



## OPEN ACCESS

# Crystal structure of the capsular polysaccharide synthesizing protein CapE of *Staphylococcus aureus*

Takamitsu MIYAFUSA\*†, Jose M. M. CAAVEIRO\*†, Yoshikazu TANAKA†<sup>1</sup>, Martin E. TANNER‡ and Kouhei TSUMOTO\*†§<sup>2</sup>

\*Medical Proteomics Laboratory, Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo 108-8639, Japan, †Department of Medical Genome Sciences, School of Frontier Sciences, The University of Tokyo, Minato-ku, Tokyo 108-8639, Japan, ‡Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, British Columbia V6T 1Z1, Canada, and §Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Tokyo 113-8656, Japan

## Synopsis

Enzymes synthesizing the bacterial CP (capsular polysaccharide) are attractive antimicrobial targets. However, we lack critical information about the structure and mechanism of many of them. In an effort to reduce that gap, we have determined three different crystal structures of the enzyme CapE of the human pathogen *Staphylococcus aureus*. The structure reveals that CapE is a member of the SDR (short-chain dehydrogenase/reductase) super-family of proteins. CapE assembles in a hexameric complex stabilized by three major contact surfaces between protein subunits. Turnover of substrate and/or coenzyme induces major conformational changes at the contact interface between protein subunits, and a displacement of the substrate-binding domain with respect to the Rossmann domain. A novel dynamic element that we called the latch is essential for remodelling of the protein–protein interface. Structural and primary sequence alignment identifies a group of SDR proteins involved in polysaccharide synthesis that share the two salient features of CapE: the mobile loop (latch) and a distinctive catalytic site (MxxxK). The relevance of these structural elements was evaluated by site-directed mutagenesis.

**Key words:** capsular polysaccharide, conformational change, pathogenic bacterium, SDR enzyme, UDP–sugar, X–ray crystallography

Cite this article as: Miyafusa, T., Caaveiro, J.M.M., Tanaka, Y., Tanner, M.E. and Tsumoto, K. (2013) Crystal structure of the capsular polysaccharide synthesizing protein CapE of *Staphylococcus aureus*. Biosci. Rep. 33(3), art:e00043.doi:10.1042/BSR20130017

## INTRODUCTION

*Staphylococcus aureus* is normally a harmless commensal bacterium residing in 25% of the adult population, yet is a dangerous human pathogen in susceptible individuals [1,2]. *S. aureus* is also notorious for its capacity to develop resistance against antibiotics. For example, methicillin- and vancomycin-resistant strains of *S. aureus* are among the major causes of nosocomial and community infections in the USA [3,4]. We are in urgent need of new targets and innovative strategies to effectively combat *S. aureus* and other dangerous pathogens [5].

The biosynthetic machinery that generates the CP (capsular polysaccharide) is absent in humans, and represents an attractive target to fight *S. aureus* [6–8]. CP forms a thick layer of carbohydrate on the cell surface conferring anti-phagocytic properties, and helping *S. aureus* to persist in the bloodstream of the infected host. More than 70% of clinical isolates of *S. aureus* belong to either the CP5 or the CP8 serotypes [8,9].

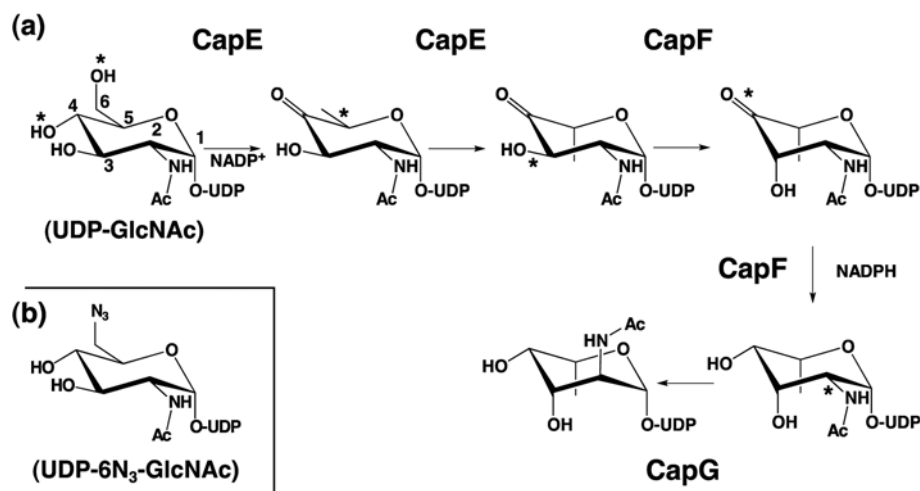
The basic structure of CP of serotypes CP5 and CP8 consists of alternating units of three types of monosaccharides: *N*-acetyl-L-fucosamine, *N*-acetyl-D-fucosamine, and *N*-acetyl-D-mannosamine uronic acid. The UDP-L-FucNAc (UDP-activated form of *N*-acetyl-L-fucosamine) is synthesized from its

**Abbreviations used:** CP, capsular polysaccharide; hUDGH, human UDP- $\alpha$ -D-glucose dehydrogenase; SDR, short-chain dehydrogenase/reductase; SEC, size-exclusion chromatography; UDP-6N<sub>3</sub>-GlcNAc, 6'-Azido-6'-deoxy-UDP- N-acetylglucosamine.

<sup>1</sup> Present address: Creative Research Initiative Sousei, Hokkaido University, Sapporo 001-0021, Japan

<sup>2</sup> To whom correspondence should be addressed (email tsumoto@ims.u-tokyo.ac.jp).

Atomic co-ordinates and structure factors of holo-CapE, holo-K126E and CapE with substrate analogue UDP-6N<sub>3</sub>-GlcNAc are deposited in the PDB under entry codes 3VVB, 3WVC, and 3W1V respectively.



**Figure 1** Biosynthetic pathway of UDP-L-FucNAc in *S. aureus*

(a) Synthesis of UDP-L-FucNAc requires the sequential activity of enzymes CapE, CapF and CapG. The enzyme CapE is a bifunctional enzyme catalysing the 4,6-dehydration of UDP-D-GlcNAc and a subsequently C5-epimerization. CapF catalyses the C3-epimerization of the previous intermediate, followed by the reduction of the keto-sugar at position C4. CapG catalyses a C2-epimerization to yield the final product UDP-L-FucNAc. The asterisks indicate the chemical groups subjected to enzymatic modification. This pathway is adapted from previous mechanistic studies [12,14]. (b) Structure of the substrate analogue UDP-6N<sub>3</sub>-GlcNAc, where the hydroxyl group at C6 of the substrate is replaced by an azide substituent [20].

glucosamine precursor (UDP-D-GlcNAc). This reaction has been described in *S. aureus* and several other pathogenic bacteria possessing CP such as *Streptococcus pneumoniae* and *Bacteroides fragilis* [10,11]. L-FucNAc is also utilized in the synthesis of the lipopolysaccharide of some Gram-negative bacteria [12,13].

The transformation of UDP-D-GlcNAc in UDP-L-FucNAc requires three enzymes (CapE, CapF and CapG) in *S. aureus* [12,14]. These three enzymes catalyse a total of five chemical reactions (Figure 1). The recently determined crystal structure of CapF unveiled a unique architecture composed of two distinctive domains [15]: a N-terminal domain belonging to the SDR (short-chain dehydrogenase/reductase) superfamily of proteins [16], and a C-terminal domain displaying a standard cupin fold [17]. A previous study of the homologous protein WbjB, together with the structural study of the enzyme CapF (just downstream of CapE), indicates that CapE is a bi-functional enzyme [14,15]. The two reactions catalysed by CapE are the C-4/C-6 dehydration and the C-5 epimerization of the UDP-sugar. The enzyme FlaA1 from *Helicobacter pylori* displays 40% homology to CapE and catalyses the same enzymatic reactions [18,19]. However, the key catalytic tyrosine residue of FlaA1 is replaced with methionine in CapE, suggesting mechanistic differences between them.

Herein we present the first crystal structure of CapE, revealing that this enzyme forms a stable and functional homo-hexamer. From the comparison of three crystal structures, we identified a novel and dynamic element unique to the capsular-polysaccharide synthesizing proteins that we called the latch. The special configuration of the active site of CapE, together with the unique features of the latch, exposed a distinctive group of enzymes of bacterial origin within the SDR superfamily sharing these two special elements.

## EXPERIMENTAL

### Substrate analogue

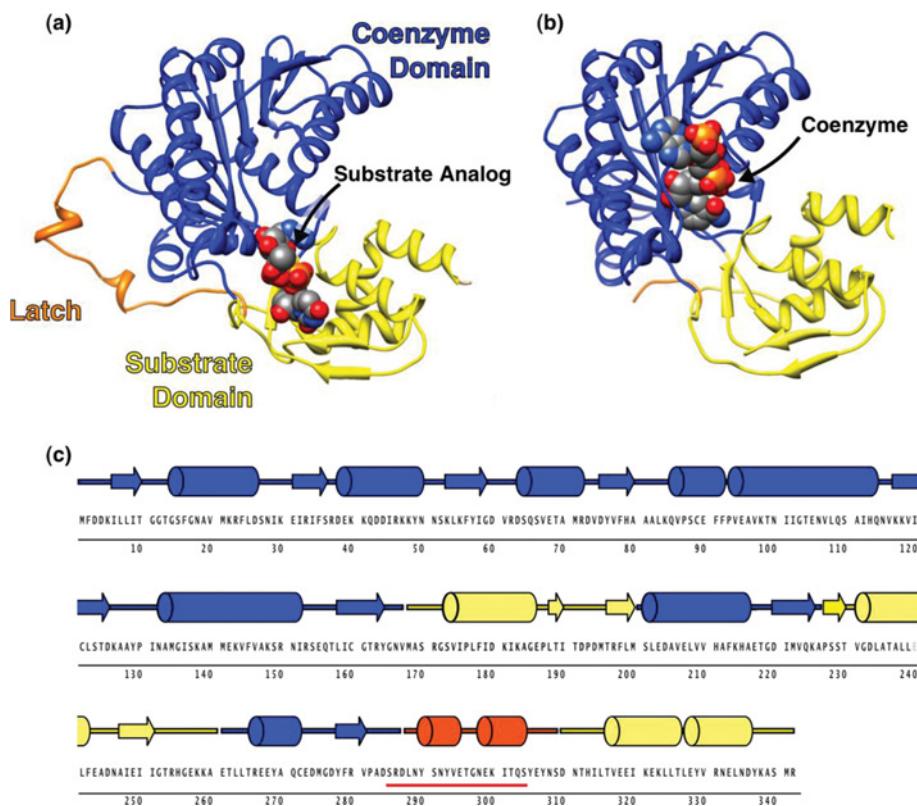
The synthesis of UDP-6N<sub>3</sub>-GlcNAc (6'-Azido-6'-deoxy-UDP-N-acetylglucosamine) was carried out as previously described [20].

### Protein expression and purification

CapE with a His<sub>6</sub> tag at the N-terminus was expressed in *Escherichia coli* BL21 (DE3) cells, and purified as described previously [15,21]. After SEC (size-exclusion chromatography), fractions of CapE were dialysed in 10 mM Tris/HCl (pH 9.0), 30 mM NaCl and 1 mM DTT (dithiothreitol), and concentrated with a 100 kDa Centrprep filtration unit (Millipore) prior to crystallization. For activity assays, protein fractions were stored at -20 °C in a solution supplemented with 40% (v/v) glycerol. Protein concentration was determined spectrophotometrically at 280 nm using the calculated molar extinction coefficient of the protein ( $\epsilon = 19\,200\text{ M}^{-1}\text{ cm}^{-1}$ ). Muteins of CapE were prepared with a Quick-Change kit following the instructions of the manufacturer. Protein expression and purification were carried out as above.

### Protein crystallization

Crystals of CapE suitable for X-ray diffraction analysis were obtained by the hanging drop method by mixing 1  $\mu\text{l}$  of fresh protein solution at 8 mg/ml and 1  $\mu\text{l}$  of crystallization solution. Crystals of wild-type CapE with coenzyme bound were obtained in 0.72 M sodium succinate (pH 7.0) and 1.4 M potassium formate. Rod-shaped crystals grew to an approximate size of



**Figure 2** Crystal structure of CapE

Structure of (a) CapE with substrate analogue UDP-6N<sub>3</sub>-GlcNAc bound, and (b) K126E mutein with coenzyme NADPH bound. The coenzyme-binding domain, substrate-binding domain, and latch are coloured in blue, yellow and orange, respectively. The ligands are depicted as spheres with CPK colours. The figure was prepared with CHIMERA [50]. (c) Primary and secondary structure of CapE. The underlined sequence (red line) corresponds to the disordered region in K126E.

30×30×100 μm<sup>3</sup> within 1 week. Crystals were transferred to a cryoprotectant solution consisting of mother liquor supplemented with 25% (v/v) glycerol, plunged into liquid N<sub>2</sub> and stored until data collection.

Crystals of CapE with substrate analogue were obtained by the co-crystallization method using 100 μM UDP-6N<sub>3</sub>-GlcNAc. Extensive screening with an Oryx8 robot (Douglas Instruments) [22] yielded a suitable solution composed of 100 mM Hepes/NaOH (pH 7.5) and 1.5 M Li<sub>2</sub>SO<sub>4</sub>. For freezing, single crystals were passed through a small drop of paratone (Hampton Research) and plunged in liquid N<sub>2</sub>. Crystals of mutein K126E were obtained in 100 mM Hepes (pH 7.5), 2% (v/v) PEG400 and 2.0 M ammonium sulfate. Suitable crystals were soaked in mother liquor supplemented with 25% (v/v) glycerol and 500 μM substrate (the substrate was not observed in the electron density) and stored in liquid N<sub>2</sub>.

### Data collection and refinement

Suitable crystals of CapE were mounted under a stream of cold nitrogen (100 K) at beamlines BL5A and AR-NE3A of the Photon Factory (Tsukuba, Japan). Data were processed with the program MOSFLM [23] and merged and scaled using the program SCALA

of the CCP4 program suite [24]. The structure of mutein K126E was determined by the method of molecular replacement with the program PHASER [25,26] using the coordinates of FlaA1 of *H. pylori* (PDB entry code 2gn4) as the search model. Coordinate refinement was carried out with PHENIX [27], REFMAC5 [28] and COOT [29]. Data were further refined with REFMAC5 using TLS parameterization. Each protomer was split in the three groups shown in Figure 2 of the Results and Discussion section [30]. The structure of the apo-CapE and apo-K126E were determined by molecular replacement using the structure of holo-CapE. Refinement was performed as above. Model quality was assessed with PROCHECK [31]. Data collection and refinement statistics are summarized in Table 1.

### Enzymatic assay

Enzymatic activity of CapE was monitored by the method described in Miyafusa et al. [15]. In a typical assay, substrate UDP-D-GlcNAc (Wako) at 200 μM was mixed with 2 μM CapE. The assay buffer was composed of 20 mM Tris/HCl (pH 8.0). Total volume was 100 μl. Assay mixtures were incubated at 37°C for 2 h, after which the reaction was stopped by addition of 100 μl of ice-cold phenol/chloroform/isoamyl alcohol at a 25:24:1 molar

**Table 1 Data collection and refinement statistics**

Values in parenthesis correspond to the highest resolution bin.

Properties	Wild-type		K126E
	UDP-6N <sub>3</sub> -GlcNAc	NADPH	NADPH
Data collection			
Space group	<i>P</i> 321	<i>P</i> 6 <sub>3</sub> 22	<i>H</i> 3 <sub>2</sub>
Unit cell			
Dimensions (Å)	a = b = 123.78, c = 104.46	a = b = 125.18, c = 101.79	a = b = 160.5, c = 80.05
Angles (°)	$\alpha = \beta = 90, \gamma = 120$	$\alpha = \beta = 90, \gamma = 120$	$\alpha = \beta = 90, \gamma = 120$
Wavelength (Å)	1.0000	1.0000	1.0000
Resolution range (Å)	34.8 – 2.10	34.05 – 2.80	40.13 – 2.20
Total observations	291899	103149	93553
Unique observations	54166	12071	20104
<i>I</i> / $\sigma$ ( <i>I</i> )	11.0 (4.1)	12.2 (2.8)	8.9 (2.5)
Completeness (%)	99.9 (100)	99.9 (100)	99.9 (100)
<i>R</i> <sub>merge</sub> (%)*	10.4 (33.8)	11.9 (79.8)	8.8 (49.8)
Multiplicity	5.4 (5.2)	8.5 (8.9)	4.7 (4.4)
Refinement			
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> (%)†	16.4/20.7	23.8/27.6	19.6/26.4
Number of protein chains	2	1	1
Number of protein residues	693	262	314
Number of protein atoms	5562	2137	2504
Number of ligands	18‡	1	6
Number of ligand atoms	168	48	73
Number of water molecules	385	3	82
B-factor, protein (Å <sup>2</sup> )	29.0	81.9	50.1
B-factor, ligands (Å <sup>2</sup> )	53.3	80.4	76.8
B-factor, water (Å <sup>2</sup> )	36.0	49.2	46.5
RMSD bonds (Å)	0.023	0.008	0.017
RMSD bonds (°)	2.19	1.363	2.01
Co-ordinate error (Å)	0.14	0.62	0.24
Ramachandran plot			
Preferred regions (%)	97.5	92.7	95.9
Allowed regions (%)	2.2	7.3	4.1
Outliers (%)	0.3	0.0	0.0
PDB codes	3W1V	3VVB	3VVC

\* $R_{\text{merge}} = \frac{\sum_{\text{hkl}} \sum_j |I(\text{hkl})_j - \langle I(\text{hkl}) \rangle|}{\sum_{\text{hkl}} \sum_j I(\text{hkl})_j}$ .

† $R_{\text{work}} = \frac{\sum_{\text{hkl}} |F(\text{hkl})_o - \langle F(\text{hkl})_c \rangle|}{\sum_{\text{hkl}} F(\text{hkl})_o}$ ;  $R_{\text{free}}$  was calculated as  $R_{\text{work}}$ , where  $F(\text{hkl})_o$  values were taken from 5% of data not included in the refinement.

‡This structure contains 14 sulfate ions from the crystallization solution.

ratio. The supernatant containing the sugars were mixed with 100  $\mu$ l of chloroform and analysed by HPLC using a CarboPac PA1 anion-exchange column (Dionex) as described previously [32]. The overall conversion to products was calculated from the consumption of substrate UDP-D-GlcNAc.

### Differential scanning calorimetry

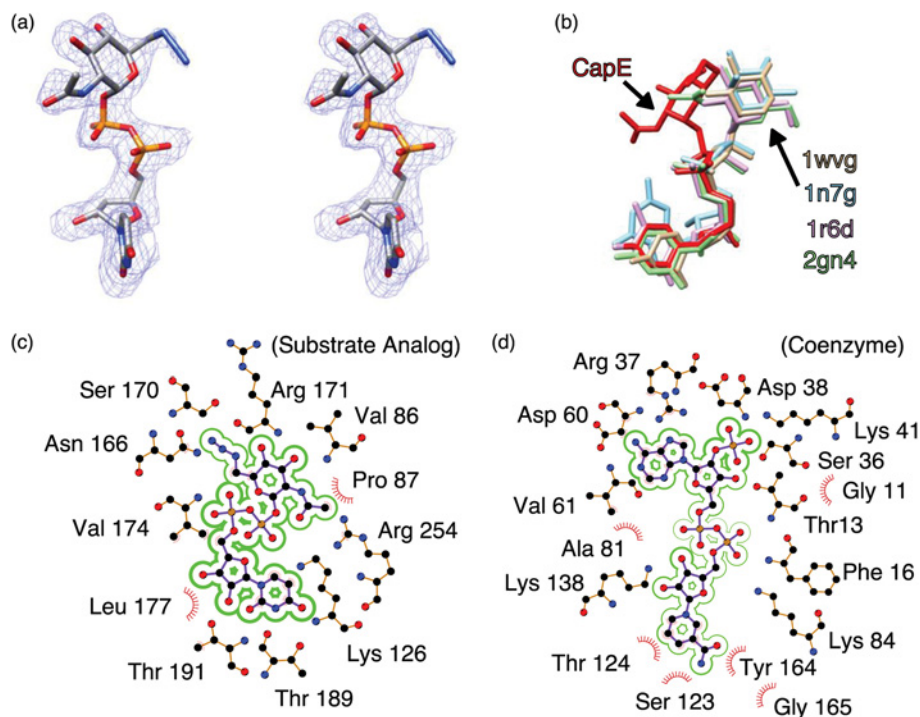
The thermal stability of CapE was determined in a VP-capillary microcalorimeter (GE Healthcare). Proteins at a concentration of 18  $\mu$ M were equilibrated in a solution containing 50 mM Hepes (pH 7.4) and 150 mM NaCl. Thermograms were recorded between 283 K and 373 K at a rate of 1 K min<sup>-1</sup>. The buffer

baseline was subtracted from the protein thermogram and the data subsequently were normalized by protein concentration and adjusted to a two-stage unfolding curve [33] using the program ORIGIN supplied by the manufacturer.

## RESULTS AND DISCUSSION

### Crystal structure of CapE

We determined three crystals structures of CapE by X-ray diffraction methods at resolutions 2.1–2.8 Å (Table 1). Two structures



**Figure 3** Conformation of substrate analogue and coenzyme

(a) Stereoview of the substrate analogue UDP-6N<sub>3</sub>-GlcNAc. The electron density correspond the sigma-A weighted 2Fo-Fc map contoured at 1.0  $\sigma$ . The substrate analogue is depicted with CPK colours. (b) Conformation of various nucleotide-sugars bound to representative SDR enzymes. The panel depicts CapE (this study, red); CDP- $\beta$ -D-xylose bound to GDP-D-glucose 4,6-dehydratase (1wvg, yellow); GDP-rhamnose bound to GDP-mannose 4,6-dehydratase (1n7g, cyan); dTDP-D-glucose bound to dTDP-glyucose 4,6-dehydratase (D128N/E129Q mutain, 1r6d, pink); and UDP-GlcNAc bound to FlaA1 (UDP-GlcNAc inverting 4,6-dehydratase, 2gn4, green). (c) Residue environment around the substrate analogue in wild-type CapE and (d) in the binding pocket of the coenzyme in K126E. Panels (c) and (d) were generated with the program LIGPLOT [51].

correspond to the binary complex of the enzyme with coenzyme (wild-type protein and inactive mutain K126E). The third structure corresponds to wild-type protein (coenzyme-free) in complex with the substrate analogue UDP-6N<sub>3</sub>-GlcNAc.

A three-dimensional homology search in the DALI server [34] indicates that CapE is a member of the extended SDR superfamily [35]. As often seen in this group of enzymes, CapE shares little or very little similarity at the primary sequence level with the other members of this family (10–25% identity). The only exception is FlaA1 of *Helicobacter pylori* (40% identity), which is also the closest structural homolog of CapE in the protein data bank (PDB entry code 2gn4; rmsd =  $2.0 \pm 0.1$  Å). We note that the enzyme following CapE in the biosynthetic route of UDP-D-FucNAc in *S. aureus*, CapF, has the lowest similarity score among the first 950 structural hits found with the program DALI [34].

CapE consists of three well-defined regions: a Rossmann domain for binding the coenzyme NADPH, a substrate binding domain, and a 23-residue loop peripheral to the protein core (Figures 2A and 2B). The first two regions are well-conserved elements of the SDR family of enzymes [35]. Segments corresponding to each domain are interspersed along the primary sequence of the enzyme (Figure 2C). The loop comprising

residues 287–309 constitutes a novel structural element in this family of proteins. We called this loop the *latch* because it engages two molecules of CapE in the crystallographic dimer. The latch is not observed in the two crystal structures of CapE with coenzyme bound because of dynamic disorder. In contrast, the latch is clearly observed in the structure with substrate analogue bound (see below). We note that the tertiary structure of CapE and FlaA1 differs significantly from each other at the latch region (Supplementary Figure S1 at <http://www.biosciencerep.org/bsr/033/bsr033e043add.htm>). Whereas the latch connects two protomers of CapE of the hexameric complex, the equivalent loop of FlaA1 (residues 292–318) folds into its own protein chain.

The substrate-binding domain consists of residues 168–199, 227–260 and 310–337. The substrate-analogue UDP-6N<sub>3</sub>-GlcNAc is clearly visible in the electron density map (Figure 3A). The electron density of the azide moiety is weaker than that in other sections of the substrate analogue, probably because of radiation damage at this functional group. Although the sugar ring of the substrate analogue protrudes into the coenzyme domain (where the active site is located), it occupies a non-catalytic conformation compared with that

of other representative SDR enzymes (Figure 3B). The conformation of UDP-6N<sub>3</sub>-GlcNAc bound to CapE is similar to that of the inhibitor UDP-D-GalNAc bound to the homologous enzyme FlaA1 [18] (Supplementary Figure S1). The substrate analogue engages residues of the binding pocket through multiple polar interactions (Figure 3C and Supplementary Table S1 at <http://www.bioscirep.org/bsr/033/bsr033e043add.htm>). In particular, the UDP moiety of the substrate analogue establishes numerous H-bond and electrostatic interactions with polar residues of the protein.

The coenzyme-binding region is the largest functional element of CapE. This domain adopts a classical Rossmann fold [16,35], comprising residues 1–167, 200–226 and 261–286. Although NADPH was not added during protein purification, or during the crystallization trials, it is clearly visible in the crystal structure of wild-type and K126E, suggesting that the cofactor binds the enzyme tightly (Supplementary Figure S2 at <http://www.bioscirep.org/bsr/033/bsr033e043add.htm>). A detailed view of the local environment around the coenzyme is depicted in Figure 3(D). The coenzyme engages in a dense network of non-covalent interactions with the enzyme, including polar and van der Waals forces, and hydrophobic interactions. Unexpectedly, the coenzyme moiety is not observed in the structure of CapE in complex with substrate analogue. This is surprising because the superposition of the two types of CapE complexes does not suggest direct clashes between the substrate analogue and the coenzyme moiety. It is possible that the departure of the coenzyme is stimulated by a kinetic mechanism similar to that of SDR proteins ArnA and UDP-xylose synthase [36]. Alternatively, it is conceivable that the substrate analogue binds to CapE in a non-native conformation that is incompatible with the binding of the coenzyme.

### CapE is a functional hexamer

Analysis of the crystallographic symmetry reveals that CapE is a homo-hexamer in the crystal form (Figure 4). The protein assembles as a trimer of dimers (3×2). The hexameric organization of CapE constitutes a rare example within the SDR family. This organization is only shared by FlaA1, which is the SDR enzyme with the highest structural homology to CapE [18]. SEC indicates that the hexameric complex is the predominant species in solution (Figure 4B). The hexamer coexists with a small fraction of monomer (~8%), but not dimer. Moreover, CapE undergoes two-state thermal unfolding without intermediates, suggesting that the hexamer is the most stable form of CapE in solution (Supplementary Figure S3 at <http://www.bioscirep.org/bsr/033/bsr033e043add.htm>) [37].

The evaluation of the contact interface with the PISA server [38] indicates that the hexamer of CapE (with substrate analogue bound) buries nearly 24000 Å<sup>2</sup> of surface area upon oligomerization, i.e. 24% of the total solvent-exposed surface. Each protomer of CapE interacts with adjacent molecules through three major interfaces (Figure 4C). The largest interface, between molecules A and B, buries 5016 Å<sup>2</sup> and involves 122 residues (28% of total) (Table 2). This interface corresponds to the typical dimerization

interface observed in numerous SDR enzymes [39,40]. The interfaces A–F and A–C are smaller, burying 2044 Å<sup>2</sup> and 934 Å<sup>2</sup>, respectively.

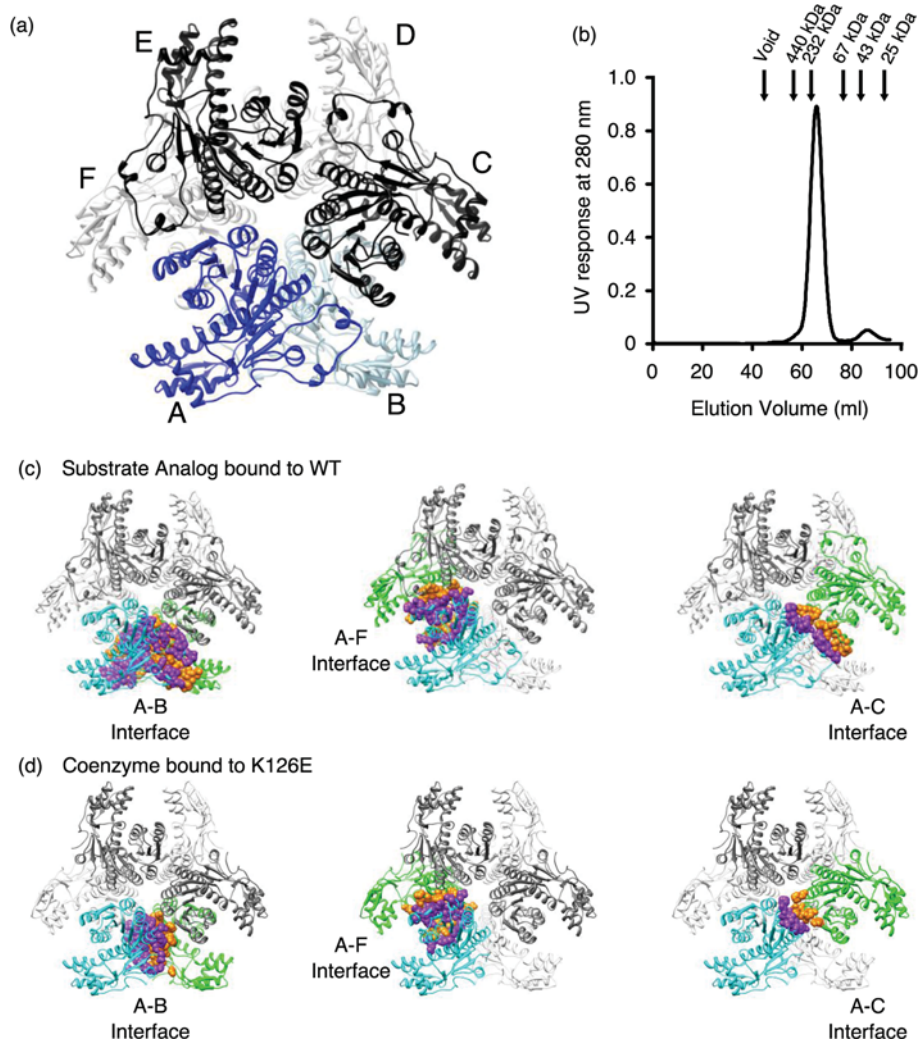
The A–B and A–C surface areas of CapE in complex with coenzyme are significantly smaller than that of CapE with substrate analogue bound. The A–B and A–C surface areas calculated from the coordinates of CapE with coenzyme bound are reduced by 60 and 65%, respectively (Figure 4C, Table 2). The reason is the dynamic disorder occurring at these two contact regions. In fact, the number of residues observed at these two interfaces is greatly diminished (Table 2). On the contrary, the A–F interaction surface remains essentially unchanged regardless of the ligand present in the crystal. The surface area, number of residues and number of non-covalent interactions in the A–F region remains constant among the three crystal structures (Supplementary Table S2 at <http://www.bioscirep.org/bsr/033/bsr033e043add.htm>). hUDGH (human UDP- $\alpha$ -D-glucose dehydrogenase) is a well-studied hexameric enzyme showing large structural changes during its catalytic cycle [41,42]. The hexamer-building interface of hUDGH is quite flexible and therefore the hexamer readily dissociates into dimers and tetramers. On the contrary, the hexamer of CapE seems more stable and remains together even after the disruptive effects described above, or after large conformational changes (see below). We cannot rule out that some of the conformational changes are influenced by crystal packing forces, and therefore may not be of functional significance.

### Conformational changes

The Rossmann domain of CapE is located at the centre of the hexamer and mediates most of the inter-chain contacts, whereas the substrate-binding domain is found at the periphery of the complex (Figure 5). This arrangement could facilitate conformational changes during the catalytic cycle at the substrate-binding domain, since this domain is more exposed to the solvent.

The crystal structures of CapE with analogue bound or with coenzyme bound were compared with the program DYNDOM [43] to quantify their conformational changes. The structure of K126E with coenzyme bound was chosen over the equivalent structure of wild-type CapE because of the better resolution achieved with the mutated protein (Table 1). The analysis with DYNDOM indicates that the substrate-binding domain rotates 11° and moves a few angstroms towards the coenzyme-binding domain upon exit of the coenzyme and binding of substrate analogue (Figure 5B). The largest shift occurs at residues Gly<sup>256</sup> and Gly<sup>252</sup> (6.0 Å). Although these observations must be understood within the context of the crystal structure, they are suggestive of a significant mobility at the substrate domain during catalysis.

Another important dynamic element is the latch (residues 287–309, Figure 6). The latch occupies a position between that of the substrate-binding and Rossmann domains (Figure 5A). The latch is disordered in the two crystal structures with coenzyme bound (Figure 6B and Supplementary Figure S4 at <http://www.bioscirep.org/bsr/033/bsr033e043add.htm>).



**Figure 4** Hexameric organization of CapE

(a) CapE forms a hexamer organized as a trimer of dimers ( $3 \times 2$ ). The structure corresponds to wild-type CapE in complex with substrate analogue (ligand is not shown). The protein subunits of one of the pseudo-dimers are depicted in dark and light blue (chains A and B). The other subunits are depicted in black and grey. (b) SEC profile of CapE. The arrows indicate the elution of the calibration standards. CapE elutes with an apparent molecular weight of 210 kDa. A minor fraction of monomer (8%) with an apparent molecular weight of 40 kDa is also observed. The predicted molecular weight of one subunit of CapE is 39 kDa, and that of the hexamer is 230 kDa. The peak corresponding to the dimer was not observed. (c) Interaction surfaces between protein subunits. The three main contact interfaces correspond to subunits A–B, A–F and A–C. (d) Same interaction surfaces in mutant K126E.

The latch connects two contiguous subunits of CapE, and forms a large fraction of the A–B and A–C interface of the hexamer with very high shape complementarity ( $S_C = 0.79$  [44]) (Figure 6C and Supplementary Figure S5 at <http://www.bioscirep.org/bsr/033/bsr033e043add.htm>). Together with the residues of neighbouring subunits of CapE, the latch contributes  $>3500 \text{ \AA}^2$  of buried surface area, including 28 H-bonds. The interaction between the latch and the adjacent protomer is also held together by contacts involving four large hydrophobic residues: Leu<sup>288</sup>, Tyr<sup>290</sup>, Tyr<sup>293</sup> and Ile<sup>301</sup>. In addition, the latch occupies a strategic position at the entrance of the substrate-binding pocket,

potentially regulating the access of the substrate to the active site (Figure 6C).

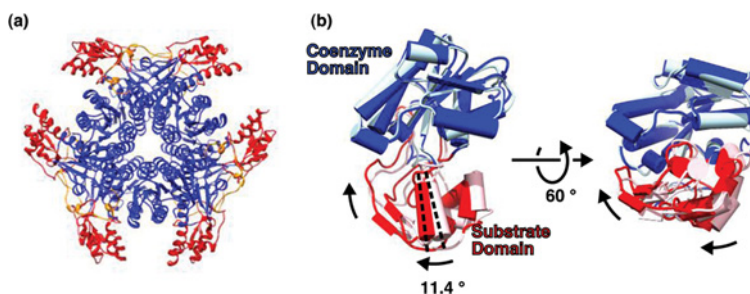
A search of tertiary structure homologues of the latch with the DALI server [34] did not find any example of this structural element in the protein data base. For example, the SDR enzyme FlaA1 (the closest structural homologue of CapE) also forms a hexamer of similar size to that of CapE, but do not exhibit this element (Supplementary Figure S1) [18]. On the contrary, the program Protein BLAST indicates that the primary sequence of the latch is well conserved among a variety of Gram-positive and Gram-negative bacteria of pathogenic potential

**Table 2 Interaction surface area between CapE protomers**

Interface*	Protein	Buried area (Å <sup>2</sup> )	Number of residues†	No. H-Bond
A–B	WT (UDP-6N <sub>3</sub> -GlcNAc)	5016	122	28
	WT (NADPH)	2162	58	4
	K126E (NADPH)	1988	58	4
A–F	WT (UDP-6N <sub>3</sub> -GlcNAc)	2044	58	12
	WT (NADPH)	2066	58	10
	K126E (NADPH)	2070	58	12
A–C	WT (UDP-6N <sub>3</sub> -GlcNAc)	934	32	6
	WT (NADPH)	346	14	0
	K126E (NADPH)	334	14	0
Total	WT (UDP-6N <sub>3</sub> -GlcNAc)	7994	212	46
	WT (NADPH)	4574	130	14
	K126E (NADPH)	4392	130	16

\*Protomers of CapE as shown in Figure 4(A).

†Some residues are counted more than once because they participate in more than one surface.


**Figure 5 Conformational changes in the substrate-binding domain**

(a) Topology of the hexamer. The coenzyme-binding domain (blue) and the substrate-binding domain (red) occupy the central and peripheral regions of the complex, respectively. The latch (orange) is located in an intermediate position. (b) Conformational change. Binding of the substrate analogue, and loss of the coenzyme, induces a rotation of 11° and a displacement of the substrate domain towards the Rossmann domain. Dark and light colours correspond to wild-type protein and K126E, respectively. The rotation angle was calculated with the program DYNDOM [43].

(Figure 6D). The genera of these pathogenic bacteria are *Pseudomonas*, *Acidovorax*, *Hyphomicrobium*, *Vibrio*, *Pasteurella*, *Fusobacterium*, *Listeria*, *Bacillus* and *Enterococcus* (in addition to *Staphylococcus*). Serotypes of a majority of these bacteria are encapsulated, suggesting a direct association between this novel group of SDR enzymes, the production of CP and pathogenesis [6,12,45–49].

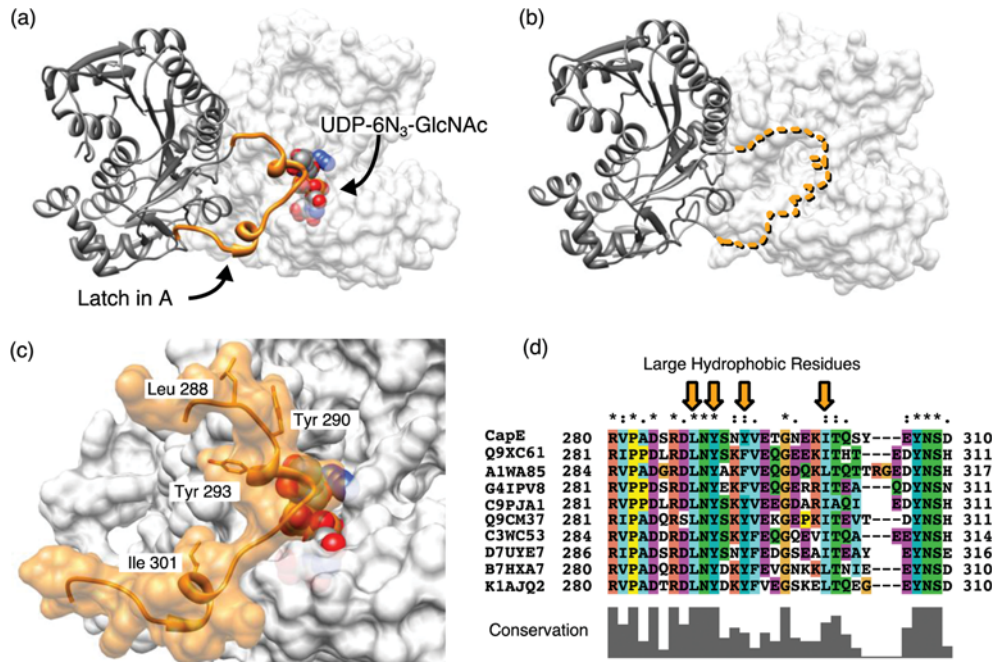
### Site-directed mutagenesis supports the structural findings

The coenzyme-binding domain hosts the canonical catalytic triad, which invariably contains the sequence YxxxK in SDR enzymes (except for a minor group of divergent SDR enzymes, which are characterized by a YxxMxxxK motif). Strikingly, the corresponding sequence of CapE is altered to M<sup>134</sup>xxxK<sup>138</sup> without the catalytic tyrosine. Importantly, the sequence alignment of the family of proteins sharing the latch

(see Figure 6D) demonstrates an absolute conservation of the catalytic residues of CapE (Supplementary Figure S6 at <http://www.bioscirep.org/bsr/033/bsr033e043add.htm>). In contrast, the other SDR structural homologues do invariably possess the canonical YxxxK motif (Supplementary Figure S7 at <http://www.bioscirep.org/bsr/033/bsr033e043add.htm>). We propose that the unique structural element that we called the latch, and the unique composition of the active site residues define a novel subfamily of polysaccharide-synthesizing enzymes within the SDR super-family of proteins.

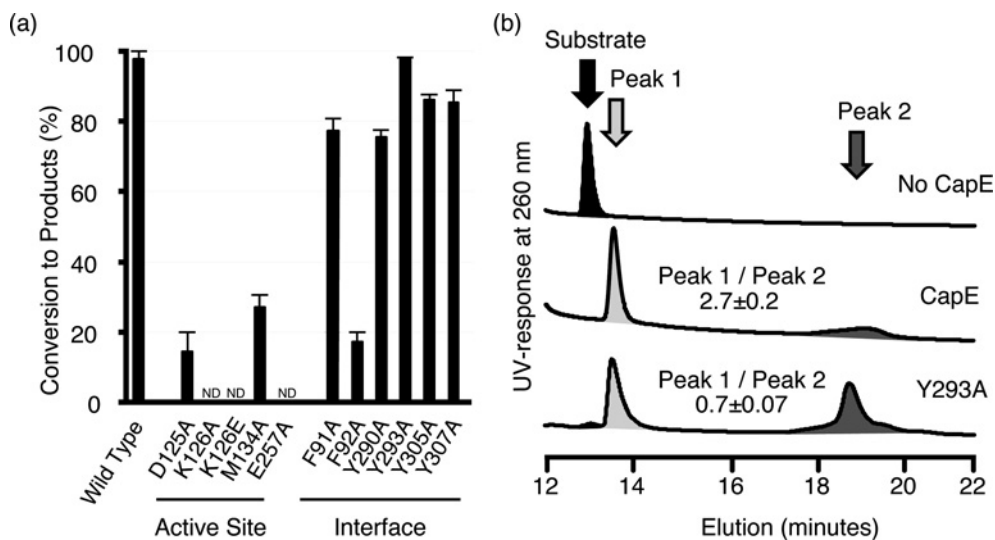
To strengthen the structural analysis we prepared several muteins of CapE by site-directed mutagenesis (Figure 7 and Supplementary Table S3 at <http://www.bioscirep.org/bsr/033/bsr033e043add.htm>). No excess coenzyme was added in the enzymatic assay because CapE was purified in complex with the coenzyme. The level of NADPH bound to CapE was essentially constant among all muteins as estimated from the ratio Abs260/Abs280 (Supplementary Table S3). Firstly,





**Figure 6** Analysis of the latch

(a) Structure of the latch (orange) of subunit A (grey ribbons) associates with the substrate-binding domain of molecule B (grey surface). The latch (orange) of subunit A (grey ribbons) associates with the substrate-binding domain of molecule B (grey surface). The substrate analogue is depicted only in subunit B (CPK representation). (b) Same view of K126E. The latch is not modelled because of disorder. The dotted line represents the hypotheticalal position of the latch in the same position as shown in panel (a). (c) Close view of the latch of subunit A interacting with the substrate-binding pocket of subunit B. The hydrophobic residues Leu<sup>288</sup>, Tyr<sup>290</sup>, Tyr<sup>293</sup> and Ile<sup>301</sup> fit on the grooves of subunit B with high surface complementarity. (d) Sequence alignment of the residues belonging to the latch. The panel shows the top-ten solutions found by BLAST [52]. The sequences belong to various genus of Gram-positive and Gram-negative pathogenic bacteria (from top to bottom: *Staphylococcus*, *Pseudomonas*, *Acidovorax*, *Hyphomicrobium*, *Vibrio*, *Pasteurella*, *Fusobacterium*, *Listeria*, *Bacillus* and *Enterococcus*). Each sequence is identified by their accession code in the UNIPROT database.



**Figure 7** Mutational analysis

(a) Wild-type CapE or mutants (2  $\mu$ M) were incubated with UDP-D-GlcNAc (200  $\mu$ M) for 2 h at 37  $^{\circ}$ C. Consumption of substrate UDP-D-GlcNAc was monitored by HPLC as described in experimental procedures. The bars correspond to the average of three independent assays  $\pm$  S.D. (b) Representative HPLC profiles. The substrate UDP-D-GlcNAc was incubated with no CapE (top), with wild-type CapE (middle) and with mutetin Y293A (bottom).

individual residues of the active site and substrate-binding pocket were mutated: Asp<sup>125</sup>, Lys<sup>126</sup>, Met<sup>134</sup> and Glu<sup>257</sup> were changed to Ala (Lys<sup>126</sup> was also mutated to Glu and its crystal structure discussed above). All these individual muteins exhibited a much-diminished activity compared with the wild-type enzyme. The muteins K126A, K126E and E257A were completely inactive, demonstrating the importance of these residues for the functional binding of the substrate. Meanwhile the activity of D125A and M134A was strongly impaired (4–6-fold lower conversion than WT protein). We note that all the muteins eluted as hexamers during their final step of purification by SEC.

Secondly, we evaluated the properties of the A–B surface (essentially the latch) by mutating separately five different residues to Ala: Phe<sup>91</sup>, Tyr<sup>290</sup>, Tyr<sup>293</sup>, Tyr<sup>305</sup> and Tyr<sup>307</sup>. The last four muteins belong to the latch region. The mutations did not alter significantly the levels of conversion compared with the wild-type enzyme (Figure 7A). However, although the values of conversion of Y293A and wild-type protein are identical, the ratio of their enzymatic products (computed from the area under the HPLC peaks) is inverted in Y293A in comparison with wild-type CapE. The relative areas of peak-1 with respect to peak-2 in wild-type or in mutein Y293A were 2.7 and 0.7, respectively (Figure 7B). Similarly, the selectivity is inverted in all muteins belonging to the latch, but not in F91A, a residue not belonging to the latch (Supplementary Table S3). Changes in the relative abundance of intermediates do not necessarily reflect a change of their thermodynamic equilibrium, but probably a slow down of their rate of interconversion. Our next challenge will consist in the identification of these unstable intermediates, which will allow an in-depth discussion of the structure/function relationship and catalytic mechanism of CapE. We note that in contrast to the results obtained with the muteins belonging to the interface A–B, the mutein F92A of the A–F interface led to a much-reduced activity (<20% conversion).

## Conclusion

We determined the first set of crystal structures of the enzyme CapE of *S. aureus* belonging to the biosynthetic pathway of CP. CapE forms a robust hexamer held together by three different interfaces, A–B, A–F and A–C. The largest A–B interface contains a mobile motif that we denominated the latch. The latch is a flexible element that increases the contact surface area of adjacent monomers of CapE in the presence of the substrate analogue UDP-6N<sub>3</sub>-GlcNAc. The structural analysis also demonstrates a rotation of the substrate-binding domain when the substrate analogue binds to the enzyme. Importantly a sequence alignment of the latch, and of the active site residues of CapE, has identified a novel subfamily of SDR enzymes involved in the synthesis of polysaccharide of Gram-positive and Gram-negative pathogenic bacteria.

## AUTHOR CONTRIBUTION

Takamitsu Miyafusa designed research, performed experiments, analysed data, and wrote the manuscript. Jose Caaveiro designed

research, performed experiments, analysed data and wrote the manuscript. Yoshikazu Tanaka designed research, performed experiments and analysed data. Martin Tanner performed experiments, analysed data and edited the manuscript. Kouhei Tsumoto designed the overall study, analysed data, provided guidance and edited the manuscript.

## ACKNOWLEDGEMENTS

We thank members of Photon Factory in Tsukuba (Japan) for their assistance during X-ray data collection. We thank Dr M. Kuroda (Tsukuba University) and Dr T. Ohta (Tsukuba University) for providing plasmids containing full-length CapF and CapE. We also appreciate useful suggestions from Dr M. Yao (Hokkaido University). Access to beamlines AR-NE3A and BL5A was granted by the Photon Factory Advisory Committee (Proposal Numbers 2009G204 and 2011G574).

## FUNDING

This work was supported in part by a Grant-in-Aid for General Research from the Japan Society for the Promotion of Science (to K.T.).

## REFERENCES

- 1 Lowy, F. D. (1998) Medical progress: *Staphylococcus aureus* infections. *N. Engl. J. Med.* **339**, 520–532
- 2 Lowy, F. D. (2011) How *Staphylococcus aureus* adapts to its host. *N. Engl. J. Med.* **364**, 1987–1990
- 3 Smith, T. L., Pearson, M. L., Wilcox, K. R., Cruz, C., Lancaster, M. V., Robinson-Dunn, B., Tenover, F. C., Zervos, M. J., Band, J. D. et al. (1999) Emergence of vancomycin resistance in *Staphylococcus aureus*. *N. Engl. J. Med.* **340**, 493–501
- 4 Rybak, M. J. and Akins, R. L. (2001) Emergence of methicillin-resistant *Staphylococcus aureus* with intermediate glycopeptide resistance – clinical significance and treatment options. *Drugs* **61**, 1–7
- 5 Payne, D. J., Gwynn, M. N., Holmes, D. J. and Pompliano, D. L. (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discovery* **6**, 29–40
- 6 O’Riordan, K. and Lee, J. C. (2004) *Staphylococcus aureus* capsular polysaccharides. *Clin. Microbiol. Rev.* **17**, 218–234
- 7 Watts, A., Ke, D. B., Wang, Q., Pillay, A., Nicholson-Weller, A. and Lee, J. C. (2005) *Staphylococcus aureus* strains that express serotype 5 or serotype 8 capsular polysaccharides differ in virulence. *Infect. Immun.* **73**, 3502–3511
- 8 Luong, T. T. and Lee, C. Y. (2002) Overproduction of type 8 capsular polysaccharide augments *Staphylococcus aureus* virulence. *Infect. Immun.* **70**, 3389–3395
- 9 Luong, T. T. and Lee, C. Y. (2006) The *arl* locus positively regulates *Staphylococcus aureus* type 5 capsule via an *mgrA*-dependent pathway. *Microbiol.-Sgm.* **152**, 3123–3131
- 10 Jones, C., Currie, F. and Forster, M. J. (1991) N.m.r. and conformational analysis of the capsular polysaccharide from *Streptococcus pneumoniae* type 4. *Carbohydr. Res.* **221**, 95–121
- 11 Manca, M. C., Weintraub, A. and Widmalm, G. (1996) Structural studies of the *Escherichia coli* O26 O-antigen polysaccharide. *Carbohydr. Res.* **281**, 155–160

- 12 Kneidinger, B., O'Riordan, K., Li, J., Brisson, J. R., Lee, J. C. and Lam, J. S. (2003) Three highly conserved proteins catalyze the conversion of UDP-N-acetyl-D-glucosamine to precursors for the biosynthesis of O antigen in *Pseudomonas aeruginosa* O11 and capsule in *Staphylococcus aureus* type 5. Implications for the UDP-N-acetyl-L-fucosamine biosynthetic pathway. *J. Biol. Chem.* **278**, 3615–3627
- 13 Manca, M. C., Weintraub, A. and Widmalm, G. (1996) Structural studies of the *Escherichia coli* O26 O-antigen polysaccharide. *Carbohydr. Res.* **281**, 155–160
- 14 Mulrooney, E. F., Poon, K. K. H., McNally, D. J., Brisson, J. R. and Lam, J. S. (2005) Biosynthesis of UDP-N-acetyl-L-fucosamine, a precursor to the biosynthesis of lipopolysaccharide in *Pseudomonas aeruginosa* serotype O11. *J. Biol. Chem.* **280**, 19535–19542
- 15 Miyafusa, T., Caaveiro, J. M. M., Tanaka, Y. and Tsumoto, K. (2012) Crystal structure of the enzyme CapF of *Staphylococcus aureus* reveals a unique architecture composed of two functional domains. *Biochem. J.* **443**, 671–680
- 16 Jornvall, H., Hoog, J. O. and Persson, B. (1999) SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Lett.* **445**, 261–264
- 17 Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R. and Goodenough, P. W. (2001) Evolution of functional diversity in the cupin superfamily. *Trends Biochem. Sci.* **26**, 740–746
- 18 Ishiyama, N., Creuzenet, C., Miller, W. L., Demendi, M., Anderson, E. M., Harauz, G., Lam, J. S. and Berghuis, A. M. (2006) Structural studies of FlaA1 from *Helicobacter pylori* reveal the mechanism for inverting 4,6-dehydratase activity. *J. Biol. Chem.* **281**, 24489–24495
- 19 Creuzenet, C., Urbanic, R. V. and Lam, J. S. (2002) Structure-function studies of two novel UDP-GlcNAc C6 dehydratases/C4 reductases. Variation from the SYK dogma. *J. Biol. Chem.* **277**, 26769–26778
- 20 Mayer, A., Gloster, T. M., Chou, W. K., Vocadlo, D. J. and Tanner, M. E. (2011) 6'-Azido-6''-deoxy-UDP-N-acetylglucosamine as a glycosyltransferase substrate. *Bioorgan. Med. Chem. Lett.* **21**, 1199–1201
- 21 Miyafusa, T., Tanaka, Y., Kuroda, M., Ohta, T. and Tsumoto, K. (2008) Expression, purification, crystallization and preliminary diffraction analysis of CapF, a capsular polysaccharide-synthesis enzyme from *Staphylococcus aureus*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **64**, 512–515
- 22 Stewart, P. D. S., Kolek, S. A., Briggs, R. A., Chayen, N. E. and Baldock, P. F. M. (2011) Random microseeding: a theoretical and practical exploration of seed stability and seeding techniques for successful protein crystallization. *Cryst. Growth. Des.* **11**, 3432–3441
- 23 Leslie, A. G. (2006) The integration of macromolecular diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 48–57
- 24 Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82
- 25 Storoni, L. C., McCoy, A. J. and Read, R. J. (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 432–438
- 26 McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. and Read, R. J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 458–464
- 27 Adams, P. D., Gopal, K., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Pai, R. K., Read, R. J., Romo, T. D. et al. (2004) Recent developments in the PHENIX software for automated crystallographic structure determination. *J. Synchrotron Radiat.* **11**, 53–55
- 28 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
- 29 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
- 30 Winn, M. D., Isupov, M. N. and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 122–133
- 31 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) Procheck – a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291
- 32 Oka, T., Nemoto, T. and Jigami, Y. (2007) Functional analysis of arabidopsis thaliana RHM2/MUM4, a multidomain protein involved in UDP-D-glucose to UDP-L-rhamnose conversion. *J. Biol. Chem.* **282**, 5389–5403
- 33 Ladbury, J. E. and Doyle, M. L. (2004) Biocalorimetry 2. Applications of Calorimetry in the Biological Sciences. John Wiley & Sons, West Sussex
- 34 Holm, L. and Rosenstrom, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–549
- 35 Kavanagh, K., Jornvall, H., Persson, B. and Oppermann, U. (2008) The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell. Mol. Life Sci.* **65**, 3895–3906
- 36 Polizzi, S. J., Walsh, R. M., Jr., Peebles, W. B., Lim, J. M., Wells, L. and Wood, Z. A. (2012) Human UDP-alpha-d-xylose synthase and *Escherichia coli* ArnA conserve a conformational shunt that controls whether Xylose or 4-Keto-Xylose is produced. *Biochemistry* **51**, 8844–8855
- 37 Huus, K., Havelund, S., Olsen, H. B., van de Weert, M. and Frokjaer, S. (2005) Thermal dissociation and unfolding of insulin. *Biochemistry* **44**, 11171–11177
- 38 Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797
- 39 Major, L. L., Wolucka, B. A. and Naismith, J. H. (2005) Structure and function of GDP-mannose-3',5'-epimerase: an enzyme which performs three chemical reactions at the same active site. *J. Am. Chem. Soc.* **127**, 18309–18320
- 40 Allard, S. T. M., Beis, K., Giraud, M. F., Hegeman, A. D., Gross, J. W., Wilmouth, R. C., Whitfield, C., Graninger, M., Messner, P., Allen, A. G. et al. (2002) Toward a structural understanding of the dehydratase mechanism. *Structure* **10**, 81–92
- 41 Sennett, N. C., Kadirvelraj, R. and Wood, Z. A. (2011) Conformational flexibility in the allosteric regulation of human UDP-alpha-D-Glucose 6-dehydrogenase. *Biochemistry* **50**, 9651–9663
- 42 Kadirvelraj, R., Sennett, N. C., Polizzi, S. J., Weitzel, S. and Wood, Z. A. (2011) Role of packing defects in the evolution of allostery and induced fit in human UDP-glucose dehydrogenase. *Biochemistry* **50**, 5780–5789
- 43 Hayward, S. and Berendsen, H. J. (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30**, 144–154
- 44 Lawrence, M. C. and Colman, P. M. (1993) Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946–950
- 45 Sue, D., Hoffmaster, A. R., Popovic, T. and Wilkins, P. P. (2006) Capsule production in *Bacillus cereus* strains associated with severe pneumonia. *J. Clin. Microbiol.* **44**, 3426–3428
- 46 Bolstad, A. I., Jensen, H. B. and Bakken, V. (1996) Taxonomy, biology, and periodontal aspects of *Fusobacterium nucleatum*. *Clin. Microbiol. Rev.* **9**, 55–71



- 47 Thurlow, L. R., Thomas, V. C. and Hancock, L. E. (2009) Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. *J. Bacteriol.* **191**, 6203–6210
- 48 Gulig, P. A., Bourdage, K. L. and Starks, A. M. (2005) Molecular pathogenesis of vibrio vulnificus. *J. Microbiol.* **43**, Spec No., 118–131
- 49 Arumugam, N. D., Ajam, N., Blackall, P. J., Asiah, N. M., Ramlan, M., Maria, J., Yuslan, S. and Thong, K. L. (2011) Capsular serotyping of *Pasteurella multocida* from various animal hosts – a comparison of phenotypic and genotypic methods. *Trop. Biomed.* **28**, 55–63
- 50 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612
- 51 Wallace, A. C., Laskowski, R. A. and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.* **8**, 127–134
- 52 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402

---

**Received 31 January 2013/9 April 2013; accepted 23 April 2013**

---

**Published as Immediate Publication 24 April 2013, doi 10.1042/BSR20130017**

---



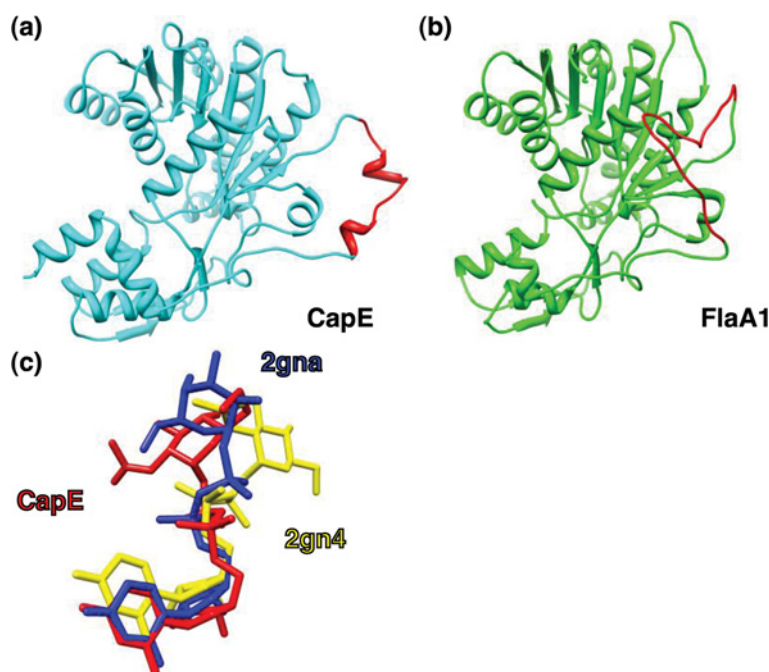
## OPEN ACCESS

## SUPPLEMENTARY DATA

# Crystal structure of the capsular polysaccharide synthesizing protein CapE of *Staphylococcus aureus*

Takamitsu MIYAFUSA\*†, Jose M. M. CAAVEIRO\*†, Yoshikazu TANAKA†<sup>1</sup>, Martin E. TANNER‡ and Kouhei TSUMOTO\*†§<sup>2</sup>

\*Medical Proteomics Laboratory, Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo 108-8639, Japan, †Department of Medical Genome Sciences, School of Frontier Sciences, The University of Tokyo, Minato-ku, Tokyo 108-8639, Japan, ‡Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, British Columbia V6T 1Z1, Canada, and §Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Tokyo 113-8656, Japan



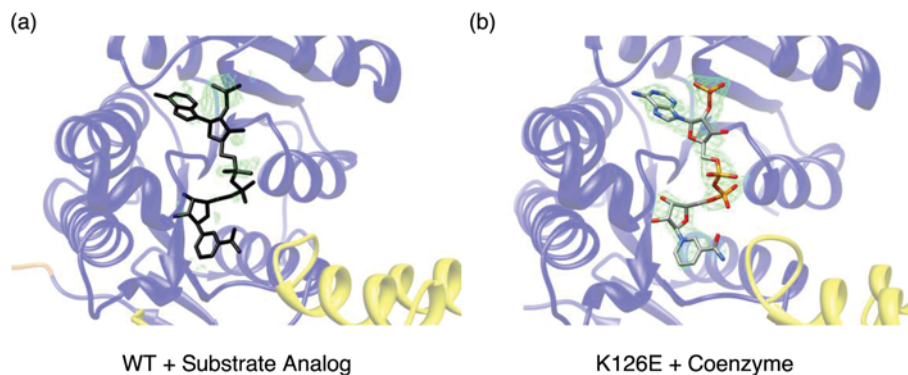
**Figure S1** The structure of the protomer of CapE

Panel (a) is compared with that of FlaA1 (b) using the software PDBe fold. The conformation and orientation of the latch region (in this figure from Leu<sup>288</sup> to Thr<sup>302</sup>) clearly differs from the equivalent region in FlaA1 (from Gln<sup>292</sup> to Ala<sup>310</sup>). (c) Structure of nucleotide-sugars bound to CapE and FlaA1. The panel depicts the substrate analogue bound to CapE (red), the substrate analogue UDP-D-GalNAc bound to FlaA1 (PDB entry code 2gna, blue) and the substrate UDP-D-GlcNAc bound to FlaA1 (PDB entry code 2gn4, yellow).

<sup>1</sup> Present address: Creative Research Initiative Sousei, Hokkaido University, Sapporo 001-0021, Japan

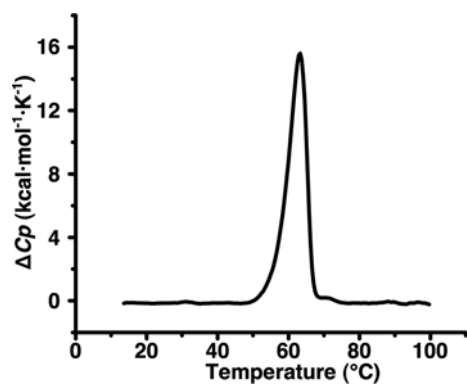
<sup>2</sup> To whom correspondence should be addressed (email tsumoto@ims.u-tokyo.ac.jp).

Atomic co-ordinates and structure factors of holo-CapE, holo-K126E and CapE with substrate analogue UDP-6N3-GlcNAc are deposited in the PDB under entry codes 3VVB, 3WVC, and 3W1V respectively.



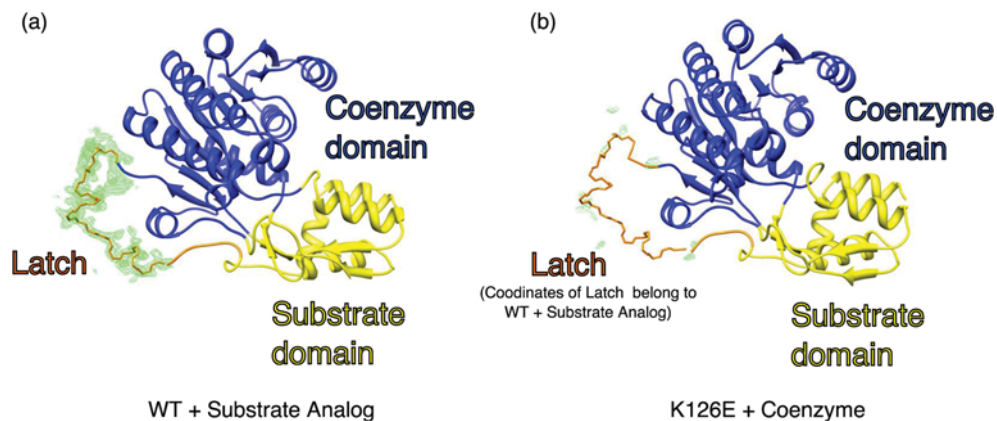
**Figure S2 Coenzyme is absent in the binary complex with substrate analogue**

Electron density map in the cofactor binding site of (a) the binary complex of wild-type CapE with UDP-6N<sub>3</sub>-GlcNAc bound and (b) mutein K126E. The green mesh corresponds to the sigma-A weighted 2Fo-Fc electron density map contoured at a level of 1.0 $\sigma$ . In panel (a), the black sticks correspond to the superimposed coordinates of NADPH from the structure in panel (b). The electron density maps cover a radius of 2.0 Å from any atom belonging to the coenzyme.



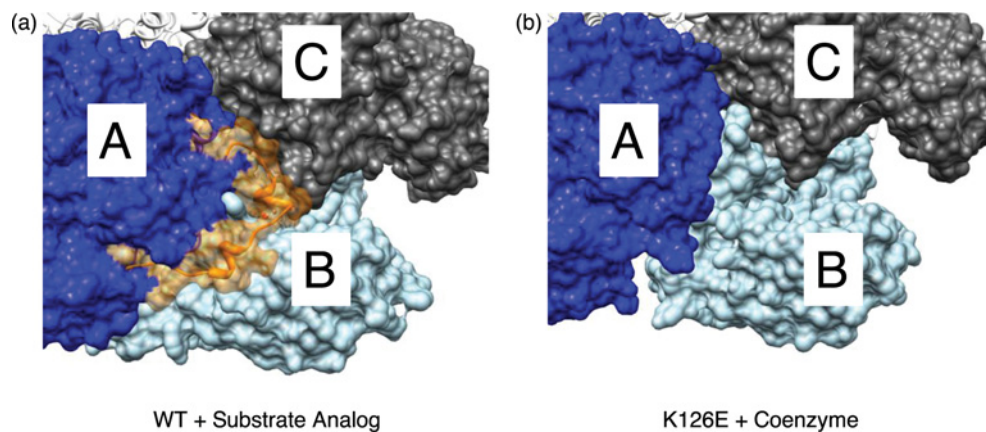
**Figure S3 Thermal unfolding of CapE monitored by DSC**

CapE (18  $\mu$ M) was heated from 283 to 373 K at a rate of 1 K min<sup>-1</sup>. The thermodynamic parameters obtained from fitting the data to a two-state thermal transition are  $T_M = 62.5 \pm 0.1$  °C and  $\Delta H^\circ = 105 \pm 1.7$  kcal mol<sup>-1</sup>.



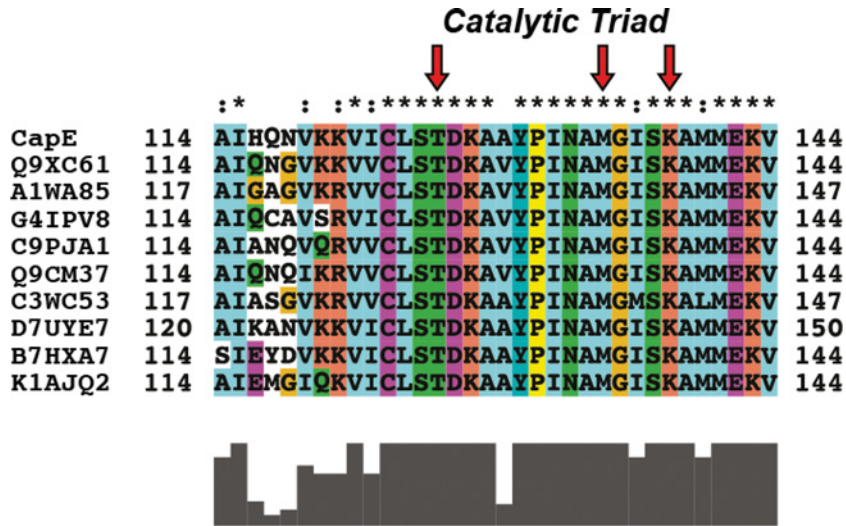
**Figure S4** The latch is disordered in the apo form of CapE

Comparison of the electron density maps in the region of the latch in (a) CapE with substrate-analogue bound and (b) K126E with coenzyme bound. The green mesh corresponds to the sigma-A weighted 2Fo-Fc electron density map contoured at a level of  $1.5 \sigma$ . For comparative purposes, the structure of the latch shown in panel (b) corresponds to that in CapE with substrate analogue. The electron density maps cover a radius of 2.0 Å from any atom belonging to the latch.



**Figure S5** The latch contacts multiple subunits of CapE

Surface-representation of the latch region in the structure of (a) CapE in complex with UDP-6N<sub>3</sub>-GlcNAc and (b) K126E in complex with coenzyme. The latch of subunit A is depicted as a semi-transparent orange surface. The latch does not appear in panel (b) because of dynamic disorder.



**Figure S6** Sequence alignment of the region of CapE containing the catalytic triad  
 The sequences correspond to the same set of proteins aligned in Figure 6 of the manuscript.





Figure S7 Sequence alignment of the top 55 structural homologues found in the DALI server (only published structures were considered)

The position of the catalytic triad in the SDR family of proteins is highlighted. Only CapE displays a Met at the catalytic site.

**Table S1 Polar interactions between UDP-N<sub>3</sub>-GlcNAc and CapE**

<b>Number</b>	<b>Substrate</b>	<b>Residue [atom]</b>	<b>Distance (Å)</b>
1	[O2']	Glu <sup>257</sup> [OE1]	2.71
2	[O2']	Glu <sup>257</sup> [OE2]	3.45
3	[O3']	Val <sup>86</sup> [N]	2.78
4	[O4']	Lys <sup>84</sup> [O]	2.80
5	[N3]	Thr <sup>189</sup> [O]	2.57
6	[O2]	Thr <sup>191</sup> [N]	2.99
7	[O4]	Thr <sup>189</sup> [OG1]	3.30
8	[O2B]	Arg <sup>254</sup> [NH2]	2.79
9	[O1A]	Val <sup>174</sup> [N]	2.84
10	[O2]	Thr <sup>191</sup> [OG1]	3.17
11	[O2']	Thr <sup>191</sup> [OG1]	2.69
12	[O1B]	Lys <sup>126</sup> [NZ]	2.86
13	[O2']	Thr <sup>191</sup> [OG1]	2.69
14	[O3']	Gln <sup>85</sup> [OE1]	3.00

**Table S2 Intermolecular polar contacts between subunits of CapE (Å)**

Since CapE is a homomer, and to avoid duplications, only one set of interactions is given.

(a)

Number	Chain A	Chain B	Substrate analogue wild-type	NADPH wild-type	NADPH K126E
Hydrogen bonds					
1	Lys <sup>143</sup> [NZ]	Glu <sup>90</sup> [OE1]	2.56	2.69	2.30
2	Lys <sup>143</sup> [NZ]	Ile <sup>131</sup> [O]	3.13	–	–
3	Arg <sup>150</sup> [NH1]	Glu <sup>90</sup> [O]	3.30	3.22	3.64
4	Thr <sup>253</sup> [N]	Thr <sup>302</sup> [OG1]*	3.14	–	–
5	Thr <sup>253</sup> [N]	Asn <sup>298</sup> [O]*	3.26	–	–
6	Thr <sup>253</sup> [OG1]	Thr <sup>302</sup> [OG1]*	2.69	–	–
7	Arg <sup>254</sup> [NH1]	Val <sup>294</sup> [O]*	2.86	–	–
8	His <sup>255</sup> [N]	Asn <sup>292</sup> [O]*	2.85	–	–
9	His <sup>255</sup> [ND1]	Asn <sup>292</sup> [O]*	3.28	–	–
10	His <sup>255</sup> [NE2]	Gln <sup>271</sup> [OE1]*	3.29	–	–
11	Gly <sup>256</sup> [N]	Glu <sup>267</sup> [OE2]	2.69	–	–
12	Arg <sup>266</sup> [NH2]	Glu <sup>257</sup> [O]	3.19	–	–
13	Asn <sup>298</sup> [N]†	Ile <sup>251</sup> [O]	3.65	–	–
14	Asn <sup>298</sup> [N]†	Thr <sup>253</sup> [O]	2.67	–	–
Salt bridges					
1	Lys <sup>143</sup> [NZ]	Glu <sup>90</sup> [OE1]	2.64	2.69	2.30

(b)

Number	Chain A	Chain F	Substrate analogue wild-type	NADPH wild-type	NADPH K126E
Hydrogen bonds					
1	Arg <sup>37</sup> [NH1]	Arg <sup>37</sup> [O]	–	3.05	3.38
2	Lys <sup>84</sup> [N]	Glu <sup>39</sup> [OE1]	2.81	2.81	2.90
3	Gln <sup>85</sup> [N]	Glu <sup>39</sup> [OE2]	2.68	2.75	2.95
4	Ser <sup>88</sup> [OG]	Glu <sup>39</sup> [OE2]	3.02	2.92	2.88
5	Lys <sup>98</sup> [NZ]	Asp <sup>63</sup> [OD2]	2.63	2.52	2.68
6	Glu <sup>95</sup> [OE2]	Ser <sup>66</sup> [OG]	2.52	–	–
7	Arg <sup>37</sup> [NH2]	Thr <sup>99</sup> [OG1]	3.79	–	–
8	Arg <sup>62</sup> [NH2]	Asp <sup>60</sup> [OD2]	–	–	3.46
Salt bridges					
1	Arg <sup>37</sup> [NH1]	Asp <sup>60</sup> [OD2]	3.11	–	–
2	Arg <sup>37</sup> [NH2]	Asp <sup>60</sup> [OD1]	2.81	–	–
3	Lys <sup>98</sup> [NZ]	Asp <sup>63</sup> [OD2]	2.63	2.52	2.68
4	Arg <sup>62</sup> [NH2]	Asp <sup>60</sup> [OD2]	–	–	3.46

\*Residues belonging to the latch are disordered in the two structures of CapE with NADPH-bound.

**Table S3 Activity of wild-type CapE and mutants**

N.D., no product was detected by HPLC; N.A., not applicable because of low conversion.

<b>Protein</b>	<b>Position of mutation</b>	<b>Absorbance ratio 280 nm/260 nm</b>	<b>Conversion</b>	<b>Product ratio</b>
Wild-type		0.95	97.6 ± 2.3	2.7 ± 0.2
D125A	Active site	1.12	14.5 ± 5.6	N.A.
K126A	Active site	1.00	N.D.	N.A.
K126E	Active site	1.05	N.D.	N.A.
M134A	Active site	1.11	27.0 ± 3.7	N.A.
E257A	Active site	0.96	N.D.	N.A.
F91A	A–B surface	1.00	77.3 ± 3.4	1.5 ± 0.1
F92A	A–F surface	1.16	17.3 ± 2.8	N.A.
Y290A	Latch	1.00	75.5 ± 2.1	0.4 ± 0.05
Y293A	Latch	1.09	97.6 ± 0.6	0.7 ± 0.07
Y305A	Latch	1.10	86.0 ± 1.6	0.3 ± 0.02
Y307A	Latch	0.98	85.4 ± 3.4	0.8 ± 0.1

**Received 31 January 2013/9 April 2013; accepted 23 April 2013****Published as Immediate Publication 24 April 2013, doi 10.1042/BSR20130017**