Published in final edited form as:

Cryst Growth Des. 2008 November 5; 8(11): 4054–4061. doi:10.1021/cg800430f.

Function-biased choice of additives for optimization of protein crystallization – the case of the putative thioesterase PA5185 from *Pseudomonas aeruginosa* PAO1

Maksymilian Chruszcz^{1,3}, Matthew D. Zimmerman^{1,3}, Shuren Wang^{1,3}, Katarzyna D. Koclega^{1,3}, Heping Zheng^{1,3}, Elena Evdokimova^{2,3}, Marina Kudritska^{2,3}, Marcin Cymborowski^{1,3}, Alexei Savchenko^{2,3}, Aled Edwards^{2,3}, and Wladek Minor^{1,3}

¹Department of Molecular Biology and Biological Physics, University of Virginia, Charlottesville, Virginia 22908, USA

²Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario M5G 1L6, Canada

³Midwest Center for Structural Genomics

Abstract

The crystal structure of PA5185, a putative thioesterase from *Pseudomonas aeruginosa* strain PAO1, was solved using multi-wavelength anomalous diffraction to 2.4 Å. Analysis of the structure and information about the putative function of the protein were used to optimize crystallization conditions. The crystal growth was optimized by applying additives with chemical similarity to a fragment of a putative PA5185 substrate (CoA or its derivative). Using new crystallization conditions containing this function-biased set of additives, several new crystal forms were produced and structures of three of them (in three different space groups) were determined. One of the new crystal forms had an improved resolution limit of 1.9 Å, and another displayed an alternative conformation of the highly-conserved loop containing Asn26, which could play a physiological role. Surprisingly, none of the additives were ordered in the crystal structures. Application of function-biased additives could be used as a standard optimization protocol for producing improved diffraction, or new crystal forms, which may lead to better understanding of the biological functions of proteins.

Keywords

Protein crystallization; additives; polymorphism; thioesterase; oligomers

Introduction

The production of protein crystals suitable for structural analysis represents one of the major bottlenecks in the structure determination process, and thus, well-diffracting single crystals of macromolecules, either by weight or by volume, are one of the most valuable pieces of matter on Earth. After obtaining initial crystallization conditions, the parameters of crystal growth are optimized in order to get well-diffracting crystals that will be used for X-ray structure determination. The process of crystal growth optimization could be performed in many ways. Traditionally (and most commonly) grid screen and (more rarely) response-surface methods

Corresponding author: Wladek Minor, University of Virginia, Department of Molecular Physiology and Biological Physics, 1340 Jefferson Park Avenue, Charlottesville, VA 22908, USA. E-mail address: wladek@iwonka.med.virginia.edu Phone: +1-434-243-0033 Fax: +1-434-982-1616.

are used,¹ where chemical parameters of the precipitating solution are systematically varied around the initially obtained conditions.

An alternative approach, used successfully in many cases, involves the usage of so-called 'additives', usually small molecular compounds which modify the crystallization condition and sometimes lead to significant improvement of crystal quality. Such an approach does not require chemical modification of the protein like reductive methylation of lysines^{2, 3} or mutations to reduce protein's surface entropy.^{4, 5} Additives can be very simple from the chemical point of view: even addition of different inorganic ions alone have been shown to strongly influence the quality of crystals.⁶⁻¹⁰ More frequently organic compounds are used as additives.^{8, 11-14} In particular, detergents have found especially broad application.^{11, 15-17}

Additives beneficially affect macromolecule crystallization in one of three ways. In the first case, the additives alter the physiochemical properties of the crystallization experiment but are not ordered in the resulting crystal structure. In such situations the crystal structure provides no simple explanation how these compounds affect the process of crystal formation. In the second case, the additive interacts with the protein and is ordered in the resulting crystal structure, but the interaction is not biologically relevant. For example, the additive may directly or indirectly mediate crystal contacts, which may be observed explicitly in the structure. In the third (and most interesting) case, the additive interacts with the protein in a way that is physiologically relevant, either because the additive is itself a natural ligand or it mimics some aspect of natural ligands. In these situations, the additive may not directly mediate crystal contacts but may indirectly promote crystallization by stabilizing the macromolecular conformation. Most importantly, because the ordered density for the additive is observed in a biologically relevant position, it may provide crucial information about the functional mechanisms of the protein.

We present an example of successful approach for optimization of initial crystallization conditions by application of compounds that are similar to protein ligand(s) or contain chemical groups that could mimic parts of the ligand(s). As the object of this study we chose PA5185 – a putative thioesterase from *Pseudomonas aeruginosa* strain PAO1.¹⁸ *P. aeruginosa* is a Gram-negative bacterium and a major opportunistic human pathogen.¹⁹ The bacterium causes infections mainly in hospitalized, immuno-compromised, and cystic fibrosis patients,²⁰ and it demonstrates increasing drug resistance.²¹

Material and Methods

Protein Expression and Purification

Selenomethionine (Se-Met) substituted PA5185 from *Pseudomonas aeruginosa* was cloned and purified using the standard protocol developed at Midwest Center for Structural Genomics (MCSG), as described previously.²² The native protein was expressed using a modified pET-15b plasmid cloned into B834(DE3)pLysS *E. coli* cells. The seed cultures (25 mL) were grown overnight at 37°C in TB media supplied with 25 µL of 100 mg/mL ampicillin and 50 µL of 15 mg/mL chloramphenicol. The cultures were transferred to 1 L flasks and their growth was monitored by checking the optical density at 600 nm. When the optical density reached 1.1 the cultures were induced with isopropyl-1-β-D-thiogalactopyranoside (final concentration of 1 mM) and expressed overnight at a temperature of 16°C. The same purification protocol used to produce Se-Met-substituted protein was used to produce native protein. For both native and Se-Met-substituted protein, after His-tag cleavage and a second subtractive step of Ni-NTA (QIAGEN) affinity chromatography the sample was dialyzed into a buffer containing 500 mM NaCl and 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH=7.5. The protein after dialysis was further purified with a gel filtration column (HiLoad 6/16 Superdex 200) on an AKTA FPLC system (GE Healthcare). The Se-Met-substituted

protein was concentrated to 10.5 mg/mL, and the native protein to 13.6 mg/mL. The protein samples were flash-frozen in liquid nitrogen and stored at -80°C .

Crystallization and Data Collection

Crystals of Se-Met-substituted protein were obtained by hanging-drop vapor diffusion at 293 K, in drops containing a 1:1 mixture of protein solution (10.5 mg/mL) and well precipitant solution (25% w/v PEG3350, 0.05 M ammonium sulfate, and 0.1 M 1,3-bis(tris(hydroxymethyl)methylamino)propane (Bis-Tris) pH=5.5). Crystals of the native protein were grown using the same method and temperature as used for the Se-Met protein, except that the ammonium sulfate in the precipitant solution were replaced with various different additives. The concentration of the native protein used for crystallization was 13.6 mg/mL. Tracking of crystals and drops, and analysis of intermediate results were performed using the crystallization expert system *Xtaldb*.²³

Structure Solution and Refinement

Prior to data collection crystals were transferred to a cryoprotectant solution and cooled by plunging into liquid nitrogen. Data collection was done at the Structural Biology Center²⁴ at sector 19 of the Advanced Photon Source (APS). Details of data collection, structure determination, and refinement statistics are summarized in Table 1. Data from Se-Met-substituted and native crystals were processed with *HKL-2000*.²⁵ The structure of the Se-Met substituted protein was solved by the multi-wavelength anomalous diffraction (MAD) method.²⁶ A Se-Met substructure was found using Se absorbance peak data with *SHELXD*.²⁷ as implemented in an early version of *HKL-3000*.²⁸ An experimental electron density map was obtained using *autoSHARP*.²⁹ The initial model, built by *RESOLVE*,³⁰⁻³² was extended by *ARP/wARP*³³ and manual building with *COOT*.³⁴ The crystal structures of the native protein were solved using the molecular replacement method as implemented by *MOLREP*,³⁵ using the refined structure of the Se-Met protein (PDB code: 2AV9) as a search model. In all cases refinement was performed using *REFMAC*³⁶ as implemented in the *CCP4* package.³⁷ In the last stages of the refinement TLS was applied.³⁸ The TLS groups were defined using the *TLMSD* server.^{39, 40} In all cases more than one protein chain was present in each asymmetric unit and non-crystallographic symmetry (NCS) provided additional restraints. Structure validation was performed with *MOLPROBITY* and *PROCHECK*.^{41, 42} The atomic coordinates for all structures, together with the structure factors, were deposited in the PDB.⁴³ All figures presenting details of the determined structures were prepared using *PYMOL*.⁴⁴

Analysis of the PDB

Entries from the November 2007 PDB release were analyzed to check how often polymorphic crystal forms (i.e., when one protein crystallized in two or more space groups) of proteins are deposited. In our analysis we extracted sequence information directly from coordinate files. Our analysis considered only crystal structures that contain one or more copies of the same polypeptides chain in the asymmetric unit. And only structures with the sequence longer than 20 amino acids were taken into account for polymorphic crystal form analysis. Specifically, all structures that contain hetero-oligomers or non-standard residues were excluded from our dataset. (An exception was that Se-Met residues were treated as methionines, due to a high degree of similarity).⁴⁵ Protein-DNA and protein-RNA complexes were omitted. The analyzed set of structures contained 31,813 different PDB deposits.

Results and Discussion

Function and overall structure of PA5185

PA5185 belongs to the thioesterase superfamily (Pfam accession number: PF03061; <http://pfam.sanger.ac.uk/>). This family contains many enzymes, and most of them are thioesterases.⁴⁶ Despite low sequence identity (Table 2), the structure PA5185 shows strong structural similarity with the 4-hydroxybenzoyl-CoA thioesterase⁴⁷ from *Pseudomonas* sp. strain CBS-3 (PDB codes: 1LO9 and 1BVQ) and other proteins that have a similar alpha/beta “hot dog” fold⁴⁸ where five anti-parallel beta strands are packed against a large helix. The proteins that are structurally most similar to PA5185 were found by a *PROFUNC*⁵⁰ search are summarized in Table 2. PA5185 contains Asp26 (Figure 1), which is structurally equivalent to the functionally critical conserved residue Asp17 of 4-hydroxybenzoyl-CoA thioesterase.^{47, 51, 52} Given that this aspartate is conserved and the structures overall are so similar, we believe that protein PA5185 could be involved in the 4-hydroxybenzoyl-CoA thioesterase reaction and thus that the 4-chlorobenzoyl-CoA dehalogenation pathway could exist in bacteria *Pseudomonas aeruginosa* strain PAO1. Moreover PA5185 is located on the same operon with genes annotated to code for a probable iron-containing alcohol dehydrogenase, a probable acetyl-CoA dehydrogenase and a probable 3-hydroxyacyl-CoA dehydrogenase (PA5186, PA5187 and PA5188 respectively). On this basis, as well on the basis of structural data, we hypothesize that PA5185 is an acetyl-CoA thioesterase and most probably, like other enzymes from this family, is involved in lipid metabolism.⁵³

Both experimental data and theoretical predictions establish that PA5185 forms a tetramer. The oligomeric state of the protein in solution was determined by a series of gel filtration chromatography experiments (data not shown). In the solutions used (0.1 M buffer concentration, 0.15 M NaCl) stable tetramers were observed in different buffers with a pH range of 5–8. As the pH values of most of the crystallization solutions reported in this work lay within this range, it is likely tetramers were the only oligomeric assemblies present in solution during crystallization. According to predictions of the oligomeric state done with *PITA* (<http://www.ebi.ac.uk/thornton-srv/databases/pita/>),⁴⁹ PA5185 forms tetramers (Figure 2) in all crystal forms described in this work.

The proposed functional tetramer of PA5185 is stabilized mainly by intermolecular contacts through anti-parallel beta-strand hydrogen bonds and through hydrogen bonds and salt bridges between Ser19 and Ser19', and between Arg21 and Asp24' from different subunits. The tetramer has 222 point group symmetry and may be treated as a dimer of dimers (Figure 2C). Complex formation is connected with a significant loss of solvent-accessible surface area (ASA) and almost 27% of the area of a single polypeptide chain is involved in the interaction with another unit forming the tetramer. The same type of tetrameric form is also observed in crystal structures of proteins that are structurally similar to PA5185 (Table 2).

Crystal forms of PA5185

The first structure of PA5185 (PDB code: 2AV9) was determined using a Se-Met derivative of the protein. After model building and refinement it was noted that sulfate ions bind to the protein in different positions (Figure 3). Most of them (11 of 15) are bound by the loop formed by the residues Gly92-Ser95, in a position equivalent to the binding site of the phosphate in acetyl-CoA in the structure of 4-hydroxybenzoyl-CoA thioesterase (Figure 3A). Four other sulfates also bind on the protein surface and mediate interactions between chains that form tetramers. For example, sulfate ion 113 (Figure 3C) is located in a region in which three different protein chains (J, G, F) meet. Sulfate groups 101 and 115 are located on a non-crystallographic twofold axis and each of them bridge two different protein chains (B-E and J-F) through interactions with residue Arg81 from each chain (Figure 3B).

None of the sulfate groups are directly involved in crystal contacts (Table 3); although some of them (101, 111 and 115) are located in regions responsible for crystal contact formation. Sulfate groups 101 and 115 most probably promote crystal contact formation between Arg81 residues and Glu50 or Glu105 residues from neighboring tetramers (Table 3). Taking into account the multiple localizations of sulfate groups and their chemical similarity to a fragment of the putative substrate of PA5185 (CoA or its derivative) (Figure 4), we decided to modify the original crystallization conditions, by removing sulfate ions and replacing them with chemical compounds that contain sulfate or phosphate moieties.

During crystallization the following additives were used: adenosine-5'-triphosphate (ATP), 4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid (EPPS), 2-(N-morpholino)ethanesulfonic acid (MES; pH=5.9), 3-(N-morpholino)propanesulfonic acid (MOPS), HEPES (pH=7.0), sodium pyrophosphate, glucose-6-phosphate, non-detergent sulfobetaine 201 (NDSB-201), NDSB-256, coenzyme A sodium salt (sodium CoA), lithium myristoyl-CoA, lithium benzoyl-CoA and lithium acetyl-CoA. Such an approach yielded several new crystal forms, from which three new structures were determined (Table 1). The reported structures differ significantly in solvent content, number of molecules in the asymmetric unit, and diffraction limits. Although the concentrations of the additives used were relatively high compared to protein concentration, none of the determined structures contained localized additive molecules or fragments of them. Despite the fact that the additives were not localized in the electron density maps, they nonetheless were able influence the conformation of the protein chain, which is particularly visible in the region of Gly52-Gly59 (Figures 5A, 5B).

PA5185 crystallized in many forms (Figures 6-8). Removal of ammonium sulfate from the crystallization conditions, its replacement with additives, or both leads to conditions in which at least two different crystal forms are present in the same drop. Moreover, the relative abundances of the crystal forms change with respect to time (Figure 7). On the basis of our current observations, we suggest that at the early stages of the crystallization experiment, crystal forms with different solvent contents may coexist. As the drop dries, crystal forms with higher solvent content (block shaped, Figure 7B) are consumed by the more slowly growing needle-shaped crystals with lower solvent content. The analysis becomes more complicated when not only the concentrations of the additives are taken into account, but also when the concentration of the precipitate (PEG 3350) is considered (Figures 8C, 8D).

More crystal forms – more biological information?

As noticed previously, two different conformations of the PA5185 chain are observed in the monoclinic crystal. The three tetramers present in the asymmetric unit of the C2 forms differ (Figure 5D). Interestingly, in all other crystal structures described in this work, only tetramers with chains in conformation 'l' (the conformation with extended Asn57 extending to bind Gly29 of a neighboring chain) are observed. Most probably the presence of the sulfate ion induces the change of the conformation from 'l' to 'h' (the conformation with Asn57 forming an α -helical fragment). It is especially worth mentioning that tetramers with all chains in the 'h' conformation are not observed in any of the reported PA5185 crystal structures. Similarly, there are no dimers where both protein chains adopt the 'h' conformation (Figure 5D). It is possible that changing the conformation in region composed of residues Gly52-Gly59 is responsible for communication between dimers and it is important for enzymatic activity of the enzyme. First of all, the change of the loop conformation may be important for substrate binding or product release (Figure 9A), which is for example observed in the binding of effectors to the binding domain of FapR.⁶⁰

The conformation change from 'h' to 'l' causes a 7 Å shift of the Asn57 C α carbon, and almost 11 Å shift of the side chain's amide group (Figure 5B). Such shift of the Asn57 allows for formation of a hydrogen bond with the carbonyl oxygen atom from Gly29 of the neighboring

chain. Comparison of these two conformations (Figure 9B) reveals changes in the conformations of the sidechains of Asp26 and His30. Both Asp26 and Gly29 are conserved in all of the most structurally similar proteins (Figure 1), while His30 is conserved in most of them. Taking all these observations into account, we suggest that changes in the conformation of the Gly52-Gly59 loop may couple catalytic activity and/or substrate binding between dimers forming the tetramer and allow for cooperative action in the oligomer.

Polymorphism of protein crystals

Analysis of a set of structures in the PDB determined by X-ray crystallography, reveals that most protein sequences (90.7%) are associated with only one crystal form. 7.7% of the sequences are associated with two crystal forms, and only 1.3% of them crystallized in three crystal forms. The case reported here, where the PA5185 protein crystallized in four different forms, is quite unusual, yet in the PDB there are over fifty such cases. Five or more crystal forms are very unusual and our PDB analysis revealed only twenty such cases. Bovine pancreatic ribonuclease is reported in 10 distinct crystal forms, followed by concanavaline A from *Canavalia ensiformis* with 9, and human FK-506 binding protein (FKBP12) with 8 forms. There have been reports in the literature of larger sets of non-isomorphous crystal forms of some proteins, such as 9 different forms of cutinase,⁶¹ or 25 different forms of T4 lysozyme.⁶² However, these analyses include crystal forms where the protein in question contains sequence mutations, while our analysis is explicitly limited to crystal forms with 100% sequence identity to one another. It is possible that the analysis presented here does not fully show the propensity of proteins to form polymorphic crystal forms, as more crystallization trials with different crystallization conditions may produce new crystal forms. However, usually only the best ‘behaving’ crystal species that produced diffraction of decent quality are reported in the PDB.

Conclusions

Optimization of protein crystals to produce well-diffracting species often takes more time and effort than finding the initial crystallization conditions. As shown above, sometimes very small changes of the crystallization conditions may significantly improve the diffraction characteristics. A similar effect may also be obtained in the case where a low resolution structure is known, and so-called ‘reverse screening’ is applied.⁶³ Improvement of the diffraction resolution is not the only reason why further crystallization optimization is worth spending the time. With every new crystal form there is also the possibility of learning something new about the function or dynamics of the protein, particularly when the protein exists in different conformations and/or oligomeric states. The approach presented in this work, where chemicals that mimic fragments of a putative protein ligand are used, is in some sense similar to fragment-based drug discovery.⁶⁴ In our case, crystals in the same crystal form with improved diffraction characteristics, or new crystal forms that display reduced twinning, protein-additive complexes, or alternative polypeptide conformations are all considered to be successes.

Acknowledgments

We thank Andrzej Joachimiak and the members of the Structural Biology Center at the Advanced Photon Source and the Midwest Center for Structural Genomics for help and discussions. The results shown in this report are derived from work performed at Argonne National Laboratory, at the Structural Biology Center of the Advanced Photon Source. Argonne is operated by University of Chicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. The work described in the paper was supported by NIH PSI grants GM62414 and GM074942

References

1. Carter CW Jr, Yin Y. *Acta Crystallogr. Sec. D* 1994;50:572–590.
2. Rayment I. *Methods Enzymol* 1997;276:171–179. [PubMed: 9048376]
3. Walter TS, Meier C, Assenberg R, Au KF, Ren J, Verma A, Nettleship JE, Owens RJ, Stuart DI, Grimes JM. *Structure* 2006;14:1617–1622. [PubMed: 17098187]
4. Derewenda ZS, Vekilov PG. *Acta Crystallogr. Sec. D* 2006;62:116–124.
5. Cooper DR, Boczek T, Grelewska K, Pinkowska M, Sikorska M, Zawadzki M, Derewenda Z. *Acta Crystallogr. Sec. D* 2007;63:636–645.
6. Kalb AJ, Yariv J, Helliwell JR, Papiz MZ. *J. Cryst. Growth* 1988;88:537–540.
7. Vaney MC, Broutin I, Retailleau P, Douangamath A, Lafont S, Hamiaux C, Prange T, Ducruix A, Ries-Kautt M. *Acta Crystallogr. Sec. D* 2001;57:929–940.
8. Ericsson UB, Hallberg BM, Detitta GT, Dekker N, Nordlund P. *Anal. Biochem* 2006;357:289–298. [PubMed: 16962548]
9. McPherson A, Cudney B. *J. Struct. Biol* 2006;156:387–406. [PubMed: 17101277]
10. Tomcova I, Smanova IK. *J. Cryst. Growth* 2007;306:383–389.
11. Cudney R, Patel S, Weisgraber K, Newhouse Y, McPherson A. *Acta Crystallogr. Sec. D* 1994;50:414–423.
12. Lu J, Wang XJ, Ching CB. *Cryst. Growth Des* 2003;3:83–87.
13. Berger BW, Blamey CJ, Naik UP, Bahnson BJ, Lenhoff AM. *Cryst. Growth Des* 2005;5:1499–1507.
14. Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ, Wasney GA, Yeung R, Arrowsmith C, Ball LJ, Berglund H, Hui R, Marsden BD, Nordlund P, Sundstrom M, Weigelt J, Edwards AM. *Proc. Natl. Acad. Sci. USA* 2006;103:15835–15840. [PubMed: 17035505]
15. Becker M, Stubbs MT, Huber R. *Protein Sci* 1998;7:580–586. [PubMed: 9541389]
16. Guan RJ, Wang M, Liu XQ, Wang DC. *J. Cryst. Growth* 2001;231:273–279.
17. Berger BW, Gendron CM, Lenhoff AM, Kaler EW. *Protein Sci* 2006;15:2682–2696. [PubMed: 17088325]
18. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warren P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrook-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock RE, Lory S, Olson MV. *Nature* 2000;406:959–964. [PubMed: 10984043]
19. Bodey GP, Bolivar R, Fainstein V, Jadeja L. *Rev. Infect. Dis* 1983;5:279–313. [PubMed: 6405475]
20. Gomez MI, Prince A. *Curr. Opin. Pharmacol* 2007;7:244–51. [PubMed: 17418640]
21. Mesaros N, Nordmann P, Plesiat P, Roussel-Delvallez M, Van Eldere J, Glupczynski Y, Van Laethem Y, Jacobs F, Lebecque P, Malfroot A, Tulkens PM, Van Bambeke F. *Clin. Microbiol. Infect* 2007;13:560–578. [PubMed: 17266725]
22. Zhang RG, Skarina T, Katz JE, Beasley S, Khachatryan A, Vyas S, Arrowsmith CH, Clarke S, Edwards A, Joachimiak A, Savchenko A. *Structure* 2001;9:1095–1106. [PubMed: 11709173]
23. Zimmerman MD, Chruszcz M, Koclega KD, Otwinowski Z, Minor W. *Acta Crystallogr. Sec. A* 2005;61:c178–c179.
24. Rosenbaum G, Alkire RW, Evans G, Rotella FJ, Lazarski K, Zhang RG, Ginell SL, Duke N, Naday I, Lazarz J, Molitsky MJ, Keefe L, Gonczy J, Rock L, Sanishvili R, Walsh MA, Westbrook E, Joachimiak A. *J. Synchr. Rad* 2006;13:30–45.
25. Otwinowski Z, Minor W. *Methods Enzymol* 1997;276:307–326.
26. Hendrickson WA. *Science* 1991;254:51–58. [PubMed: 1925561]
27. Schneider TR, Sheldrick GM. *Acta Crystallogr. Sec. D* 2002;58:1772–1779.
28. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. *Acta Crystallogr. Sec. D* 2006;62:859–866.
29. de la Fortelle E, Bricogne G. *Methods Enzymol* 1997;276:472–494.
30. Terwilliger TC. *J. Synchr. Radiat* 2004;11:49–52.
31. Terwilliger TC. *Acta Crystallogr. Sec. D* 2002;58:1937–1940.
32. Terwilliger TC. *Methods Enzymol* 2003;374:22–37. [PubMed: 14696367]

33. Perrakis A, Morris R, Lamzin VS. *Nat. Struct. Biol* 1999;6:458–463. [PubMed: 10331874]
34. Emsley P, Cowtan K. *Acta Crystallogr. Sec. D* 2004;60:2126–2132.
35. Vagin A, Teplyakov A. *J. Appl. Crystallogr* 1997;30:1022–1025.
36. Murshudov GN, Vagin AA, Dodson EJ. *Acta Crystallogr. Sec. D* 1997;53:240–255.
37. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sec. D* 1994;50:760–763.
38. Winn MD, Isupov MN, Murshudov GN. *Acta Crystallogr. Sec. D* 2001;57:122–133.
39. Painter J, Merritt EA. *Acta Crystallogr. Sec. D* 2006;62:439–450.
40. Painter J, Merritt EA. *J. Appl. Crystallogr* 2006;39:109–111.
41. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. *Proteins* 2003;50:437–450. [PubMed: 12557186]
42. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. *J. Appl. Crystallogr* 1993;26:283–291.
43. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C. *Acta Crystallogr. Sec. D* 2002;58:899–907.
44. DeLano WL. The PyMOL Molecular Graphics System.
45. Chruszcz M, Cymborowski M, Gawlicka-Chruszcz A, Yasukawa S, Ferrara JD, Minor W. *Acta Crystallogr. Sec. C* 2004;60:o868–o871.
46. Leesong M, Henderson BS, Gillig JR, Schwab JM, Smith JL. *Structure* 1996;4:253–264. [PubMed: 8805534]
47. Benning MM, Wesenberg G, Liu R, Taylor KL, Dunaway-Mariano D, Holden HM. *J. Biol. Chem* 1998;273:33572–33579. [PubMed: 9837940]
48. Dillon SC, Bateman A. *BMC Bioinformatics* 2004;5:109. [PubMed: 15307895]
49. Ponstingl H, Kabir T, Thornton JM. *J. Appl. Crystallogr* 2003;36:1116–1122.
50. Laskowski RA, Watson JD, Thornton JM. *Nucleic Acids Res* 2005;33:W89–93. [PubMed: 15980588]
51. Thoden JB, Zhuang Z, Dunaway-Mariano D, Holden HM. *J. Biol. Chem* 2003;278:43709–43716. [PubMed: 12907670]
52. Zhuang Z, Song F, Zhang W, Taylor K, Archambault A, Dunaway-Mariano D, Dong J, Carey PR. *Biochemistry* 2002;41:11152–11160. [PubMed: 12220180]
53. Hunt MC, Alexson SE. *Prog. Lipid. Res* 2002;41:99–130. [PubMed: 11755680]
54. Thompson JD, Higgins DG, Gibson TJ. *Nucleic Acids Res* 1994;22:4673–4680. [PubMed: 7984417]
55. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. *J. Mol. Biol* 1990;215:403–410. [PubMed: 2231712]
56. Gouet P, Robert X, Courcelle E. *Nucleic Acids Res* 2003;31:3320–3323. [PubMed: 12824317]
57. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. *Proc. Natl. Acad. Sci. USA* 2001;98:10037–10041. [PubMed: 11517324]
58. Holst M, Baker N, Wang F. *J. Comput. Chem* 2001;22:475–475.
59. Thoden JB, Holden HM, Zhuang Z, Dunaway-Mariano D. *J. Biol. Chem* 2002;277:27468–27476. [PubMed: 11997398]
60. Schujman GE, Guerin M, Buschiazzi A, Schaeffer F, Llarrull LI, Reh G, Vila AJ, Alzari PM, de Mendoza D. *EMBO J* 2006;25:4074–4083. [PubMed: 16932747]
61. Jelsch C, Longhi S, Cambillau C. *Proteins* 1998;31:320–333. [PubMed: 9593202]
62. Zhang XJ, Wozniak JA, Matthews BW. *J. Mol. Biol* 1995;250:527–552. [PubMed: 7616572]
63. Stura EA, Satterthwait AC, Calvo JC, Kaslow DC, Wilson IA. *Acta Crystallogr. Sec. D* 1994;50:448–455.
64. Hajduk PJ, Greer J. *Nat. Rev. Drug Discov* 2007;6:211–219. [PubMed: 17290284]

Figure 1.

Sequence alignment, done with CLUSTALW⁵⁴, of PA5185 (2AV9) and five other homologous sequences with corresponding structures in the PDB. The other structures were selected using BLAST⁵⁵. The sequences are annotated with the secondary structure identified from the structure of PA5185. The figure was prepared using ESPRIPT⁵⁶. Asp26 is marked with a green square.

Figure 2.

(A) A ribbon diagram showing monomers forming the dimer of PA5185. One of the monomers is shown in red, second is painted according to secondary structure elements. (B) The surface of the dimer colored with the electrostatic charge distribution as calculated by the APBS plugin^{57, 58} of PYMOL⁴⁴ (blue, positive charge; red, negative charge), shown in two different orientations. (C) The ribbon diagram of the PA5185 tetramer. (D) The surface of the tetramer colored with the electrostatic charge distribution shown in two different orientations.

Figure 3. Binding of sulfate groups in the monoclinic form of PA5185 (PDB code: 2AV9). (A) Loop formed by residues Gly92-Ser95. (B) Sulfate ion located on a non-crystallographic twofold axis and bridging Arg81 residues from two different protein chains. (C) Sulfate ion 113 located in the vicinity of three different protein chains (J, G, F).

Figure 4.

(A) Superposition of PA5185 monomer (in light green) and 4-hydroxybenzoyl CoA thioesterase59 from *Pseudomonas* sp. strain CBS-3 (PDB code: 1LO7) shown in light blue. Sulfate ion and 4-hydroxyphenacyl CoA are shown in stick representation. (B) Superposition of PA5185 (blue) and 4-hydroxybenzoyl CoA thioesterase (green) dimers shown in two different orientations. Sulfate ion is shown in red and 4-hydroxyphenacyl CoA in yellow.

Figure 5.

(A) Superposition of PA5185 chains in two different conformations. The conformation 'h' is shown in green and conformation 'l' in light blue. (B) Movement of Gln57 during change of the protein conformation. (C) Tetramer of PA5185 with loops changing conformation shown in red. (D) The three different types of PA5185 tetramers found in the asymmetric unit of the monoclinic crystal form. Capital letters correspond to different protein chains, while the subscripts 'h' and 'l' refer to the conformation of the specified chain.

Figure 6.

Crystals of PA5185 grown using vapor diffusion method. (A) Crystals of the PA5185 monoclinic form (space group: C2; PDB code: 2AV9) grown using a well solution containing 25% w/v PEG3350, 0.05 M ammonium sulfate and 0.1 M Bis-Tris pH=5.5. (B) Crystals grown in conditions without (NH₄)₂SO₄, and in the presence of 0.05 M MES pH=5.9. The needle shaped crystals were identified to be orthorhombic (space group: C222; PDB code: 2O5U). (C) Crystals of PA5185 grown in the presence of 0.01 M HEPES pH 7.0. Block shaped crystals belong to hexagonal system (space group: P3221; PDB code: 2O6U). (D) Crystal grown in solution containing 0.5% w/v NDSB-201. The needle shaped crystals are orthorhombic (space group: P2221; PDB code: 2O6T).

A **B**

Figure 7.

(A) Crystals grown using a well solution containing 25% w/v PEG3350, 0.1 M Bis-Tris pH=5.5 and 10 mM ATP (A) or 10 mM EPPS (B). In solutions containing EPPS, only needle shaped crystals are present after six days while block and prism shaped crystals disappear.

Figure 8.

Crystals of PA5185 grown with a well solution containing 25% w/v PEG3350, 0.1 M Bis-Tris pH=5.5 and 1 mM lithium myristoyl-CoA (A); 1 mM sodium CoA (B); or 1 mM lithium benzoyl-CoA (C). Changes of the PEG concentration also affect crystal morphology. Needle shaped crystals (D) were grown in condition similar to this presented in part (C) of the figure, but using a lower PEG 3350 concentration (20% w/v).

Figure 9. Superposition of two different conformations of PA5185. (A) Most probably three loops (11, 12 and 13) participate in binding of the substrate and/or release of the substrate. A fragment of a superimposed substrate (from 1LO7) is presented in stick representation. (B) Changes in dimer contacts in the region of the putative active site. 4-Hydroxyphenacyl CoA and sulfate ions are shown as sticks.

Table 1

Summary of data collection, processing, and refinement statistics. Data for the highest resolution shells are given in parentheses. The two crystallization conditions marked with asterisks could be treated as functionally equivalent, as HEPES was used as the buffering agent in the concentrated protein solution.

PDB code	2AV9	2O5U	2O6T	2O6B	2O6U
Additives	Original crystallization conditions ($(\text{NH}_4)_2\text{SO}_4$)				
pH ^{**}	5.7	MES	NDSB-201	-*	HEPES*
Data collection					
Beamline	19BM	19ID	19ID	19ID	19BM
Wavelength (Å)	0.9794 0.9793 0.9642	0.9794	0.9794	0.9794	0.9792
Unit cell (Å, °)	a=241.7, b=64.4, c=117.5, β=105.4	a=124.4, b=148.2, c=58.3 C222	a=58.9, b=97.1, c=191.4 P222 ₁	a=92.7, c=86.0 P3 ₂ 21	a=93.1, c=85.7 P3 ₂ 21
Space group	C2	54	55	62	62
Solvent content (%)	45	3	6	2	2
Number of protein chains in AU	12	50.0–1.9	40.0–2.55	25.0–3.2	37.9–3.0
Resolution range (Å)	37.0–2.4	1.97–1.90	2.62–2.55	3.31–3.20	3.11–3.00
Highest resolution shell (Å)	2.49–2.40	41694	36563	6998	8614
Unique reflections	69141 68671 68766	47(3.6)	6.8(4.4)	8.1(5.5)	8.0(8.3)
Redundancy	4.1(3.6) 4.1(3.6) 4.0(3.2)	98.4(87.9)	99.7(98.2)	96.4(83.4)	97.8(98.9)
Completeness (%)	99.6(96.4) 99.8(98.9) 99.7(97.2)	0.091(0.427)	0.121(0.429)	0.108(0.553)	0.086(0.499)
R _{merge}	0.066(0.348)	17.2(1.9)	16.5(2.3)	18.0(1.8)	23.4(2.5)
Average I/σ (I)	12.6(2.2) 12.6(2.2) 11.6(1.6)				
Refinement					
Highest resolution shell (Å)	2.46–2.40	1.96–1.90	2.62–2.55	3.29–3.20	3.09–3.00
R	0.204(0.237)	0.191(0.230)	0.180(0.248)	0.218(0.306)	0.195(0.238)
R _{free}	0.253(0.326)	0.223(0.278)	0.248(0.413)	0.262(0.323)	0.238(0.285)
Mean B value (Å ²)	42.5	22.7	21.3	88.4	61.2
RMS deviation bond lengths (Å)	0.008	0.020	0.018	0.018	0.017
RMS deviation angles (°)	1.9	1.7	1.7	1.8	1.8
Ramachandran plot					
Most favored regions (%)	93.7	95.4	93.6	81.2	93.0
Additional allowed regions (%)	6.1	4.3	6.0	18.0	6.2
Generously allowed regions (%)	0.2	0.3	0.4	0.8	0.8

^{**}The pH of the solution was determined after mixing the well solution with the protein storage buffer (500 mM NaCl and 10 mM HEPES pH=7.5) in a 1:1 ratio. Ramachandran plot statistics were calculated with *PROCHECK*. (AU - asymmetric unit.)

Table 2

Proteins that display highest structural similarity to PA5185. The structures were identified with *PROFNC* and *SSM*. All proteins reported in the table are tetrameric according to calculations done using the Protein Quaternary Structure (PQS) server (<http://pqs.ebi.ac.uk/>).⁴⁹

PDB code	Organism	Sequence identity (%)	Number of aligned residues	Z-score	RMS D (Å)	Space group	Number of chains in AU	Solvent content (%)
2EGR	<i>Aquifex aeolicus</i>	18	121	8.1	1.4	P6 ₄	2	53
2ALI	<i>Pseudomonas aeruginosa</i>	33	126	7.9	1.5	I222	1	35
1S5U	<i>Escherichia coli</i>	19	124	7.7	1.4	P1	8	43
2OAF	<i>Jannaschia</i> sp.	25	130	6.6	1.7	P6 ₁ 22	2	52
2OIW	<i>Bacillus stearothermophilus</i>	22	117	7.0	1.6	P2 ₁ 2 ₁ 2	4	42
2FUJ	<i>Xanthomonas campestris</i>	25	107	6.7	1.4	I23	1	63
1LO9	<i>Pseudomonas</i> sp.	17	127	6.2	1.7	I222	1	40
1Z54	<i>Thermus thermophilus</i>	18	122	6.5	1.7	P2 ₁	4	41
2HX5	<i>Prochlorococcus marinus</i>	23	125	5.4	1.8	I23	1	49
2GF6	<i>Sulfolobus</i>	19	120	6.6	1.7	P222 ₁	4	49
2C9E	<i>Solfataricus</i>	30	122	6.9	1.8	C2	4	40
2NUJ	<i>Thermus thermophilus</i> <i>Jannaschia</i> sp.	24	134	6.7	1.9	P4 ₁ 22	2	61

Table 3

Crystal contacts in PA5185 polymorphs. The contacts were calculated using program *CONTACT* from *CCP4* package, and contacts were considered only where hydrogen bond interactions with distances between donor and acceptor were shorter than 3.2 Å. Contacts bridged by water molecules were omitted.

	C2 (2AV9)	C222 (2O5U)	P222 ₁ (2O6T)	P3 ₂ 21 (2O6U)
Crystal contacts	Arg6A-Asn93A(-x+1/2, y-1/2, -z+1) Arg6A-Leu91A(-x+1/2, y-1/2, -z+1) Arg90A-Gln107A(-x+1/2, y-1/2, -z+1) Asn93A-Glu109A(-x+1/2, y+1/2, -z+1) Ser145A-Ser125B(-x+1/2, y+1/2, -z+1) His14B-Glu134F(x-1/2, y+1/2, z) Arg81B-Glu105J(x-1/2, y+1/2, z) Arg81B-Glu50F(x-1/2, y+1/2, z) Glu105B-Arg81J(x-1/2, y+1/2, z) Gly106B-Gln107J(x-1/2, y+1/2, z) Gln107B-Gly106J(x-1/2, y+1/2, z) Glu122B-Arg108C(-x+1/2, y-1/2, -z+1) Glu122B-Glu109C(-x+1/2, y-1/2, -z+1) Gln133B-Arg124K(-x+1/2, y-1/2, -z+1) Glu134B-Ser125K(-x+1/2, y-1/2, -z+1) Arg108C-Glu122B(-x+1/2, y+1/2, -z+1) Phe15E-Glu105F(x-1/2, y+1/2, z) Glu50E-Arg81J(x-1/2, y+1/2, z) Glu105E-Phe15F(x-1/2, y+1/2, z) Gly105E-Glu10F(x-1/2, y+1/2, z) Asp137K-Ala138K(-x+1, y, -z+1)	Glu10A-Arg124C(-x, -y, z) Asn93A-Arg136C(-x, y, -z) Glu109A-Asp55C(-x, y, -z+1) Ala130A-Asp137C(-x, y, -z) Ile131A-Gln133C(-x, y, -z) Gln133A-Gln133C(-x, y, -z) Asn93B-Gln147B(-x, y, -z+1) Asn93B-Arg51C(-x, y, -z+1) Gln133B-Ser145B(-x, y, -z+1) Arg136B-Ser145B(-x, y, -z+1) Ser145B-Asp137B(-x, y, -z+1) Gln143A-Ile27C(x, -y, -z+1)	Glu10A-Arg127E(x, -y+1, -z) Asn93A-Gln143G(x+1, y, z) Asn93A-Arg136G(x+1, y, z) Ala130A-Asp137G(x+1, y, z) Ser145A-Pro128E(x, -y+1, -z) Gln133E-Gln133I(x+1, y, z) Gln133E-Ile131I(x+1, y, z) Gln143E-Asn93I(x+1, y, z) Arg127G-Glu10I(x, -y+1, -z) Asn93I-Arg136E(x-1, y, z)	Arg6A-Glu109B(-y, x-y+1, z-1/3) Pro75A-Arg108B(-y, x-y+1, z-1/3)