the FEBS Journal

Structural and functional studies on a mesophilic stationary phase survival protein (Sur E) from *Salmonella typhimurium*

A. Pappachan[1], H. S. Savithri[2] and M. R. N. Murthy[1]

1 Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India
2 Department of Biochemistry, Indian Institute of Science, Bangalore, India

Keywords
divalent metal ion; domain swapping; mononucleotidase; stationary phase; Sur E

Correspondence
M. R. N. Murthy, Molecular Biophysics Unit, Indian Institute of Science, Bangalore- 560 012, India
Fax: +91 80 23600535
Tel: +91 80 22932458
E-mail: mrn@mbu.iisc.ernet.in

Database
The coordinates and structure factors of the crystal structures described in this study have been submitted to the Protein Data Bank, and the structures have been assigned the accession codes 2v4n and 2v4o for the F222 SurE structure and the C2 SurE structure, respectively

(Received 13 August 2008, revised 24 September 2008, accepted 26 September 2008)

doi:10.1111/j.1742-4658.2008.06715.x

SurE, the stationary-phase survival protein of *Salmonella typhimurium*, forms part of a stress survival operon regulated by the stationary-phase RNA polymerase alternative sigma factor. SurE is known to improve bacterial viability during stress conditions. It functions as a phosphatase specific to nucleoside monophosphates. In the present study we reported the X-ray crystal structure of SurE from *Salmonella typhimurium*. The protein crystallized in two forms: orthorhombic F222; and monoclinic C2. The two structures were determined to resolutions of 1.7 and 2.7 Å, respectively. The protein exists as a domain-swapped dimer. The residue D230 is involved in several interactions that are probably crucial for domain swapping. A divalent metal ion is found at the active site of the enzyme, which is consistent with the divalent metal ion-dependent activity of the enzyme. Interactions of the conserved DD motif present at the N-terminus with the phosphate and the Mg^{2+} present in the active site suggest that these residues play an important role in enzyme activity. The divalent metal ion specificity and the kinetic constants of SurE were determined using the generic phosphatase substrate para-nitrophenyl phosphate. The enzyme was inactive in the absence of divalent cations and was most active in the presence of Mg^{2+}. Thermal denaturation studies showed that *S. typhimurium* SurE is much less stable than its homologues and an attempt was made to understand the molecular basis of the lower thermal stability based on solvation free-energy. This is the first detailed crystal structure analysis of SurE from a mesophilic organism.

During stress and the stationary phase of growth, bacterial cells undergo a variety of morphological and physiological changes. To tide over these unfavorable conditions, several genes are induced. The rpoS-encoded stationary-phase RNA polymerase alternative sigma factor σS (RpoS) plays a major role as a regulator of genes involved in the response to stress. In *Escherichia coli*, rpoS clusters with three other genes: *pcm*, *surE* and *nlpD*. *nlpD* codes for a lipoprotein, *pcm* codes for an L-isoaspartate *O*-methyltransferase and *surE* codes for a stationary-phase survival protein. The *surE* gene was first discovered in *E. coli* by Clarke and co-workers [1]. *E. coli* strains with a mutant *surE* gene survived poorly in the stationary

Abbreviations
Aa SurE, *Aquifex aeolicus* SurE; C2-SurE, monoclinic SurE; Ec SurE, *Escherichia coli* SurE; F222-SurE, orthorhombic SurE; IPTG, isopropyl thio-β-D-galactoside; Pa SurE, *Pyrobaculum aerophilum* SurE; pNPP, para-nitrophenyl phosphate; SFE, solvation free-energy; St SurE, *Salmonella typhimurium* SurE; Tm SurE, *Thermotoga maritima* SurE; Tt SurE, *Thermus thermophilus* SurE.

phase and under conditions of high temperature and high salt compared with parent strains that had the intact *surE* gene [1]. The *surE* gene duplicated in a strain of *E. coli* subjected to 2000 generations of high-temperature growth, which again emphasizes its role in the stress response [2]. *surE* is an ancient and well-conserved gene distributed across various kingdoms.

The exact biochemical role of SurE is still not very clear. PHO2 protein (from the yeast *Yarrowia lipolytica*) that had a low level of sequence similarity with the N-terminal domain of SurE proteins complemented mutations in two acid phosphatases of *Saccharomyces cerevisiae*, which led to the proposal that SurE might be an acid phosphatase [3]. Later studies on SurEs clearly demonstrated that in contrast to nonspecific acid phosphatases, SurE proteins showed specificity towards nucleoside monophosphates and hence it was proposed that SurE could be designated a nucleotidase. These studies also revealed that SurE shows a divalent metal ion-dependent nucleotidase activity and can dephosphorylate various ribose and deoxyribonucleoside monophosphates with highest affinity towards 3'-AMP. *E. coli* SurE (Ec SurE) also showed exopolyphosphatase activity with preference for short chain-length substrates (P_{20-25}) [4]. It has been shown that tyrosine phosphorylation regulates guanosine-5′-*O*-γ thiotriphosphate-stimulated L-isoaspartyl methyl transferase in rat kidney cytosol [5]. Studies have shown that there is a functional relationship between SurE and Pcm. L-isoaspartate *O*-methyltransferase converts isoapartyl residues to L-aspartyl residues and thereby repairs damaged proteins that can accumulate with time in senescent cells. A copy of either *pcm* or *surE* appears to be sufficient to avoid isoaspartyl damage during stress [6]. This observation suggested that the two genes might represent parallel pathways by which *E. coli* responds to protein damage.

X-ray crystal structures of SurE from three thermophilic organisms – *Thermotoga maritima* (Tm SurE), *Thermus thermophilus* (Tt SurE), *Aquifex aeolicus* (Aa SurE) and an archaic organism, *Pyrobaculum aerophilum* (Pa SurE) have been determined. The structure of Aa SurE has been deposited in the Protein Data Bank, but has not been described in the literature. Here we report a detailed analysis of the crystal structure of SurE from a mesophilic organism – *Salmonella typhimurium* (St SurE). Results of activity studies with the substrate para-nitrophenyl phosphate (pNPP), are also presented. Thermal denaturation studies have shown that St SurE is much less stable than its homologues and an attempt was made to understand the molecular basis of the lower thermal stability.

Results and Discussion

Preliminary characterization and crystallization of the protein

The *surE* gene from *S. typhimurium* was cloned in an isopropyl thio-β-D-galactoside (IPTG)-inducible vector, overexpressed in *E. coli* and purified to homogeneity using Ni-nitrilotriacetic acid affinity chromatography. The purified protein showed a single polypeptide band in SDS/PAGE corresponding to a molecular mass of 28 kDa, which agreed with the theoretically calculated molecular mass from the sequence including the additional amino acids resulting from the cloning strategy. The molecular mass of the protein was also confirmed by MALDI-TOF MS. CD spectra indicated a well-folded protein, and dynamic light-scattering measurements showed a monodisperse distribution. Hence the enzyme was considered suitable for crystallization.

The protein crystallized under oil in two forms. Orthorhombic crystals (F222-SurE) were obtained in the presence of 0.2 M trisodium citrate dihydrate, 0.1 M Hepes (pH 7.5) and 30% 2-methyl-2,4-pentanediol, and monoclinic crystals (C2-SurE) were obtained under conditions of 0.1 M Mes (pH 6.5) and 12% poly(ethylene glycol) 20 000. The asymmetric unit contained one protein subunit in the F222-SurE crystal and four subunits in the C2-SurE crystal with Mathew's coefficients of 2.8 and 3.1, respectively. Data-collection statistics are given in Table 1.

Structure solution and quality of the model

F222-SurE was solved by molecular replacement to a resolution of 1.7 Å using Tm SurE as the phasing model. Subsequently, this refined model was used to solve the structure of the C2-SurE to a resolution of 2.7 Å. In both crystal forms the electron density accounted for the majority of the residues. There were no disordered regions in the map. The quality of refinement of the two crystal forms is shown in Table 1. Cis peptides were found between R81 and P82 and between G90 and I91. A similar cis peptide between G93 and V94 is found in Tm SurE also.

Overall structural features of the protein

St SurE is an αβα sandwich protein and adopts the Rossmann fold as in archaic and thermophilic homologues. The St SurE monomer consists of 13 β-strands, six α-helices and three 3_{10}-helices. The core of the protein is made up of a nine-stranded β-sheet flanked by α1, α5 and η2 on one side and by α2, α3, α4 and η1

Table 1. Data-collection and refinement statistics. Values in parentheses refer to the highest-resolution shell.

Unit cell parameters	Crystal form 1 (orthorhombic)	Crystal form 2 (monoclinic)
Data-collection statistics		
a, b, c (Å)	73.73, 121.64, 143.26	161.08, 95.30, 94.57
α, β, γ (°)	90, 90, 90	90, 98.9, 90
Space group	F222	C2
Resolution range (Å)	50–1.70 (1.76–1.7)	50–2.75 (2.85–2.75)
Total number of reflections	471 878	880 212
No. of unique reflections	35 384	37 552
Multiplicity	5.6	4.1
Mean ((I)/s(I))	26.3 (2.33)	14.3 (2.32)
R(merge)[a]	4.5 (47.7)	8.1 (49.9)
Completeness	99.7 (97.8)	99.6 (99.8)
Protomers in the asymmetric unit	1	4
Solvent content	55.6	60.2
Refinement statistics		
R_{work} (%)	18.3	18.9
R_{free} (%)	21.2	25.4
Model quality		
Number of atoms	2228	7676
Protein	1930	7582
Magnesium	2	4
Phosphate	5	10
Glycerol	6	6
Water	285	74
Mean B-factor (Å^2)		
Protein atoms	23.2	46.8
Water	37.0	43.8
Glycerol	43.5	54.5
Magnesium	21.1	42.2
Phosphate	18.9	36.6
rmsd from ideal values		
Bond length (Å)	0.007	0.016
Bond angle (degrees)	1.07	1.80
Residues in Ramachandran plot (%)		
Most allowed region	89.1	87.2
Allowed region	10.9	12.3
Generously allowed region	0.0	0.5
Disallowed region	0.0	0.0

[a] $R_{merge} = \sum_{hj} |I_{hj} - \langle I_h \rangle| / \sum I_{hj}$, where $\langle I_{hj} \rangle$ is the *j*th observation of reflection *h* and $\langle I_h \rangle$ is its mean intensity.

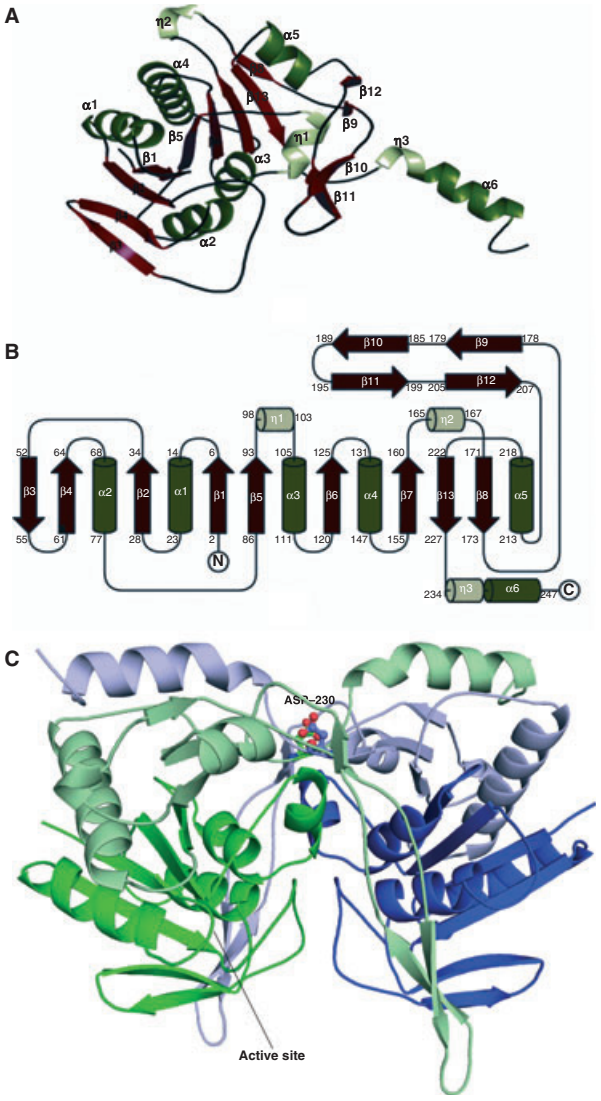

Fig. 1. (A) Monomeric structure of St SurE. (B) Topology of the secondary structural elements of St SurE. Arrows represent β-strands, while small and large cylinders represent 3_{10}-helices and α-helices, respectively. Numbers shown indicate the starting and ending residues of the secondary structural elements. (C) The domain-swapped dimer of St SurE. The A and B subunits are shown in green and blue, respectively. The darker and lighter shades of the respective colors represent the N-terminal and C-terminal domains of the corresponding subunits. D230, which is involved in crucial interactions that may be responsible for domain swapping, is shown at the top of the drawing. The arrow points towards the active site.

on the other side (Fig. 1A,B). One noticeable difference between St SurE and the other SurEs is the short length of η1 and α3 followed by a longer loop region. This region includes three conserved active-site residues (S104, G105 and T106) and occurs at the dimeric interface. A monomer comprises two domains (Fig. 1C). Residues 1–125 form the N-terminal domain that is mostly conserved among members of the SurE family. This domain consists of β1 to β6, α1 to α3 and η1. The sequence identity among pairs of SurEs varies from 31% to 42%. Residues 126 to 253 form the C-terminal domain, which consists of β7 to β13, α4, α5, η2 and η3. This domain is more divergent among the various SurEs and the sequence identities among the

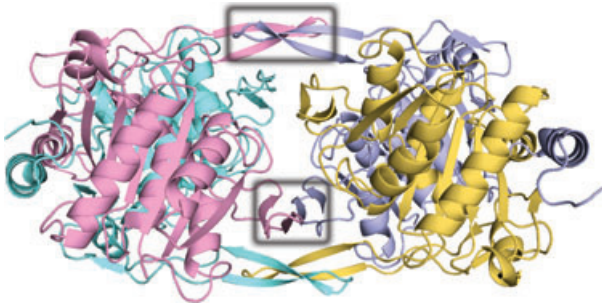

Fig. 2. Structure of the C2-SurE tetramer. The A to D subunits are represented by pink, cyan, blue and yellow, respectively. The tetrameric interactions are between the subunits A and C, and B and D, by way of the β hairpins and the loop, which connects β2 to β3 (boxed).

polypeptides in this region vary from 20% to 35%. The C-terminal domain consists of two protruding arms: (a) a C-terminal helical tail (η3, α6), which extends to the neighboring subunit to form a domain-swapped dimer; and (b) a β-hairpin, which is involved in intersubunit interactions, leading to a loose tetrameric organization of subunits (Fig. 2). The overall structure and the St SurE N-terminal domain superposed well with the corresponding domains in other SurEs yielding rmsd values of 0.81–1.26 Å. However, there was a larger variation in the C-terminal domains (yielding rmsd values ranging from 1.84 to 2.89 Å). A search using the DALI server with the C terminal domain as the query did not produce any significant hits other than the SurEs already known. However, the N-terminal domain aligned with certain other proteins such as the 1lss-A trk system potassium-uptake protein trka homolog, which is present ubiquitously in a variety of prokaryotic and eukaryotic K channels and transporters, and another universal stress protein with *Z*-scores of around 6. Similar results have been reported for Tm and Pa SurEs. Superposition of the protomers of monoclinic and orthorhombic forms using the program ALIGN [7] gave an rmsd of 0.54 Å between corresponding Cα atoms. The variation was larger with respect to the C-domain, which gave an rmsd of 0.70 Å.

Oligomeric status of the protein

The oligomeric status of SurE proteins is of much interest. Pa SurE has been reported to exist as a dimer in solution, whereas Ec SurE is a tetramer. Lee *et al.* [8] report a dimeric structure for Tm SurE, whereas Zhang *et al.* [9] report that Tm SurE exists as a tetrameric protein in solution. Iwasaki & Miki [10] demonstrated a dimer–tetramer equilibrium in Tt SurE based on sedimentation equilibrium experiments. In St SurE, the four subunits in the asymmetric unit of the monoclinic form, which were labeled as A, B, C and D, are related by 222 noncrystallographic symmetry. In the orthorhombic form, the subunits related by crystallographic twofolds form a tetramer very similar to that of the monoclinic form. The total surface-accessible area buried on dimerization of A and B subunits of the monoclinic form is 7633 $Å^2$, which suggests tight association. The area buried in the interface of the subunits related by twofold along **a** in F222-SurE is comparable (7000 $Å^2$). In contrast, the AD and AC interfaces of the monoclinic form bury only 419 and 902 $Å^2$ respectively. These buried areas are not sufficient for tight association. Gel filtration analysis and dynamic light-scattering experiments also indicated that SurE is a dimer in solution. Although the tetrameric unit is weakly held by AD and AC interfaces, it is interesting that a highly similar oligomeric structure is observed in Tt SurE and Tm SurE, which might imply that tetramerization has some physiological role.

All SurEs, with the exception of Pa SurE, have been reported to show domain swapping between the monomers of the dimer. Eisenberg and coworkers [11] have reported that Pa SurE predominantly exists in a nondomain-swapped dimeric form. Domain swapping is avoided in Pa SurE by a sharp turn in the segment of residues 242–245, bringing the polypeptide chain back to the same subunit. We analyzed the interactions of corresponding residues in SurEs to understand the reasons for domain swapping. A few strong interactions were observed that impart rigidity to this segment, preventing it from turning backwards in most SurEs. These interactions in St SurE include D230 OD1-T232 N (2.95 Å), D230 OD1-T232 OG1 (2.66 Å) and D230 OD2-H234 NE2 (2.89 Å) (Fig. 3). Thus, D230 is a crucial residue that appears to be essential for domain swapping. The equivalent residue is also conserved in Aa SurE (D234). H234 of St SurE is replaced with Y238 in Aa SurE. The hydroxyl group of Y238 is hydrogen bonded to D234 OD1 (2.55 Å). The other interactions of D234 in Aa SurE correspond to the interactions of D230 in St SurE. A different set of interactions appears to promote a rigid structure leading to domain swapping in Tt and Tm SurE. These or other interactions that impart a rigid structure are not observed in Pa SurE.

Intersubunit contacts

St SurE forms a strong domain-swapped dimer (Fig. 1C). The dimer is held mainly by hydrogen bonds

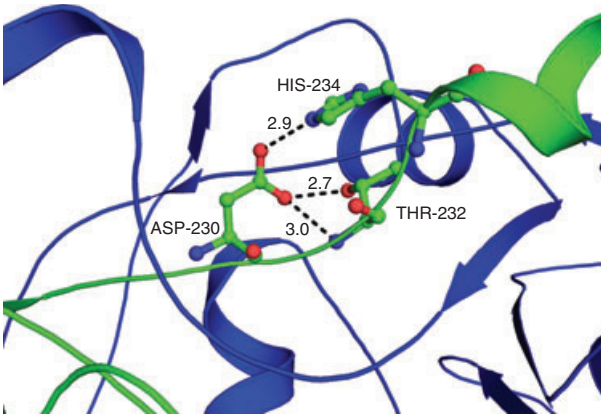

Fig. 3. Interactions of D230 that might be responsible for domain swapping in St SurE. The A and B subunits are shown in green and blue, respectively.

and van der Waals interactions. Most of the residues that are involved in the dimerization are found on the C-terminal helix responsible for the domain swapping. There are two salt bridges – between R52 and D190 and between R114 and D99 (cut off value 4.5 Å) – and 33 strong hydrogen bonds (cut off value 3 Å) across the dimeric interface. The thermophilic SurEs have a larger number of salt bridges compared with St SurE (Tm Sure-5, Tt SurE-7, Aa SurE-4). However, Pa SurE, which is also thermostable, has only a single salt bridge. In the tetrameric unit of St SurE, the number of contacts between the A and C subunits are more than the contacts between the A and D subunits. The stretch of residues from 186 to 198 that are present in the β-hairpin extension are mainly involved in the tetrameric interactions at the AC interface. Q188, P191 and W198 are the major contributors to several symmetry-related pairs of interactions at this interface. In the AD interface, there is a salt bridge between E48 and R192.

Stability of St SurE

Urea denaturation studies were carried out by incubating the protein at a concentration of 0.5 mg·mol^{-1} with varying concentrations of urea (0–7 M) for 4 h after which CD measurements were taken (Fig. 4A). The minimum near 222 nm, representative of helical structure, was disrupted by incubation of the protein with urea at a concentration of 2 M and it almost completely disappeared at around 4 M urea. It is noteworthy that the minimum near 209 nm, which represents the β-structure, was still retained, although reduced in intensity, even after incubation with 7 M urea. This might be because the β-sheets, which form the core of the protein, are still intact while the helices that form the outer layers of the sandwich protein have denatured.

Following this, thermal melting studies were conducted using circular diachronic monitoring at both 209 and 222 nm. The protein lost its secondary structure at a temperature of around 45 °C (Fig. 4B). This was also confirmed by differential scanning calorimetry (data not shown). Thus, the biphasic denaturation profile is observed with urea but not with thermal denaturation. The St SurE melting temperature is very low compared with its homologues, which are active even at temperatures above 80 °C. A detailed analysis of the amino acid residue composition, hydrophobic buried residues, salt bridges, hydrogen bonding, etc., could not account for the lower thermostability of St SurE. Following this, an analysis of the solvation free-energy of folding for the five SurEs was carried out using the software MSD PISA (Table 2). The solvation free-energy (SFE) follows a linear relationship with the number of residues and satisfies an empirical relationship SFE = 15.30−1.13N, where N is the number of residues in the polypeptide chain [12]. The table also shows the difference in percentage between SFEs calculated using PISA and expected on the basis of the linear relationship. This difference is small (< 10%) for well folded and stable proteins, whereas for misfolded proteins it is usually larger than 10% [12]. The large difference in SFEs for St SurE suggests that the lower value of SFE is the most probable reason for its lower thermal stability. This would be a consequence of the precise amino acid sequence of St SurE and its relationship to the polypeptide fold.

Phosphatase activity studies

The phosphatase activity of SurE against pNPP was measured at various temperatures, pH values and in the presence of different metal ions. The enzyme showed negligible activity at temperatures above 50 °C. This is in agreement with stability studies, which indicated that the denaturation temperature of St SurE is around 45 °C. The activity of St SurE was maximal at neutral pH (Fig. 4C). This was similar to that of Ec SurE. However, Tm SurE and Pa SurE have maximal activity towards pNPP at an acidic pH, around 5.5, and Tt SurE was maximally active at pH 8.2. St SurE shows almost no activity in the absence of divalent metal ions. Activation by various metal ions was in the order $Mg^{2+} > Mn^{2+} > Ca^{2+} > Zn^{2+} > Ni^{2+} > Co^{2+}$ (Fig. 4D). Tm SurE and Tt SurE also show maximum activity with Mg^{2+}. However, unlike Ec SurE and Pa SurE, there was negligible activity in the presence of

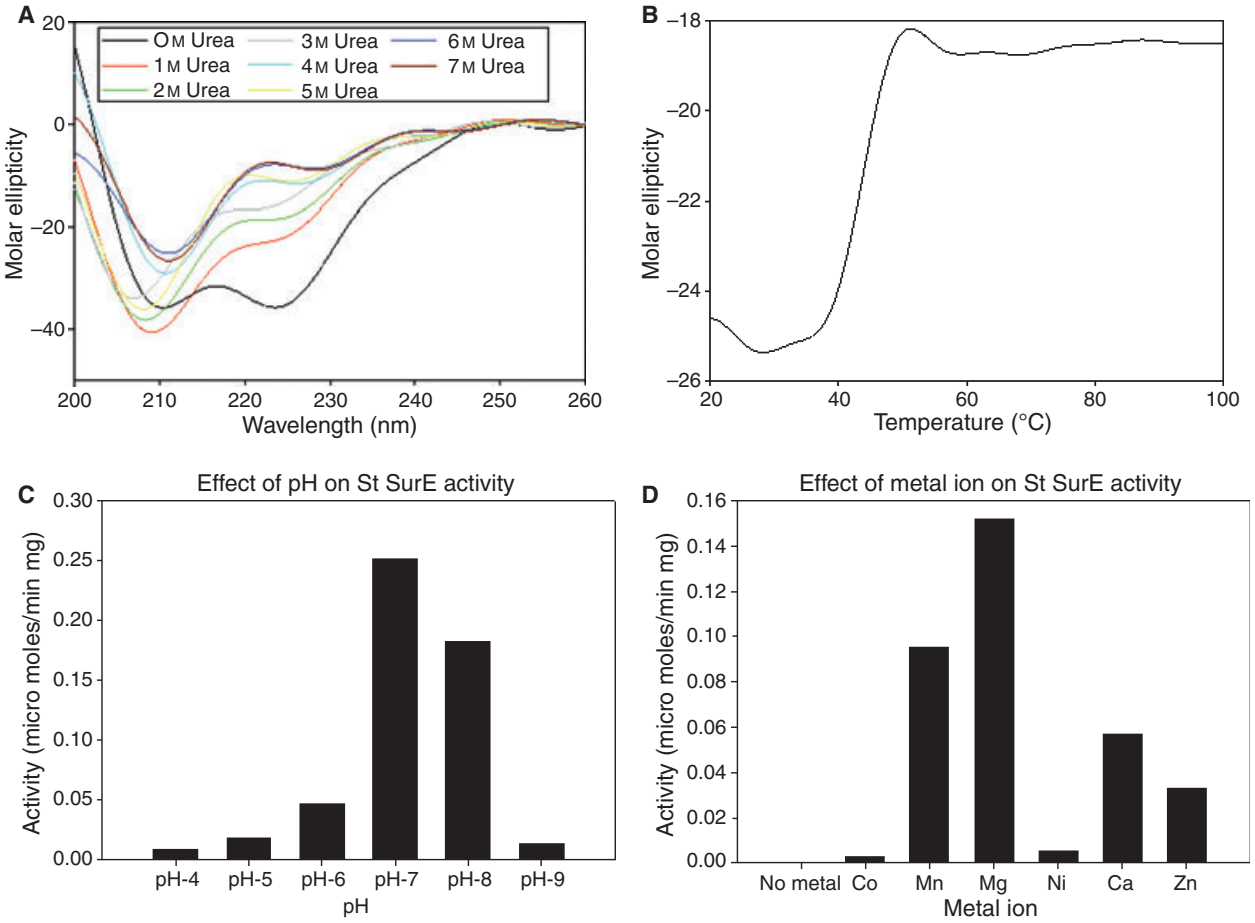

Fig. 4. (A) Urea denaturation profile of St SurE. (B) Thermal melting profile of St SurE. (C) Effect of pH on the phosphatase activity of St SurE using pNPP as the substrate. (D) Effect of divalent cations on the phosphatase activity of St SurE using pNPP as the substrate.

Table 2. Comparison of solvation free-energy of folding −ΔG (kcal·mol^{-1}) for SurEs from different sources.

Organism	No. of residues	ΔG (calculated)[a]	ΔG (predicted)[b]	Per cent difference[c]
Salmonella typhimurium	254	−203.6	−271.7	25.1
Thermotoga maritima	247	−239.7	−263.8	9.1
Thermus thermophilus	213	−207.0	−225.4	8.2
Aquifex aeolicus	248	−241.6	−264.9	8.8
Pyrobaculum aerophilum	277	−255.1	−297.7	14.3

[a] Calculated using the software MSD-PISA. [b] Calculated based on the equation SFE = 15.30–1.13*N*. [c] [ΔG(calculated)–ΔG(predicted)/ΔG predicted] × 100.

Co^{2+} for St SurE. With pNPP, St SurE showed a K_m of 4.8 mM and a V_{max} of 10.69 μmoles·min^{-1}·mg^{-1}, which were comparable with those of Ec SurE.

Geometry of the active site

The putative active site could be identified by the presence of bound magnesium and phosphate ions and by comparison with the active sites of other SurEs. There are 15 residues in the putative active site that are mostly conserved across the various SurEs. These residues are solely located in the N-terminal domain. The active site is found near the interface between the two monomers (Fig. 1C) and is partly acidic because of the presence of the conserved DD motif (D8 and D9) (Fig. 5A). As the precipitating condition used for crystallization did not contain any divalent cations, the magnesium ion is probably co-purified with the protein. The ion is coordinated by the carboxyl oxygen atoms of D8, D9, by the carboxamide oxygen of N92, by the hydroxyl oxygen of S39 and by the oxygen atoms of two water molecules, and has an approximate octahedral geometry (Fig. 5B). Not all of the six ligands coordinated to the metal ion in F222-SurE can

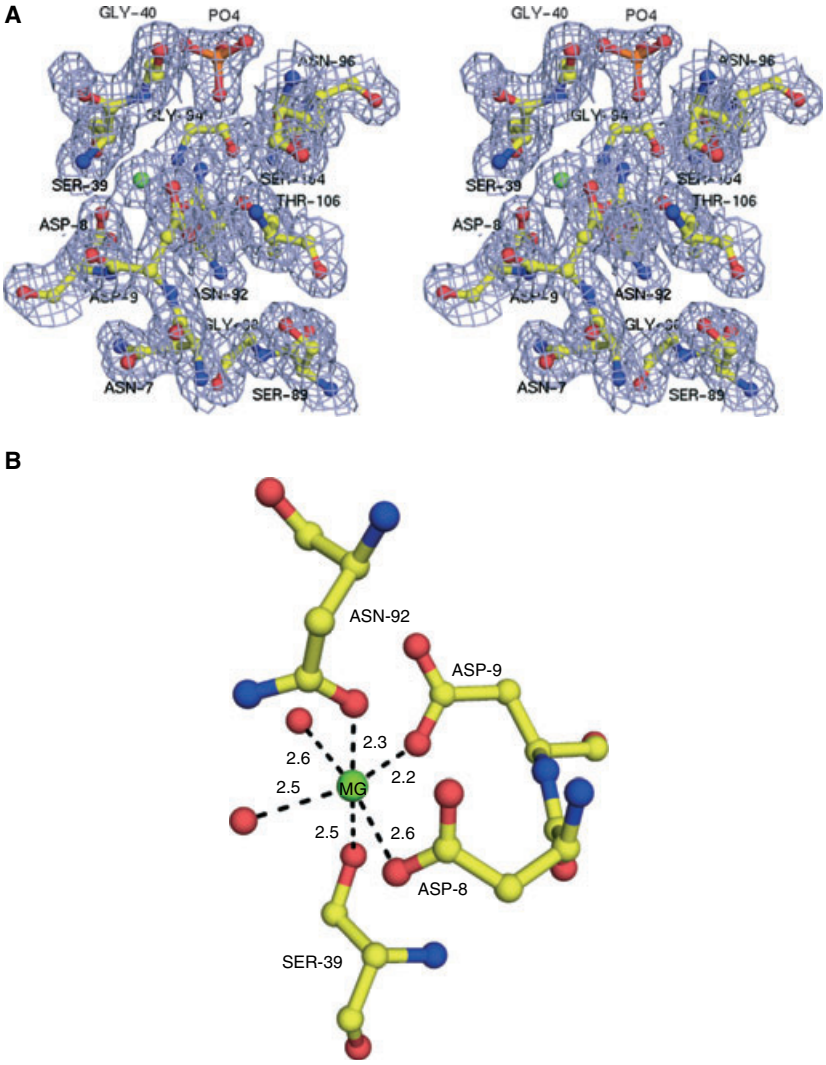

Fig. 5. (A) Stereo view of active-site residues with Mg^{2+} and phosphate with the (2Fo-Fc) electron density map contoured at the 1σ level. (B) Octahedral metal ion co-ordination in the active site of F222-SurE.

be seen in C2-SurE probably because of the lower resolution of the structure. When Tt SurE binds Mg^{2+} and AMP, the loop formed by residues 34–50 undergoes conformational change from a 'closed' to an 'open' form. The shift in the position of the loop away from the metal ion leads to the formation of two interprotomer active sites where the O_ε of E37 from one subunit coordinates with the metal ion bound to the other subunit. In St SurE, this loop is in a closed conformation. As Iwasaki and Miki [10] report, only if the loop is in an open conformation there is enough space in the active site to accommodate the substrate. In the closed form, the hydroxyl group of S39 coordinates with the metal ion of the same subunit. When AMP binds, its side chain clashes with the ribose moiety of the bound AMP if the loop does not adopt an open conformation. Co-crystallization and soaking trials with various putative substrates such as AMP, GMP and CMP have not so far been successful.

Apart from Mg^{2+} and water molecules, a strong tetrahedral density was found in the active site (Fig. 5A). A phosphate ion was modeled at this site assuming that it could have come from the cell because it is a putative phosphatase enzyme. The phosphate fitted well in the density. Refinement resulted in reasonable B values for the phosphates when the occupancy was set to 0.5. In the F222-SurE structure, the phosphate hydrogen bonds to several molecules of water and to OD1 of N96 whereas in the C2-SurE structure, phosphate bonds with Mg^{2+} (2.18 Å) and hydrogen bonds with the OG of S104 (2.87 Å), with OD1 and OD2 of D8 (2.82, 3.41 Å), with ND2 and OD1 of N96 (2.92, 3.02 Å), with OG1 of T106 (3.03 Å), with N of G105 (3.05 Å) and G40 (3.05 Å). An oxygen atom of

the phosphate coordinates with the metal ion in two of the four subunits in the asymmetric unit of C2-SurE. The interaction of the phosphate with G40 may prevent loop opening in St SurE.

Possible mechanism of phosphatase activity

The molecular mechanism of the phosphatase activity of SurE has not yet been established. The substrate might enter through an open channel that is present at the dimeric interface (Fig. 1C). This channel covers a volume of 889 $Å^3$ and an area of 535.5 $Å^2$ and is lined by the residues R36, N37, R38, S39, G40, Al41, S44, L45, T46, L47, E48, L51, M65 and T67.

The positions of phosphate in C2-SurE and F222-SurE are not identical. The distance between these positions is 2.76 Å. The position corresponding to the phosphate in C2-SurE is occupied by a water molecule (water 228) in F222-SurE. The strict conservation of the DD motif in SurEs strongly suggests that these residues are involved in catalysis. Only in C2-SurE is D8 at an appropriate distance from phosphate for a nucleophilic attack. The distance between OD1 of D8 and the phosphate is 3.26 Å in C2-SurE and 5.74 Å in F222-SurE. The distance of the phosphate from the metal ion is also greater in F222-SurE (5.96 Å) when compared with C2-SurE (3.46 Å). These observations suggest that the phosphate position in C2-SurE is most likely to correspond to that of the substrate mononucleoside phosphate. The interactions of phosphorus with the metal ion in C2-SurE may help to polarize the phosphate bond for nucleophilic attack. The divalent metal ion might also orient the nucleophilic D8 so that it can attack the phosphorus atom of the monophosphate. This nucleophilic attack may lead to a phosphorylated D8 enzyme intermediate. The next step in catalysis requires an activated water molecule to produce an OH^- ion, which will hydrolyze the D8 phosphate intermediate to release a phosphate and regenerate the active enzyme. Unfortunately, the water molecule suitable for such activation was not found in C2-SurE, perhaps because of its low resolution. Several water molecules have been located in the active site of F222-SurE. However, these water molecules may not correspond to that required for catalysis as the phosphate is probably not optimally placed in F222-SurE. The mechanism proposed here is similar to that of TA0175, a phosphoglycolate phosphatase belonging to the HAD superfamily of proteins [13] with which SurE shows several common active-site features. The nucleotidase activity of SurE may be required for phosphorus scavenging and remobilization when the cells are under stress and consequently mononucleotide phosphate concentrations are high.

Materials and methods

Cloning and purification of St SurE

The *SurE* gene was PCR amplified from *Salmonella enterica* Typhimurum strain IFO12529 genomic DNA as the template using Deep Vent DNA polymerase (New England Biolabs, Ipswich, MA, USA), a sense primer (CATATGGCTAGC ATGCGCATATTGCTGAGTAAC) containing an *Nhe*I site and an antisense primer (TTAGGATCCTTACCATTGCG TGCCAACTCCCAC) containing a *Bam*HI site. The gene was cloned at the *Nhe*I and *Bam*HI sites of the pRSET-C vector (Invitrogen, Carlsbad, CA, USA). The sequence of the recombinant SurE clone obtained was determined and confirmed by comparison with the *S. typhimurium* genome. The cloning strategy resulted in 14 additional amino acids from the vector at the N terminus of the expressed protein, including a hexahistidine tag that facilitated protein purification.

The recombinant plasmid was transformed into BL21(DE3) pLysS cells and transformants were selected on LB agar plates containing 100 $\mu g \cdot mL^{-1}$ of ampicillin. A single colony was picked and cultured in 25 mL of Terrific broth overnight at 37 °C then inoculated into 500 mL of Terrific broth containing 2 mL of glycerol and grown until the *D* reached 0.6 at 600 nm. Protein expression was then induced by 0.3 mM IPTG. The culture was further incubated at 30 °C for 6 h. Cells were harvested by centrifugation, resuspended in lysis buffer that contained 50 mM Tris (pH 8), 200 mM NaCl, 2% Triton X-100 and 30% glycerol, and then lysed by sonication on ice. The supernatant obtained after centrifugation of the cell lysate was gently mixed with Ni-nitrilotriacetic acid resin for 2–3 h and then loaded onto a glass column. Nonspecifically bound proteins were washed from the column using 50 mL of lysis buffer containing 20 mM imidazole. The recombinant protein was eluted with 5 mL of lysis buffer containing 200 mM imidazole. The protein was dialyzed extensively against 25 mM Tris (pH 8) containing 100 mM NaCl to remove imidazole. St SurE was obtained with a final yield of 30 $mg \cdot L^{-1}$ of cell culture. The protein was concentrated using a 10-kDa molecular mass cut-off Amicon Ultra-15 Centrifugal Filter Unit (Millipore, Billerica, MA, USA).

Initial characterization of the protein

The purity and molecular mass of the protein were checked by 12% SDS/PAGE and MALDI-TOF MS. Gel-permeation chromatography with 200 μL of a 1 $mg \cdot mL^{-1}$ protein solution was performed using a Superdex S-200 column with a bed volume of 28 mL and a void volume of 8 mL. Dynamic light-scattering experiments were carried out using a VISCOTEK dynamic light-scattering particle size analyzer with a data-acquisition time of 10 s. The hydrodynamic radius was calculated using the OMNISIZE 3 software. CD measurements for St SurE were recorded at a protein concentration of

0.3 mg·mL^{-1} in a buffer containing 25 mM Tris (pH 8) using a Jasco J715 spectropolarimeter. Urea denaturation experiments were performed by incubating the protein at a concentration of 0.5 mg·mL^{-1} with varying concentrations of urea (0–7 M at 1-M intervals). After 4 h, far-UV CD-spectra were recorded for these samples. Thermal melting studies were carried out using a differential scanning calorimeter VP-DSC (Microcal Inc., Northampton, MA, USA). The differential scanning calorimetry scan was carried out with the protein in the sample cell at a concentration of 0.5 mg·mL^{-1} and buffer in the reference cell. Data were analyzed using the ORIGIN software provided by Microcal Inc., along with the instrument.

Crystallization, data collection and processing

Crystallization experiments were carried out with Hampton crystallization screens using the microbatch method with a mixture of silicon and paraffin oil in equal proportions layered over the crystallization droplets. F222 crystals were obtained under various conditions containing trisodium citrate as the precipitant. The crystal used for data collection was obtained with 0.2 M trisodium citrate dihydrate, 0.1 M Hepes (pH 7.5) and 30% 2-methyl-2,4-pentanediol. C2 crystals were produced in the presence of 12% poly(ethylene glycol) 20 000 in 0.1 M Mes (pH 6.5). The crystals were transferred to the crystallization buffer containing 20% glycerol as the cryoprotectant for a few seconds and then mounted in a cryo-loop. X-ray diffraction data were collected at 100 K using a RU300 rotating-anode X-ray generator and a MAR Research (Hamburg, Germany) image plate detector system. The data sets were processed using DENZO and the resulting intensities were scaled using SCALEPACK [14]. Data-collection and processing statistics are given in Table 1.

Structure solution and refinement

The structure of the orthorhombic St SurE was determined by molecular replacement with AMoRe [15] using the atomic coordinates of Tm SurE, which shared 36% sequence identity with St SurE as the phasing model. A dimer of St SurE was used as the search model for the structure solution of the monoclinic form. Both structures were refined using REFMAC5 [16]. The protein models were improved by visual inspection and manual model building using the graphics program COOT [17]. The progress of refinement was monitored by calculation of R*free* [18] using 5% of the total independent reflections that were not included in the refinement. Stereochemical qualities of the models were verified using PROCHECK [19].

Structural analysis

DSSP was used to assign the secondary structure of the protein. NACCESS [20] was used to calculate buried surface areas. The program CONTACT from the CCP4 suite was used for the identification of intersubunit contacts. The PISA [21] server was used for calculating the solvation free-energy of folding. All structural superpositions and the rmsd values for SurE were determined using the program ALIGN [7]. The DALI server [22] was used for homology model searching. Average B-factors for protein atoms, water molecules and ligands were calculated using the BAVERAGE program of the CCP4 suite. Void analysis was carried out using the CASTp server [23]. The figures were prepared using PYMOL [24]. The topology diagram was prepared using TOPDRAW [25].

Enzyme activity assays

Phosphatase activity of St SurE was measured against pNPP, a general substrate for all phosphatases. pNPP is converted to para-nitrophenol by the phosphatase, the formation of which can be monitored directly at 405 nm. Absorbance measurements were carried out on a JASCO UV-visible spectrophotometer model V-530 (Japan Specroscopic Co., Japan). The protein used for assays was dialyzed against 10 mM EDTA to remove the metal ion bound in the active site (this was identified in the crystal structure). The reaction mixture (150 μL) contained 100 mM cacodylate buffer (pH 7), 5–10 mM metal, 0.5–40 mM pNPP and 1–20 μg of protein. All experiments were replicated three times using a fresh batch of purified protein and checked for consistency of results. The pH dependence of phosphatase activity towards pNPP (5 mM) was determined in the presence of 10 mM $MgCl_2$ and 10 μg of the enzyme using buffers (100 mM) of varying pH from 4 to 9 (sodium citrate, pH 4; sodium acetate, pH 5; CAPSO, pH 6; cacodylate/Hepes, pH 7; Tris/HCl, pH 8; and MOPS, pH 9). The metal dependence of activity was determined at the optimum pH of 7 (100 mM cacodylate buffer, pH 7) using various metals (10 mM), substrate (pNPP, 5 mM) and 10 μg of protein. For determination of K_m and V_{max}, the phosphatase assays contained substrate at concentrations of 0.5 to 40 mM. Kinetic parameters were determined by nonlinear curve fitting to the Lineweaver–Burk plot using Graphpad PRISM software.

Acknowledgements

The diffraction data were collected at the X-ray facility for structural biology at the Molecular Biophysics Unit, Indian Institute of Science, supported by the Department of Science and Technology (DST) and the Department of Biotechnology (DBT) of the Government of India. MRN and HSS thank DST and DBT for financial support. AP acknowledges the Council for Scientific and Industrial Research (CSIR), Government of India

for the award of a senior research fellowship. We thank Simanshu, Garima and Eugene Krissinel for help with experiments and useful discussions.

References

1 Li C, Ichikawa JK, Ravetto JJ, Kuo HC, Fu JC & Clarke S (1994) A new gene involved in stationary-phase survival located at 59 minutes on the Escherichia coli chromosome. *J Bacteriol* **176**, 6015–6022.

2 Riehle MM, Bennett AF & Long AD (2001) Genetic architecture of thermal adaptation in Escherichia coli. *Proc Natl Acad Sci U S A* **98**, 525–530.

3 Treton BY, Le Dall MT & Gaillardin CM (1992) Complementation of Saccharomyces cerevisiae acid phosphatase mutation by a genomic sequence from the yeast Yarrowia lipolytica identifies a new phosphatase. *Curr Genet* **22**, 345–355.

4 Proudfoot M, Kuznetsova E, Brown G, Rao NN, Kitagawa M, Mori H, Savchenko A & Yakunin AF (2004) General enzymatic screens identify three new nucleotidases in Escherichia coli. Biochemical characterization of SurE, YfbR, and YjjG. *J Biol Chem* **279**, 54687–54694.

5 Bilodeau D & Beliveau R (1999) Inhibition of GTP-gammaS-dependent L-isoaspartyl protein methylation by tyrosine kinase inhibitors in kidney. *Cell Signal* **11**, 45–52.

6 Visick JE, Ichikawa JK & Clarke S (1998) Mutations in the Escherichia coli surE gene increase isoaspartyl accumulation in a strain lacking the pcm repair methyltransferase but suppress stress-survival phenotypes. *FEMS Microbiol Lett* **167**, 19–25.

7 Cohen GE (1997) ALIGN: a program to superinpose protein coordinates accounting for insertions and deletions. *J Appl Cryst* **30**, 1160–1161.

8 Lee JY, Kwak JE, Moon J, Eom SH, Liong EC, Pedelacq JD, Berendzen J & Suh SW (2001) Crystal structure and functional analysis of the SurE protein identify a novel phosphatase family. *Nat Struct Biol* **8**, 789–794.

9 Zhang RG, Skarina T, Katz JE, Beasley S, Khachatryan A, Vyas S, Arrowsmith CH, Clarke S, Edwards A, Joachimiak A *et al.* (2001) Structure of Thermotoga maritima stationary phase survival protein SurE: a novel acid phosphatase. *Structure* **9**, 1095–1106.

10 Iwasaki W & Miki K (2007) Crystal structure of the stationary phase survival protein SurE with metal ion and AMP. *J Mol Biol* **371**, 123–136.

11 Mura C, Katz JE, Clarke SG & Eisenberg D (2003) Structure and function of an archaeal homolog of survival protein E (SurEalpha): an acid phosphatase with purine nucleotide specificity. *J Mol Biol* **326**, 1559–1575.

12 Chiche L, Gregoret LM, Cohen FE & Kollman PA (1990) Protein model structure evaluation using the solvation free energy of folding. *Proc Natl Acad Sci USA* **87**, 3240–3243.

13 Kim Y, Yakunin AF, Kuznetsova E, Xu X, Pennycooke M, Gu J, Cheung F, Proudfoot M, Arrowsmith CH, Joachimiak A *et al.* (2004) Structure- and function-based characterization of a new phosphoglycolate phosphatase from Thermoplasma acidophilum. *J Biol Chem* **279**, 517–526.

14 Otwinowsky Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326.

15 Navaza J (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr A* **50**, 157–163.

16 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.

17 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.

18 Brunger AT (1993) Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallogr D* **49**, 24–36.

19 Laskowski RA, Moss DS & Thornton JM (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231**, 1049–1067.

20 Hubbard SJ & Thornton JM (1993) *NACCESS, Computer Program*. Department of Biochemistry and Molecular Biology, University College, London.

21 Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.

22 Holm L & Sander C (1995) Dali: a network tool for protein structure comparison. *Trends Biochem Sci* **20**, 478–480.

23 Binkowski TA, Naghibzadeh S & Liang J (2003) CASTp: computed atlas of surface topography of proteins. *Nucleic Acids Res* **31**, 3352–3355.

24 DeLano WL (2002) *The PYMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA.

25 Bond CS (2003) TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* **19**, 311–312.