated, based on their X-ray crystallographic and computationally-constructed mutant structures. The docking energies that were obtained by a simple computational method correlated well with their experimentally determined K_m values. The present study demonstrates the usefulness of computational approaches for understanding and engineering the enzymatic properties of antibiotic acylases.

Keywords: antibiotics, β lactam, acylase, docking simulation, substrate selection

1. Introduction

Since Alexander Fleming discovered penicillin [1], the first known natural antibiotic produced by fungi, many antibiotics that kill microorganisms or inhibit their growth have been found and developed to produce medicines, by chemical modification of the original natural compounds. Among the many antibiotic classes, β-lactam-related compounds containing penicillin and cephalosporin are the most commonly utilized in medicine (Fig. 1) [2, 3]. Modifications of the original compounds produced by microorganisms are essential to manufacture antibiotic drugs [4]. Several β-lactam acylases, which are members of the N-terminal nucleophile (Ntn) hydrolase superfamily [5], are utilized to remove the acyl moiety of these compounds to obtain the common β-lactam nuclei under non-hazardous, mild conditions, from which various antibiotic drugs are semi-synthesized [6]. Cephalosporin acylase (CA), or glutaryl-7-aminocephalosporanic-acid acylase (EC 3.5.1.93), and penicillin G acylase (PGA) (EC 3.5.1.11) are the most intensively studied enzymes, and an abundance of kinetic and other biochemical data have been accumulated [7-13]. In addition, the X-ray crystallographic structures of CA and PGA have been solved in various forms, with and without substrate ligands [14-20].

CA and PGA are initially synthesized *in vivo* as inactive precursors, composed of a single polypeptide with a molecular mass (*Mr*) of 80-90 kDa. The precursor enzymes are transformed into the active, mature forms comprising the small (α) and large (β) subunits, with approximate *Mr* values of 20 and 60 kDa, respectively, by the cleavage of a short, intervening polypeptide, called the spacer sequence [17-20]. The enzyme activation occurs via either intermolecular catalysis by other activated acylase molecules [17, 18] or intramolecular catalysis by catalytic residues within the precursor enzyme [19, 20]. Thus, the enzyme activation mechanism is one of the major subjects in acylase biochemistry.

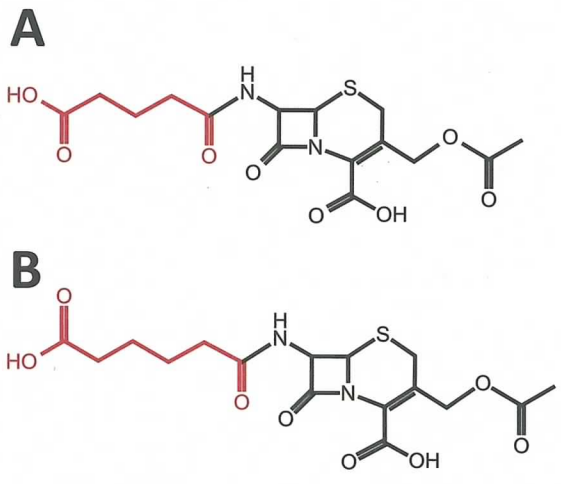


Figure 1. Substrate compounds for cephalosporin acylase (CA). **A**, glutaryl-7-aminocephalosporanic acid (GL-7-ACA). **B**, adipyl-7-aminocephalosporanic acid (AD-7-ACA). The acyl side chains indicated in red are cleaved by the hydrolysis reaction with CA.

Modeling and docking studies of penicillin acylases were performed to elucidate their substrate specificity and stereoselectivity [21, 22]. However, no computational analyses on the structure-based enzymatic properties of CA have been reported. In the present study, docking simulations of the wild-type CA and its variants with substrate compounds were performed, based on the X-ray crystallographic structures. The docking energies obtained by our calculations were compared with the corresponding experimental data. The results are discussed in terms of the mechanisms of substrate selection by antibiotic acylases and of engineering the enzymes for biomedical purposes.

2. Materials and Methods

The X-ray crystallographic structures of native CA from *Pseudomonas* sp. SY-77 (strain GK-16) in the inactive precursor and activated mature forms (Protein Data Bank (PDB) codes 1oqz and 1or0, respectively) [19, 20] were used as the receptor molecules for the docking simulation. The X-ray structure of CA in the activated form, 1or0, was also used for constructing 3D models of the CA mutants analyzed in the present study. The molecular structures of the substrate ligands for the docking simulation, glutaryl-7-aminocephalosporanic acid (GL-7-ACA) and adipyl-7-aminocephalosporanic acid (AD-7-ACA) (Fig. 1), were computationally constructed with the 'Build-Fragment' tool of Discovery Studio (Accelrys, Inc.).

Before the docking simulation, hydrogen atoms were added to the protein and substrate molecules by applying the CHARMm force field [23], and the molecular conformations of these ligands were minimized and refined with partial Momany-Rone charges [24]. The simulation was performed with the ligand-docking program CDOCKER [25], which is based on a molecular dynamics (MD) docking algorithm with the CHARMm force field. To define the substrate binding site in CA from *Pseudomonas* sp. SY-77, the 3D structures were superimposed on that of another CA from *Brevundimonas diminuta*, PDB code 1jvz, in the complex with GL-7-ACA [16], based on the sequence alignment [26] between the CA molecules, to minimize the RMSD between the corresponding backbone atoms. In the structure of CA from the *Pseudomonas* sp. superimposed on that of CA from *B. diminuta*, a virtual sphere, with a radius of 15 Å from the center of the GL-7-ACA molecule associated with CA from *B. diminuta*, was defined as the substrate binding site of CA from the *Pseudomonas* sp. (Fig. 2A). The CDOCKER program was executed under the computational conditions of 100 random initial conformations.

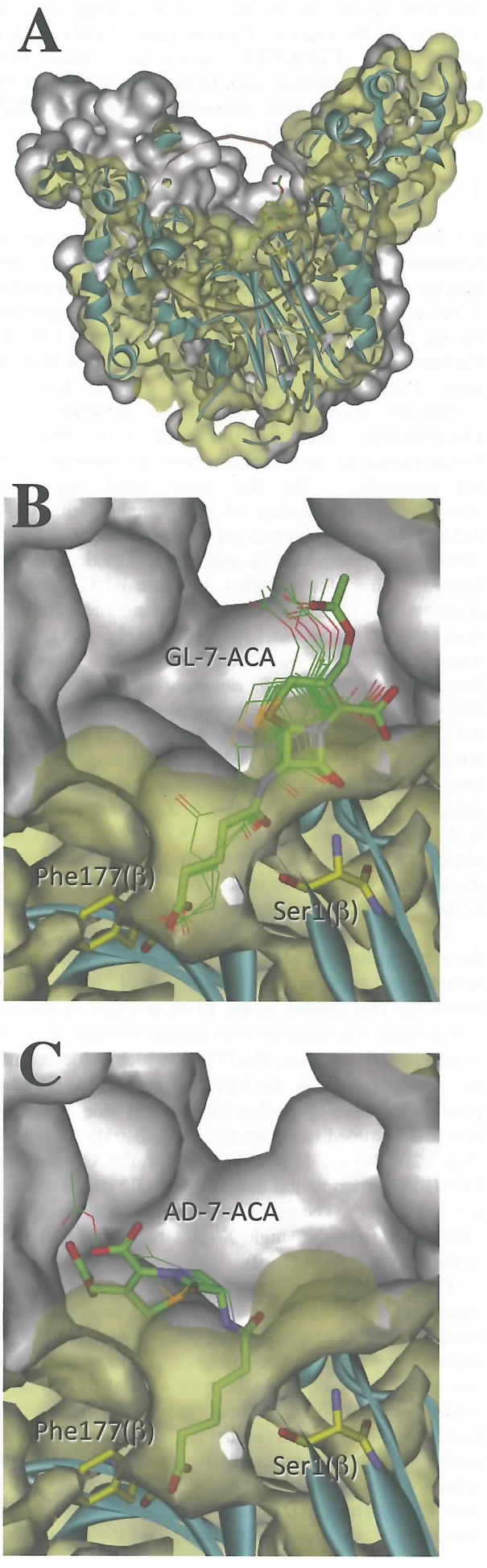
Generally, many docking poses are obtained from the simulations of a protein receptor with a substrate ligand. Each pose yields a 3D structure of the protein in complex with the ligand, and was evaluated by the CDOCKER energy [25], comprising the energy terms for the ligand conformation and for the protein-ligand interaction. In the present study, the docking poses were examined from the viewpoint of the catalytic reaction. We defined the poses as "reactive" when the acyl-side chain was captured by the substrate-binding pocket in the hydrophobic interior of the acylases [16] (see Fig. 2). Unless otherwise noted, all of the docking energy data are reported for the reactive docking poses. Only docking poses with a -CDOCKER energy larger than -1,000 kcal/mol were recorded, and, in the text, NP means that no docking poses were obtained with a -CDOCKER energy larger than -1,000 kcal/mol.

3. Results

CA from *Pseudomonas* sp. SY-77 and its two normal substrates, GL-7-ACA and AD-7-ACA (Fig. 1), were selected for the docking simulation, since high-resolution X-ray crystallographic structures have been determined for the inactive precursor and activated forms [19, 20]. Furthermore, the enzyme kinetics of native CA from the same *Pseudomonas* sp. and its mutants has been extensively investigated with those substrates [7-9]. Unfortunately, the structures of CA from the *Pseudomonas* sp. in complexes with its substrates were not available. On the other hand, the X-ray crystallographic structure of the complex with GL-7-ACA (PDB code 1jvz) was solved for another CA from *B. diminuta* [16], although its mutant and kinetic data were limited. Thus, the complex structure of CA from *B. diminuta* was used to define the binding site within CA from the *Pseudomonas* sp. (1or0) (Fig. 2A). The sequence identity between the CAs from the *Pseudomonas* sp. and *B. diminuta* is 98.1%, and the RMSD between their superimposed main-chain atoms in the 3D structures is 0.36 Å for the 670 overlapping residues. The detailed procedure for defining the binding site in CA from the *Pseudomonas* sp., based on the complex structure of CA from *B. diminuta*, is described in the Materials and Methods. No docking poses were obtained for CA in the precursor form (1oqz) with any of the ligands, because the spacer chain blocked the access of the ligands to the catalytic site.

To compare the computational docking energies with the experimental K_m values, we used the enzyme kinetics data for a series of CA mutants, Phe177(β)X [7], in which the Phe residue at the 177th position in the wild type β chain was replaced with another residue, X. In the wild-type CA structure, Phe177(β), which is located near the catalytic residue Ser1(β), is one of the residues forming the binding pocket for the acyl side chain, and thus it is assumed to be involved in substrate recognition (Figs. 2B and C) [16, 17]. As for the CA mutants, F177(β)X, experimental data for their 3D structures were not available. Therefore, the mutant structures were constructed from the single X-ray structure of native CA in the activated form (1or0).

For each mutant, three structural models were constructed and used for the docking simulation. In model 1, the conformation of the new residue at the replaced site remained unchanged; i.e., its rotamer state was either the same as or the most similar to that of the wild-type Phe177(β), without any refinement in the mutant molecule. In model 2, only the side-chain conformation at the replaced site was minimized for the refinement by molecular mechanics (MM) calculations with CHARMm [23], before docking. Except for the replaced site, the remaining structure of the mutants was



the same as that of the original, native enzyme. In model 3, the conformations of three residues including the new residue at the replaced site, X177(β), as well as the two residues Thr176(β) and Asn68(β) that are closest to X177(β), were refined together, before docking. Furthermore, we performed the docking simulation based on the "induced fit" method [27]. In this simulation, the side-chain conformation of X177(β), as well as those of four other residues, Tyr149(α), Tyr33(β), Arg57(β), and Phe58(β), which form the substrate binding site together with X177(β), were minimized during the docking simulation, starting from model 1 of each mutant, by the MM calculation.

The docking simulation was performed for all combinations of either the wild-type or a mutant CA with GL-7-ACA or AD-7-ACA. Many docking poses were obtained from the simulation for a protein receptor with a substrate ligand. Each pose yielded a 3D structure of the protein in complex with the ligand (Fig. 2), and was evaluated by the CDOCKER energy [25], comprising the energy terms for the ligand conformation and the interaction between protein and ligand. We collected the ten best poses and report their energy values for both the minimum (E_{min}) and the average (E_{avr}) ± standard deviation (SD).

The computationally obtained CDOCKER energies, expressed as the negative values, $-E_{min}$ and $-E_{avr}$, for the protein models were plotted against the experimental data of -ln K_m [7] (for model 2, see Fig. 4). The correlation coefficients and their *P*-values were estimated from these plots. The calculated energies significantly depend on the protein model, suggesting their high sensitivity to the local protein environment at the binding site. The correlation data in which statistically significant relationships were detected; i.e., in the cases of $P<0.05$,

---

Figure 2. Docking models of native CA from *Pseudomonas* sp. SY-77 (PDB code 1or0) with GL-7-ACA and AD-7-ACA. (**A**) A docking model with GL-7-ACA, indicating the docking sphere defined with the structure of CA from *Brevundimonas diminuta* in the complex with GL-7-ACA (1jvz) (see Materials and Methods). It is represented in sections to allow visualization of the inside of the protein molecule. The solvent accessible surface of the protein is depicted in solid white, and the back side of the surface is transparent light yellow. (**B**) Enlarged view of the substrate binding site of CA, showing the ten best docking poses with GL-7-ACA. The best ligand pose is depicted by solid sticks, and the other nine ligands are depicted by thin lines, colored by atom type (C, green; N, blue; O, red). Phe177(β) at the mutated site and the catalytic nucleophile residue, Ser1(β), are also displayed with solid sticks colored by an atom type (C, yellow; N, blue; O, red). (**C**) Enlarged view of the substrate binding site of CA, showing the ten best docking poses with AD-7-ACA. The rendering and coloring of the modes are similar to those in (B). The conformational variations of the AD-7-ACA poses is small, in comparison with that of GL-7ACA (B), and they can almost be superimposed on the two representative conformations.

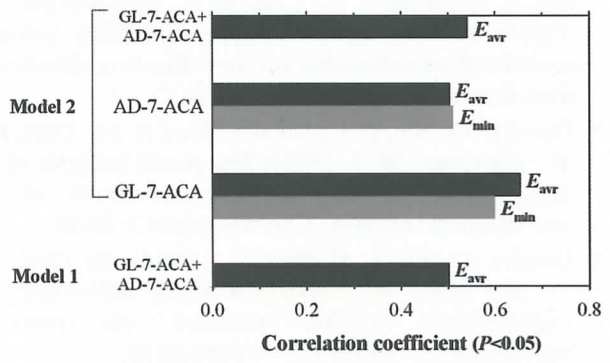


Figure 3. Correlation coefficients between the computational and experimental data. All the combinations of the four computational methods with the two substrates were tested, and only the data with P values less than 0.05 are plotted in this graph.

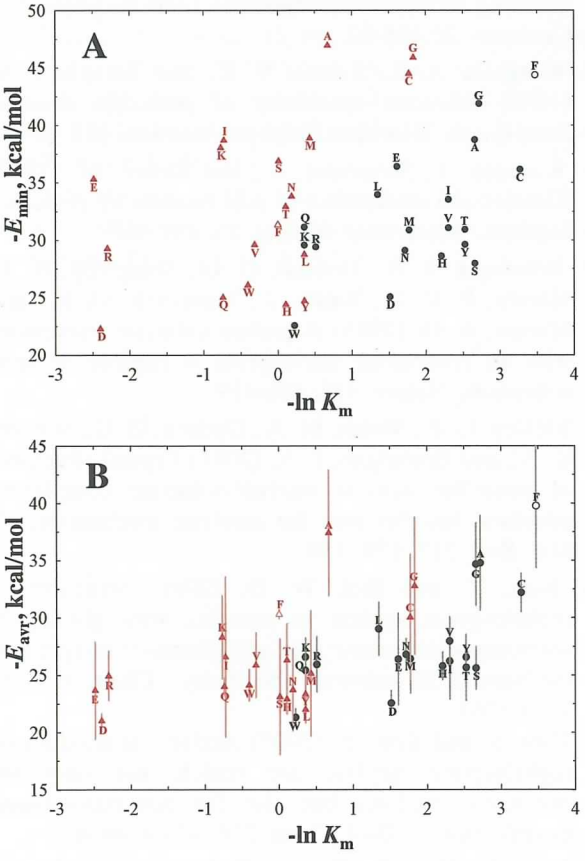


Figure 4. Plots of calculated docking energy as a function of experimental K_m (mM) values for F177(β)X mutants of CA with substrate ligands, GL-7-ACA and AD-7-ACA. Negative values of the minimum and average CDOCKER energies, $-E_{min}$ and $-E_{avr}$ ± SD, are plotted against -ln K_m for protein model 2 in **A** and **B**, respectively. The capital letters accompanying each data point in these plots indicate the amino acid residues at position 177 in the β-chain of the CA mutants. Open circles and open triangles show the data of native CA; i.e., Phe at position 177(β), for GL-7-ACA and AD-7-ACA, respectively. The closed circles and closed triangles show the data of mutant CAs, F177(β)X, for GL-7-ACA and AD-7-ACA, respectively.

were presented in Fig. 3. Among these, model 2, in which only the substituted residue was refined, gives the best correlation. The plots of the energies against -ln K_m for model 2 and their correlation data are displayed in Fig. 4 and Table 1, respectively. The plots for the substrate GL-7-ACA showed better correlation than those for AD-7-ACA. In addition, E_{avr} exhibits better correlation than E_{min}. The best correlation was obtained for the plot of E_{avr} in model 2 with GL-7-ACA, the best substrate of CA with K_m values less than 1 mM.

4. Discussion

In the present work, the substrate binding by CA from *Pseudomonas* sp. SY-77 and its mutants was extensively studied by MM and MD calculations. The 3D structures of these mutants, in which a single residue close to the catalytic center was replaced with other amino acid residues, were constructed based on an X-ray crystallographic structure of the native enzyme, in different manners. Among these structure sets for the CA mutants, one set exhibited the best correlation with the experimentally obtained kinetic data [7] (Table 1, Figs. 3 and 4). In these mutant structures exhibiting the best correlation of the calculated docking energies with the experimental data, only the replaced residues were subjected to the side-chain refinement, and the remaining structures other than the replaced site were the same as that of the native enzyme. This suggests that the residue replacements at this site actually exert few effects on the surrounding structure, and that the side-chain refinement for several residues including the replaced residue cannot predict their correct conformations in the present calculation.

The range of experimental K_m values, 10^{+3}~10^{-4}, on the x-axis of the plots in Fig. 4 is expected to be the range of energies estimated from RT ln K, approx. 10 kcal/mol, at 37°C. This energy value is smaller than the range, 30 kcal/mol, for the calculated docking energies on the y-axis, but is consistent with the order of magnitude. For the correlation of the docking energy with K_m, the rate of acyl-enzyme intermediate formation [13] is assumed to be faster than its hydrolysis. The docking energy for

Table 1. Correlation coefficients between experimentally obtained -ln K_m and -CDOCKER energy obtained from docking simulation of CA with substrate ligands, GL-7-ACA and AD-7-ACA.

Energy[1]	GL-7-ACA	AD-7-ACA	GL-7-ACA+ AD-7-ACA
E_{min}	0.599 (0.007)	0.510 (0.026)	0.3170 (0.052)
E_{avr}	0.653 (0.002)	0.504 (0.028)	0.542 (0.0004)

[1]The data were obtained from the plots of the calculated energy values against -ln K_m in Fig. 4. The numbers in parentheses below the correlation coefficient values are the P-values for each correlation. Lower P-values indicate higher statistical significance of the correlation.

GL-7-ACA correlated well with the experimental K_m values, indicating the faster formation of the acyl enzyme in this case. However, the correlation for AD-7-ACA is not so significant (Table 1, Figs. 3 and 4), and AD-7-ACA is a worse substrate with higher K_m values. Thus, the lower correlation for AD-7-ACA may be due to the slower formation of the acyl-enzyme.

Docking simulations were also performed with PGA and homology models of aculeacin A, a cyclic lipopeptide antibiotics, acylase [28-32]. These data are compared with those of CA, and are discussed for elucidation of mechanisms of the substrate selection by the antibiotic acylases, which will be reported elsewhere. Based on the present results, mutational studies of the antibiotic acylases with experiments are now in progress.

In conclusion, the docking simulations of the native and mutant CA were performed with the substrate ligands, and one of the computational methods was validated for the enzyme. The present study demonstrates usefulness of the computational approach for analyzing and engineering the substrate specificities of antibiotic acylases.

Acknowledgements

The authors thank Mr. Kosuke Asakura (Toyama Prefectural University) and Dr. Tadahiro Ohmura, (Accelrys K. K.) for assistance and support of the computations. This work was supported by Grant-in-Aids for Scientific Research (23510280) from Japan Society for the Promotion of Science.

References

1. Fleming, A. (1929) On the antibacterial action of cultures of a Penicillium, with special reference to their use in the isolation of *B. influenzae*. British Journal of Experimental Pathology 10: 226-236.
2. Clark, H. T., Johnson, J. R., and Robinson, R. (1949) The chemistry of penicillin, Princeton Univ. Press, Princeton.
3. Newton, G. G. and Abraham, E. P. (1955) Cephalosporin C, a new antibiotic containing sulphur and D-alpha-aminoadipic acid. Nature 175: 548.
4. Huber, F., Chauvette, R., and Jackson, B. (1972) Preparative methods for 7-aminocephalosporanic acid and 6-aminopenicillanic acid, p. 27–48. *In* Flynn, E. (ed.), Cephalosporins and Penicillins, Chemistry and Biology, Academic Press, New York.
5. Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. (1995) SCOP: a structural classification of protein database for the investigation of sequences and structures. J. Mol. Biol., 247: 536–540.
6. Barends, T. R. M., Yoshida, H., and Dijkstra, B. W. (2004) Three-dimensional structures of enzymes useful for β-lactam antibiotic production. Curr. Opin. Biotechnol., 15: 356-363.
7. Sio, C. F., Otten, L. G., Cool, R. H., and Quax, W. J. (2003) Analysis of a substrate specificity switch residue of cephalosporin acylase. Biochem. Biophys. Res. Commun., 312: 755-760.
8. Otten, L. G., Sio, C. F., van der Sloot, A. M., Cool, R. H., and Quax, W. J. (2004) Mutational analysis of a key residue in the substrate specificity of a cephalosporin acylase. Chembiochem, 5: 20-25.
9. Otten, L. G., Sio, C. F., Reis, C. R., Koch, G., Cool, R. H., and Quax, W. J. (2007) A highly active adipyl-cephalosporin acylase obtained via rational randomization. FEBS J. 274: 5600-5610.
10. Li, Y., Chen, J., Jiang, W., Mao, X., Zhao, G., and Wang, E. (1999) *In vivo* post-translational processing and subunit reconstitution of cephalosporin acylase from *Pseudomonas* sp. 130. Eur. J. Biochem. 262: 713-719.
11. Oh, B., Kim, K., Park, J., Yoon, J., Han, D., and Kim Y. (2004) Modifying the substrate specificity of penicillin G acylase to cephalosporin acylase by mutating active-site residues. Biochem. Biophys. Res. Commun. 25:486-92.
12. Morgolin, A. L., Svedas, V. K., and Berezin, I. V. (1980) Substrate specificity of penicillin amidase from E. coli. Biochim. Biophys. Acta 616: 283-289.
13. Konecny, J., Schneider, A., and Sieber, M. (1983) Kinetics and mechanism of acyl transfer by penicillin acylases. Biotechnol. Bioeng. 25: 451-467.
14. Brannigan, J. A., Dodson, G. G., Duggleby, H. J., Moody, P. C. E., Smith, J., Tomchick, D. R., and Murzin, A. G. (1995) A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. Nature, 378: 416-419.
15. McVey, C. E., Walsh, M. A., Dodson, G. G., Wilson, K. S., and Brannigan, J. A. (2001) Crystal structures of penicillin acylase enzyme-substrate complexes: structural insights into the catalytic mechanism. J. Mol. Biol. 313: 139–150.
16. Kim, Y. and Hol, W. G. (2001) Structure of cephalosporin acylase in complex with glutaryl-7-aminocephalosporanic acid and glutarate: insight into the basis of its substrate specificity. Chem. Biol. 8: 1253-1264.
17. Kim, S. and Kim, Y. (2001) Active site residues of cephalosporin acylase are critical not only for enzymatic catalysis but also for post-translational modification. J. Biol. Chem. 276: 48376-48381.
18. Kim, Y., Kim, S., Earnest, T. N., and Hol, W. G. (2002) Precursor structure of cephalosporin acylase. Insights into autoproteolytic activation in a new N-terminal hydrolase family. J. Biol. Chem. 277: 2823-2829.
19. Kim, J. K., Yang, I. S., Rhee, S., Dauter, Z., Lee, Y. S., Park, S. S., and Kim, K. H. (2003) Crystal structures of glutaryl 7-aminocephalosporanic acid acylase: insight into autoproteolytic activation. Biochemistry 42: 4084-4093.
20. Kim, J. K., Yang, I. S., Shin, H. J., Cho, K. J., Ryu, E. K., Kim, S. H., Park, S. S., and Kim, K. H. (2006) Insight into autoproteolytic activation from the structure of cephalosporin acylase: a protein with two

proteolytic chemistries. Proc. Natl. Acad. Sci. USA. 103: 1732-1737.

21. Braiuca, P., Ebert, C., Fischer, L., Gardossi, L., and Linda, P. (2003) Homology model of Penicillin acylase from *Alcaligenes faecalis* and in silico evaluation of its selectivity. Chembiochem 4: 615-622.
22. Guncheva, M., Ivanov, I., Galunsky, B., Stambolieva, N., and Kaneti, J. (2004) Kinetic studies and molecular modelling attribute a crucial role in the specificity and stereoselectivity of penicillin acylase to the pair ArgA145-ArgB263. Eur. J. Biochem. 271: 2272–2279.
23. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMm: A program for macromolecula r energy, minimization, and dynamics calculations. J. Comp. Chem. 4: 187-217.
24. Momany, F. A. and Rone, R. (1992) Validation of the general purpose QUANTA 3.2/CHARMm force field. J. Comp. Chem. 13: 888-900.
25. Wu, G., Robertson, D. H., Brooks, C. L. 3rd, and Vieth, M. (2003) Detailed analysis of grid-based molecular docking: A case study of CDOCKER—A CHARMm-based MD docking algorithm. J. Comput. Chem. 24: 1549-1562.
26. Marti-Renom, M. A., Madhusudhan, M. S., and Sali, A. (2004) Alignment of protein sequences by their profiles. Protein Sci. 13: 1071-1087.
27. Koska, J., Spassov, V. Z., Maynard, A. J., Yan, L., Austin, N., Flook, P. K., and Venkatachalam, C. M. (2008) Fully automated molecular mechanics based induced fit protein-ligand docking method. J. Chem. Inf. Model. 48: 1965-1973.
28. Denning, D. W. (2003) Echinocandin antifungal drugs. Lancet 362: 1142-1151.
29. Hashimoto, S. (2009) Micafungin: a sulfated echinocandin. J. Antibiot. 62: 27-35.
30. Takeshima, H., Inokoshi, J., Takada, Y., Tanaka, H., and Omura, S. (1989) A deacylation enzyme for aculeacin A, a neutral lipopeptide antibiotic, from *Actinoplanes utahensis*: purification and characterization. J. Biochem. 105: 606-610.
31. Inokoshi, J., Takeshima, H., Ikeda, H., and Omura, S. (1992) Cloning and sequencing of the aculeacin A acylase-encoding gene from *Actinoplanes utahensis* and expression in *Streptomyces lividans*. Gene 119: 29-35.
32. Torres-Bacete, J., Hormigo, D., Stuart, M., Arroyo, M., Torres, P., Castillon, M. P., Acebal, C., Garcia, J. L., and de la Mata, I. (2007) Newly discovered penicillin acylase activity of aculeacin A acylase from *Actinoplanes utahensis*. Appl. Environ. Microbiol. 73: 5378-5381.