Tavarekere S. Girish,^a Vivek B,^b
Melwin Colaco,^b Sandra
Misquith^b and B. Gopal^{a*}

^aMolecular Biophysics Unit, Indian Institute of
Science, Bangalore 560 012, India, and

^bDepartment of Chemistry, St Joseph's College,
Bangalore 560 027, India

Correspondence e-mail:
bgopal@mbu.iisc.ernet.in

Received 25 September 2012

Accepted 5 December 2012

PDB Reference: *S. aureus* amidohydrolase
SACOL0085, 4ewt

Structure of an amidohydrolase, SACOL0085, from methicillin-resistant *Staphylococcus aureus* COL

Staphylococcus aureus is an opportunistic pathogen that rapidly acquires resistance to frontline antibiotics. The characterization of novel protein targets from this bacterium is thus an important step towards future therapeutic strategies. Here, the crystal structure of an amidohydrolase, SACOL0085, from *S. aureus* COL is described. SACOL0085 is a member of the M20D family of peptidases. Unlike other M20D peptidases, which are either monomers or dimers, SACOL0085 adopts a butterfly-shaped homotetrameric arrangement with extensive intersubunit interactions. Each subunit of SACOL0085 contains two Mn²⁺ ions at the active site. A conserved cysteine residue at the active site distinguishes M20D peptidases from other M20 family members. This cysteine, Cys103, serves as bidentate ligand coordinating both Mn²⁺ ions in SACOL0085.

1. Introduction

The staphylococci are a diverse group of pathogenic bacteria that are implicated in a diverse range of diseases ranging from minor skin infections to life-threatening bacteraemia (Cunningham *et al.*, 1996; Maki, 1981). Since the emergence of methicillin-resistant *Staphylococcus aureus* (MRSA), this pathogen has continued to acquire antimicrobial resistance to the point where some isolates are resistant to many frontline antimicrobial agents (Mwangi *et al.*, 2007). The structural and biochemical characterization of potential drug-target proteins is thus an important input for future drug-discovery programs. The specific inhibition of bacterial proteases constitutes a significant component of these drug-development efforts. An analysis of the genome sequences of several *S. aureus* strains revealed several encoded unannotated proteases in addition to the annotated genes (Highlander *et al.*, 2007; Girish & Gopal, 2010). Here, we describe the crystal structure of an *S. aureus* amidohydrolase, SACOL0085.

SACOL0085 belongs to the M20D family of peptidases. Other members of this enzyme family for which three-dimensional structures are known include *Bacillus subtilis* YXEP (PDB entry 1ysj; Midwest Center for Structural Genomics, unpublished work) and ILL2 (PDB entry 1xmb; Bitto *et al.*, 2009), an auxin-conjugate amidohydrolase from *Arabidopsis thaliana*. These structures revealed that M20D proteases also adopt a two-domain architecture like other enzymes in the M20/Acy-I family (Rowell *et al.*, 1997; Jozic *et al.*, 2002; Lundgren *et al.*, 2007). The N-terminal domain in these enzymes, which is also referred to as the catalytic domain, retains residues involved in metal coordination and substrate binding. The second domain, which is often smaller, is referred to as a satellite or dimerization domain because of its role in quaternary association. SACOL0085 shows significant sequence similarity to both *B. subtilis* YXEP and *A. thaliana* ILL2. While the overall structure is similar to those of other M20D proteases, the quaternary arrangement differs across these enzymes, ranging from monomeric (ILL2) and dimeric (YXEP) to tetrameric (SACOL0085) association (Bitto *et al.*, 2009).

A consistent finding from biochemical and structural studies is that enzymes in the M20D family require two metal ions for their catalytic activity (Bitto *et al.*, 2009). Most of these enzymes prefer Mn²⁺ as a metal cofactor. While a few M20D enzymes have been noted to have a Cu²⁺ ion in the active site (LeClere *et al.*, 2002), the structure of *B. subtilis* YXEP contained two Ni²⁺ ions bound in the active site. The

Table 1

Data, refinement and model statistics.

Values in parentheses are for the outer shell.

Data collection	
Resolution (Å)	2.10 (2.21–2.10)
Space group	<i>P</i> 1
Unit-cell parameters (Å, °)	<i>a</i> = 44.62, <i>b</i> = 120.11, <i>c</i> = 132.41, <i>α</i> = 115.40, <i>β</i> = 94.64, <i>γ</i> = 96.55
Total No. of observations	468901 (59549)
Total No. of unique observations	133741 (18498)
<i>R</i> _{merge} [†] (%)	11.1 (40.4)
Mean <i>I</i> / <i>σ</i> (<i>I</i>)	7.5 (2.9)
Completeness (%)	93.8 (88.6)
Multiplicity	3.5 (3.2)
Refinement statistics	
<i>R</i> _{work} [‡] (%)	20.03
<i>R</i> _{free} [§] (%)	22.94
No. of residues	1542
No. of water molecules	471
Ligands	
Mn ²⁺ ions	8
Polyethylene glycol (PEG)	1
1-Deoxy-1-thio-heptaethylene glycol (PE7)	3
R.m.s.d., bond lengths (Å)	0.010
R.m.s.d., bond angles (°)	1.205
Ramachandran plot analysis	
Most favoured (%)	94.7
Additionally allowed (%)	4.9
Disallowed (%)	0.3

[†] $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th observation of reflection *hkl* and $\langle I(hkl) \rangle$ is the average intensity. [‡] $R_{\text{work}} = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$. [§] R_{free} was calculated as for R_{work} but on 5% of the data excluded from the refinement calculation.

structure of *A. thaliana* ILL2 represents the apo form of this enzyme without a bound metal cofactor (Bitto *et al.*, 2009). However, based on a comparison between the YXEP and ILL2 structures, five residues were identified to play a role in metal coordination (Bitto *et al.*, 2009). These include a cysteine, a glutamate and three histidine residues. The cysteine residue serves as a bidentate ligand coordinating both metal ions. This observation, primarily based on the *B. subtilis* YXEP structure, was proposed to distinguish enzymes of the M20D family from the other M20 proteases, which have an aspartate residue with an equivalent role and position (Lundgren *et al.*, 2007; Lindner *et al.*, 2003).

The structure of SACOL0085 was determined at 2.1 Å resolution. A distinctive feature of the SACOL0085 structure is a butterfly-shaped tetrameric quaternary arrangement with extensive inter-subunit interactions. The active site of SACOL0085 has two bound Mn²⁺ ions. A comparison between the four monomeric units of the SACOL0085 tetramer shows variations in the inter-domain orientation. This feature is likely to aid domain movements and may play a role in the regulation of catalytic activity.

2. Materials and methods

2.1. Cloning, overexpression and purification of recombinant SACOL0085

The SACOL0085 gene (UniProt ID Q5HJR7; 1.179 kb) was PCR-amplified from the genomic DNA of *S. aureus* COL using 5'-TGTCGCTAGCATGAATCAACAATTAATTGAAACT-3' and 5'-CGGTCTCGAGGTTATCTCCTTTAAGGTAATCTAAAA-3' as the forward and reverse primers, respectively. The gene was subsequently cloned into the pET15b bacterial expression vector between *Nhe*I and *Xho*I restriction sites. The recombinant protein was expressed in *Escherichia coli* BL21 (DE3) cells with a hexahistidine tag at the N-terminus. When cell growth reached an optical density

of 0.6, 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce protein overexpression. Post-induction, the cells were grown at 290 K for 6–8 h. The cells were harvested by spinning at 6000 rev min^{−1} for 10 min. The cell pellet was resuspended in lysis buffer (50 mM Tris–HCl pH 7.5, 250 mM NaCl) and lysed by sonication. The supernatant from this stage was incubated with Ni²⁺–NTA affinity beads (Sigma–Aldrich) for 1 h at 277 K. The bound protein was eluted with a gradient of imidazole concentration ranging from 5 to 200 mM in a buffer consisting of 50 mM Tris–HCl pH 7.5, 250 mM NaCl. The partially purified protein was further purified on a Sephacryl S-200 gel-filtration column equilibrated with a buffer consisting of 25 mM Tris–HCl pH 7.5, 200 mM NaCl, 1 mM MnCl₂, 1 mM DTT.

2.2. Crystallization, data collection and structure determination

Initial crystallization conditions were determined using commercial crystallization kits (Hampton Research). Crystallization trials were performed using both the hanging-drop method and the microbatch-under-oil method. The protein concentration was maintained at 10.0 mg ml^{−1} at 293 K. Each drop (4 μl) consisted of 2 μl protein solution and 2 μl crystallization reagent. Initial crystals were obtained in a condition consisting of 0.2 M magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 30% PEG 400. These conditions were subsequently modified to yield diffraction-quality crystals. The inclusion of 2 mM MnCl₂ in the protein-purification buffers improved both the size and the diffraction quality of the crystals. The crystals were flash-cooled in liquid nitrogen directly from the crystallization wells, as the crystallization condition contained a cryoprotectant (PEG 400). A complete data set was collected on beamline BM-14 at the European Synchrotron Radiation Facility (ESRF). The diffraction data were processed using *MOSFLM* (Leslie, 2006) and scaled using *SCALA* (Evans, 2006). The phase information used to solve the structure was obtained by molecular replacement (MR) with the program *Phaser* (McCoy *et al.*, 2007). The structure of *B. subtilis* YXEP was used as a search model (PDB entry 1ysj; 34.0% sequence identity) in an MR trial performed using the *PHENIX* suite (Adams *et al.*, 2010). The phase information was sufficient to model the SACOL0085 structure, which was subsequently refined using *REFMAC5* from the *CCP4* suite of programs (Winn *et al.*, 2003,

Figure 1
Overall structure of SACOL0085. The structure of SACOL0085 reveals a tightly packed tetramer.

2011). Iterative cycles of manual model building were performed using *Coot* (Emsley *et al.*, 2010). The use of TLS restraints was important to achieve convergence in refinement. Three TLS restraint groups were applied per subunit based on input from the *TLSMD* server (Painter & Merritt, 2006).

3. Results and discussion

3.1. The structure of SACOL0085

The structure of SACOL0085 was determined at 2.1 Å resolution. The model was refined to an *R* factor of 20.0% (*R*_{free} of 22.9%). The data, refinement and model statistics are compiled in Table 1. The structure of SACOL0085 consists of four monomeric subunits. Each of the four monomers in the asymmetric unit comprises residues 1–389. Three residues from the C-terminus (390–392) in each chain as well as the stretches of residues 320–324, 354–359 and 387–389 in chain *B* could not be modelled owing to poor electron density.

Prominent electron density observed in the active site of each subunit could be modelled by two Mn²⁺ ions. An additional tube-like electron density observed in the electron-density map could be best interpreted as a polyethylene glycol (PEG) moiety. The PEG molecules are involved in contacts with symmetry-related molecules and may assist in crystal packing.

Each monomer of SACOL0085 consists of a typical two-domain architecture. This feature is common to most M20/Acy-I proteases (Girish & Gopal, 2010; Jozic *et al.*, 2002). The larger catalytic domain in SACOL0085 consists of residues 1–184 and 298–389. The catalytic domain has a $\alpha\beta\alpha$ sandwich fold in which a large central β -sheet is sandwiched between α -helices on either side. The large central β -sheet consists of eight β -strands arranged in the order b1–b2–b4–b3–b5–b8–b6–b7. The smaller domain (the satellite/dimerization domain in M20 terminology) comprises residues 185–297. This domain adopts an $\alpha\beta$ fold with a four-stranded β -sheet (b'4–b'1–b'3–b'2) flanked by two α -helices (a'1–a'2). The β -strands are arranged in an antiparallel manner. Both β -strands and α -helices in the satellite

(a)

(b)

Figure 2

Sequence and structural conservation in M20D peptidases. (a) Sequence alignment of SACOL0085 and M20D homologues for which crystal structures have been determined. (b) The sequence conservation mapped onto the SACOL0085 structure using the *ConSurf* server (<http://consurf.tau.ac.il/>).

domain show extensive interactions with their counterparts from neighbouring monomers, thereby contributing to the stability of the tetrameric arrangement (Fig. 1). The inter-domain region at the hinge consists of two short β -strands (b9–b10).

3.2. The tetrameric arrangement of SACOL0085

The four monomers of SACOL0085 in the asymmetric unit reveal extensive interactions and assume a tetrameric ‘butterfly’-like structure. The satellite domain from each subunit forms the body, while the catalytic domains are oriented sideways to form the wings of the ‘butterfly’ structure. An analysis of the quaternary structure using the *PISA* server at EBI suggests a calculated ΔG_{diss} of 36.8 kcal mol^{−1}. The tetrameric arrangement in SACOL0085 buries a total interface area of 15 530 Å². This analysis also suggests a dimer of dimers. The stable dimeric arrangements consisting of *AB* and *CD* dimers with

Table 2 Analysis of the quaternary association in M20D proteases.			
Name/PDB code	Quaternary arrangement	Buried surface area at the interface (Å ²)	ΔG_{diss} (kcal mol ^{−1})
SACOL0085	Tetramer	15530	36.8
	Dimer <i>AB</i>	3920	6.2
	Dimer <i>CD</i>	3460	4.5
1ysj	Dimer	2630	3.9
1xmb	Monomer	—	—

calculated ΔG_{diss} values of 6.2 and 4.2 kcal mol^{−1}, respectively, form the tetramer (Figs. 1 and 2). The closest sequence homologues of SACOL0085 are *B. subtilis* YXEP and *A. thaliana* ILL2, an auxin-conjugate amidohydrolase (Fig. 2). Although both YXEP and ILL2 show significant sequence homology (34 and 32%, respectively) to SACOL0085, their quaternary arrangements are different (Table 2).

Figure 3
The active site of SACOL0085. (a) Stereoview of the superposition of the active site of the open (chain *D*; blue) and closed (chain *A*; green) conformations of SACOL0085. (b) Stereoview of the superposed active sites of SACOL0085 (chain *A*; green), *A. thaliana* ILL2 (PDB entry 1xmb; wheat) and *B. subtilis* YXEP (PDB entry 1ysj; orange).

This aspect of SACOL0085 can be rationalized by mapping the sequence conservation onto the structure of SACOL0085 (Fig. 2). The residues involved in oligomerization are poorly conserved across these homologues (Fig. 2).

3.3. The active site of SACOL0085

As in the case of the other two M20D enzyme structures, the active site of SACOL0085 is located in the vicinity of the inter-domain region. The loop connecting β -strand b7 to α -helix a5 forms one side of the cleft, while the other side and base of the cleft are defined by residues that coordinate the metal cofactor. The exit of this active-site cleft is exposed to the bulk solvent. Two Mn^{2+} ions could be modelled in each active site. The distance between the two Mn^{2+} ions ranges from 3.38 Å (chain A) to 3.21 Å (chain B). The Mn^{2+} ions are coordinated by five residues. While Cys103, Glu139 and His362 coordinate one Mn^{2+} ion, the other Mn^{2+} is coordinated by Cys103, His105 and His164 (Fig. 3). Water molecules complete the metal coordination. Cys103 serves as a bridging residue by coordinating both metal ions (Fig. 3). This cysteine (Cys103) is equivalent to Cys98 in *B. subtilis* YXEP and Cys137 in *A. thaliana* ILL2. This bridging cysteine in M20D proteases serves to distinguish these enzymes from other M20 family members (Bitto *et al.*, 2009). The similar active-site configuration and identical metal-coordinating residues in the ILL2 and SACOL0085 structures suggests that SACOL0085 could potentially function as an auxin-conjugate amidohydrolase in *S. aureus*.

3.4. Inter-domain orientation

Proteases of the M20 family demonstrate large differences in the orientation between the catalytic and the satellite/dimerization domains (Lundgren *et al.*, 2007). Large inter-domain motions have

Table 3

Variation in the inter-domain orientation of structurally characterized M20D proteases.

Protein	Calculated angle (°)
SACOL0085, chain A	139.8
SACOL0085, chain B	141.3
SACOL0085, chain C	145.0
SACOL0085, chain D	146.0
YXEP (PDB entry 1ysj)	141.5
ILL2 (PDB entry 1xmb)	147.3

previously been noted for many enzymes in this class of proteins (Girish & Gopal, 2010). An inference from these biochemical and structural data is that the location of the active site at the inter-domain region, in addition to the large variations in the inter-domain orientation, is crucial for these enzymes to hydrolyze a variety of substrates with different sizes and chemical properties. A structural comparison between the different subunits of SACOL0085 also reveals differences in the inter-domain orientation. The inter-domain angles between the different subunits of SACOL0085 were calculated by considering the centre of masses of the catalytic domain, the hinge region and the satellite domain. We note that although all four monomers have Mn^{2+} ions in the active site, variation in the inter-domain orientation reveals that chain A is in the most closed conformation, while chain D is in an open conformation (Fig. 4). An angular difference of about 5° was noted between the open and closed conformations. In this context it is worth noting that while YXEP and ILL2 have similar overall structures, with a C^α root-mean-square deviation (r.m.s.d.) of ~1.5 Å, the orientation between the catalytic and satellite domains varies in these two cases. While *B. subtilis* YXEP with two bound Ni^{2+} ions adopts a more closed

Figure 4

Variations in the orientation of the catalytic and satellite domains. (a) A comparison between the open (chain D; blue) and closed (chain A; green) conformations of SACOL0085. The centre of mass of the catalytic domain, the satellite domain and the hinge region were considered in calculating the domain movements in each subunit. (b) A comparison between the monomeric units of SACOL0085, ILL2 (PDB entry 1xmb; wheat) and YXEP (PDB entry 1ysj; orange).

conformation (inter-domain angle of 141°), *A. thaliana* ILL2, which does not have a metal ion in the active site, is more open (147°). The crystal structure of SACOL0085 suggests that metal-ion binding is not an essential prerequisite for inter-domain reorientation. While chain *A* of SACOL0085 is comparable to the YXEP structure (closed form), chain *D* is comparable to the open conformation, as in the case of ILL2 (Table 3, Fig. 4).

4. Conclusions

The structure of SACOL0085 reveals the extent of the structural plasticity in the active site of M20D proteases. The influence of the changes in the inter-domain orientations that are observed between the monomers of SACOL0085 on substrate and cofactor specificity remains to be established. Put together, the differences in the quaternary association and the choice of metal cofactor as well as substrate preferences suggest diverse functional roles and regulatory mechanisms for M20D proteases.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Bitto, E., Bingman, C. A., Bittova, L., Houston, N. L., Boston, R. S., Fox, B. G. & Phillips, G. N. Jr (2009). *Proteins*, **74**, 61–71.
- Cunningham, R., Cockayne, A. & Humphreys, H. (1996). *J. Med. Microbiol.* **44**, 157–164.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Girish, T. S. & Gopal, B. (2010). *J. Biol. Chem.* **285**, 29406–29415.
- Highlander, S. K. *et al.* (2007). *BMC Microbiol.* **7**, 99.
- Jozic, D., Bourenkow, G., Bartunik, H., Scholze, H., Dive, V., Henrich, B., Huber, R., Bode, W. & Maskos, K. (2002). *Structure*, **10**, 1097–1106.
- LeClere, S., Tellez, R., Rampey, R. A., Matsuda, S. P. & Bartel, B. (2002). *J. Biol. Chem.* **277**, 20446–20452.
- Leslie, A. G. W. (2006). *Acta Cryst.* **D62**, 48–57.
- Lindner, H. A., Lunin, V. V., Alary, A., Hecker, R., Cygler, M. & Ménard, R. (2003). *J. Biol. Chem.* **278**, 44496–44504.
- Lundgren, S., Andersen, B., Piskur, J. & Dobritzsch, D. (2007). *J. Biol. Chem.* **282**, 36037–36047.
- Maki, D. G. (1981). *Am. J. Med.* **70**, 719–732.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Mwangi, M. M., Wu, S. W., Zhou, Y., Sieradzki, K., de Lencastre, H., Richardson, P., Bruce, D., Rubin, E., Myers, E., Siggia, E. D. & Tomasz, A. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 9451–9456.
- Painter, J. & Merritt, E. A. (2006). *Acta Cryst.* **D62**, 439–450.
- Rowsell, S., Pauptit, R. A., Tucker, A. D., Melton, R. G., Blow, D. M. & Brick, P. (1997). *Structure*, **5**, 337–347.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Winn, M. D., Murshudov, G. N. & Papiz, M. Z. (2003). *Methods Enzymol.* **374**, 300–321.