Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Three-dimensional structure of the type III secretion chaperone SycE from *Yersinia pestis*

Artem G. Evdokimov, Joseph E. Tropea, Karen M. Routzahn and David S. Waugh*

Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, PO Box B, Frederick, MD 21702-1201, USA

Correspondence e-mail: waughd@ncifcrf.gov

Received 5 October 2001
Accepted 2 January 2002

PDB Reference: SycE, 1k6z, r1k6zsf.

Many bacterial pathogens utilize a type III (contact-dependent) secretion system to inject cytotoxic effector proteins directly into host cells. This ingenious mechanism, designed for both bacterial offense and defense, has been studied most extensively in *Yersinia* spp. To be exported efficiently, at least three of the effectors (YopE, YopH and YopT) and several other proteins that transit the type III secretion pathway in *Yersinia* (YopN, YopD and YopB) must first form transient complexes with cognate-specific Yop chaperone (Syc) proteins. The cytotoxic effector YopE, a selective activator of mammalian Rho-family GTPases, associates with SycE. Here, the structure of *Y. pestis* SycE at 1.95 Å resolution is reported. SycE possesses a novel fold with an unusual dimerization motif and an intriguing basic cavity located on the dyad axis of the dimer that may participate in its interaction with YopE.

1. Introduction

Gram-negative bacterial pathogens of plants and animals have evolved a number of versatile mechanisms to transport virulence factors across the bacterial and host cell membranes (Koster *et al.*, 2000; Donnenberg, 2000; Lee & Schneewind, 2001). One of the more common secretion/translocation systems, which has been termed type III, has been studied most extensively in *Yersinia* spp. (Cornelis *et al.*, 1998; Cornelis, 2000; Cheng & Schneewind, 2000). Yet, despite the fact that virtually all of the *Yersinia* proteins involved in type III secretion have been identified, remarkably little is known about how they interact to orchestrate the secretion of the effectors from the bacterium and their translocation across the eukaryotic cell membrane. The molecular machinery involved in type III secretion in *Y. pestis* is encoded by a 70 kbp plasmid known as pCD1 (Perry *et al.*, 1998). When activated by contact with mammalian cells, the type III secretion apparatus somehow pierces both bacterial membranes as well as the eukaryotic cell membrane and injects anti-host effectors called Yops (*Yersinia* outer proteins) directly into the cytosol of the host.

The primary secretion signals for the Yops are thought to be located near their N-termini, roughly within the first 20 amino acids (Sory *et al.*, 1995; Schesser *et al.*, 1996). A competitive hypothesis, which posits that the signals are encoded in the *yop* mRNAs rather than in the amino-acid sequence of the proteins (Anderson & Schneewind, 1997), has been vigorously disputed (Lloyd *et al.*, 2001). Regardless of the true nature of the primary secretion signal, it is widely accepted that several of the Yop proteins depend on additional signals for efficient secretion and translocation (Wattiau & Cornelis, 1993; Wattiau *et al.*, 1994; Cheng *et al.*, 1997). These secondary

signals, which are located immediately adjacent to the primary secretion signals near the N-termini of the Yops, correspond to the binding sites for cognate type III secretion chaperones known as Sycs (specific Yop chaperones). The primary role of the Sycs may be to assist in establishing a hierarchy of secretion for the various effectors, thereby controlling the timing of their delivery into eukaryotic cells (Lloyd *et al.*, 2001). Alternatively or in addition, the Sycs may act to maintain their cognate Yops in a translocation-competent (perhaps partially unfolded) state, protect them from proteolytic degradation and prevent premature interactions between different Yops or between the Yops and the common components of bacterial cells (Wattiau *et al.*, 1996). Known *Yersinia* chaperone/escortee pairs include SycE/YopE, SycH/YopH, SycT/YopT, SycN+YscB/YopN, LcrH/YopB and LcrH/YopD (Cornelis & Van Gijsegem, 2000). Type III secretion chaperones have also been identified in other Gram-negative bacterial pathogens (reviewed by Wattiau *et al.*, 1996; Bennett & Hughes, 2000). Although a few of the *Yersinia* secretion chaperones have recognizable counterparts in other organisms (*e.g.* SycE, SycN and YscB), there is little or no sequence similarity between the different members of this family. However, it has been suggested that at least several of the type III secretion chaperones may possess a leucine-zipper motif followed by a C-terminal amphipathic α-helix (Wattiau *et al.*, 1996).

SycE, which is also known as YerA (Forsberg & Wolf-Watz, 1990), is the cognate secretion chaperone for YopE, a selective activator of mammalian Rho-family GTPases. SycE reportedly binds to amino acids 15–50 of YopE (Woestyn *et al.*, 1996). According to Cheng & Schneewind (1999), SycE is a homodimer in solution and each YopE polypeptide binds one SycE homodimer. Here, we describe the three-dimensional structure of SycE determined by means of cryogenic X-ray crystallography and discuss its impact on our understanding of the SycE–YopE interaction and the possible structural relationships between the type III secretion chaperones in general.

2. Materials and methods

2.1. Protein expression and purification

The MBP-FLAG-SycE-His_6 expression vector was constructed by polymerase chain reaction (PCR) amplification of the open reading frame (ORF) encoding SycE from *Y. pestis* genomic DNA with the following oligonucleotide primers: 5′-GAG AAC CTG TAC TTC CAG GAT TAC AAA GAC GAG ATG TAT TCA TTT GAA CAA GCT ATC-3′ and 5′-ATT AGT GAT GAT GGT GGT GAT GAC TAA ATG ACC GTG GTG GTG AG-3′. This PCR amplicon was subsequently used as the template for a second PCR with the following primers: PE-277, 5′-GGG GAC AAG TTT GTA CAA AAA AGC AGG CTC GGA GAA CCT GTA CTT CCA G-3′ and PE-278, 5′-GGG GAC CAC TTT GTA CAA GAA AGC TGG GTT ATT AGT GAT GAT GGT GGT GAT G-3′, generating a second amplicon that was inserted by recombinational cloning into the entry vector pDONR201 (Invitrogen) to create pKM937. The nucleotide sequence of the entire insert was then confirmed experimentally. Next, the SycE ORF, now bracketed by a hexahistidine tag on its C-terminus and a recognition site for tobacco etch virus (TEV) protease followed by an abbreviated FLAG tag on its N-terminus (ENLYFQDYKDE), was moved from pKM937 by recombinational cloning into the destination vector pKM596 (Evdokimov *et al.*, 2000) to create pKM941. pKM941 directs the expression of FLAG-SycE-His_6 as a fusion to the C-terminus of *Escherichia coli* maltose-binding protein (MBP). The MBP moiety can be removed by cleaving the fusion protein with TEV protease at a designed site in the linker to yield a recombinant *Y. pestis* SycE protein with an abbreviated FLAG tag (DYKDE) on its N-terminus and a hexahistidine tag on its C-terminus. The MBP-FLAG-SycE-His_6 fusion protein was expressed in *E. coli* BL21(DE3) cells that also contained the *argU*/*ileX* tRNA accessory plasmid pKC1 and the TEV protease expression vector pRK603 (Kapust & Waugh, 2000). Cells containing the expression vector were grown to mid-log phase (OD_{600nm} = 0.5) at 310 K in Luria broth containing 100 μg ml^{-1} ampicillin, 30 μg ml^{-1} chloramphenicol and 25 μg ml^{-1} kanamycin, at which time isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 m*M* and the temperature was reduced to 303 K. 4 h after induction, the cells were pelleted by centrifugation and stored at 193 K. Selenomethionine-substituted FLAG-SycE-His_6 was produced in the same way, using the media formulation described by Doublié (1997).

The native (untagged) SycE expression vector (pKM913) was constructed in essentially the same manner as described previously (Cheng & Schneewind, 1999), except that the PCR amplicon was inserted into pET11c (Novagen) instead of pET9a. The nucleotide sequence of the entire insert was then confirmed experimentally. The untagged SycE protein was expressed in *E. coli* BL21(DE3)-CodonPlus-RIL cells (Stratagene) as described above for FLAG-SycE-His_6, except that kanamycin was omitted from the medium.

E. coli cell paste containing FLAG-SycE-His_6 was suspended in ice-cold 50 m*M* sodium phosphate pH 8, 300 m*M* NaCl (buffer *A*) containing Complete protease inhibitor cocktail (Roche Molecular Biochemicals) and disrupted with an APV Gaulin Model G1000 homogenizer at 69 MPa. The homogenate was centrifuged at 20 000*g* for 30 min at 277 K to remove insoluble material. The supernatant was filtered through a 0.45 μm cellulose acetate membrane and applied to a 50 ml Ni–NTA Superflow affinity column (Qiagen) equilibrated in buffer *A*. The column was washed with five column volumes of equilibration buffer followed by five column volumes of buffer *A* containing 25 m*M* imidazole to remove non-specifically bound proteins. Elution was carried out with a linear gradient from 25 to 250 m*M* imidazole in buffer *A*. Fractions containing recombinant FLAG-SycE-His_6 were pooled and ethylenediaminetetraacetic acid (EDTA) was added to a final concentration of 1 m*M*. The sample was concentrated by diafiltration and passed through a 100 ml amylose affinity column (New England Biolabs) equilibrated

in buffer *A* containing 1 m*M* EDTA to remove the undigested fusion protein. The column effluent was mixed with an equal volume of 50 m*M* sodium phosphate pH 7.2, 2 *M* ammonium sulfate and reduced with dithiotheitol (DTT) at a final concentration of 5 m*M*. The sample was applied to a HiPrep 16/10 butyl agarose column (Amersham Pharmacia Biotech) equilibrated in 50 m*M* sodium phosphate pH 7.2, 1 *M* ammonium sulfate. The column was washed with five column volumes of equilibration buffer after which the bound material was eluted with a linear gradient from 1 to 0 *M* ammonium sulfate over five column volumes. Fractions containing recombinant protein were pooled, concentrated and fractionated on a HiPrep 26/60 Sephacryl S-100 HR column (Amersham Pharmacia Biotech) equilibrated in buffer *A* containing 1 m*M* EDTA and 5 m*M* DTT. The relevant fractions were pooled and dialyzed against 25 m*M* bis-tris pH 6, 5 m*M* DTT (buffer *B*). The sample was then applied to a 1 ml Mono Q column (Amersham Pharmacia Biotech) equilibrated in buffer *B* and eluted with a linear gradient of 0–0.4 *M* NaCl in buffer *B* over 50 column volumes. Fractions containing FLAG-SycE-His_6 were pooled, dialyzed against 20 m*M* Tris–HCl pH 7.4, 50 m*M* NaCl, 5 m*M* DTT and then concentrated to 15 mg ml^{-1} (determined spectrophotmetrically using a molar extinction coefficient of 15 220 M^{-1} cm^{-1}). Aliquots were flash-frozen with liquid nitrogen and stored at 193 K until use. The final product was judged to be >95% pure on the basis of silver staining after SDS–PAGE (data not shown). The molecular weight was confirmed by electrospay mass spectrometry; the efficiency of selenomethionine incorporation was >99%.

E. coli cell paste containing native (untagged) SycE was suspended in ice-cold 50 m*M* bis-tris pH 5.8 (buffer *C*) containing Complete protease inhibitor cocktail (Roche Molecular Biochemicals) and 1 m*M* benzamidine and disrupted with an APV Gaulin Model G1000 homogenizer at 69 MPa. The homogenate was centrifuged at 20 000*g* for 30 min at 277 K to remove insoluble material. The supernatant was filtered through a 0.45 µm cellulose acetate membrane and applied to a HiPrep 16/10 Q FF column (Amersham Pharmacia Biotech) equilibrated in buffer *C*. The column was extensively washed with equilibration buffer and then eluted with a linear gradient of 0–0.5 *M* NaCl in buffer *C*. Fractions containing native SycE were pooled, reduced with DTT at a final concentration of 5 m*M*, diluted with buffer *C* and then rechromatographed on the HiPrep Q column. The peak fractions were pooled, reduced with DTT and mixed with an equal volume of 50 m*M* sodium phosphate pH 7.2, 2 *M* ammonium sulfate. The sample was fractionated on a HiPrep 16/10 butyl agarose column as described for FLAG-SycE-His_6. Fractions containing SycE were pooled, diluted with buffer *A* and applied to a 50 ml Ni–NTA Superflow affinity column which absorbed the remaining contaminants. The column effluent was pooled and concentrated to 3 mg ml^{-1} (determined spectrophotmetrically using a molar extinction coefficient of 13 940 M^{-1} m^{-1}). Aliquots were flash-frozen with liquid nitrogen and stored at 193 K until use. The final product was judged to be >90% pure on the basis of silver staining after SDS–PAGE (data not shown). The molecular weight was confirmed by electrospay mass spectrometry.

2.2. Peptide synthesis

Peptides SP-3485 (NH_2-MKISSFISTSLPLPTSVSGSSSV-GEMSGRSVSQ-COOH), SP-3486 (NH_2-QTSDQYANN-LAGRTESPQGSSLASRIIE-COOH), SP-3487 (NH_2-RLSS-VAHSVIGFIQRMFSEGSHKPVVTP-COOH) and SP3461 (Biotin-NH-TSVSGSSSVGEMSGRSVSQQTSNQYANN-LAGRTESP-COOH) were assembled on an amide resin using 9-fluorenylmethoxycarbonyl chemistry. All residues except for the biotin moiety were added automatically on a model 430A peptide synthesizer (Applied Biosystems). Following deprotection, the peptides were purified by reverse-phase high-performance liquid chromatography with an increasing gradient of acetonitrile in water. The purified peptides had the correct amino-acid composition and electrospray mass spectrometry yielded the predicted molecular weights.

2.3. Calorimetry

Native (untagged) SycE was dialyzed against three changes of 100 volumes of 25 m*M* Tris–HCl pH 7.5, 30 m*M* NaCl, after which the protein concentration was adjusted to 68.2 µ*M* on the basis of the theoretical specific optical density at 280 nm, as determined for the protein unfolded in 6 *M* guanidinium hydrochloride. An aliquot of SycE in solution was degassed, loaded into the 1.5 ml chamber of a MicroCal VP-ITC calorimeter and allowed to equilibrate to 293 K. The protein was titrated (30 injections of 10 µl) with the same buffer containing 560 µ*M* peptide concentration. A single-site binding model in the *Origin VI* software (MicroCal Inc.) was used to obtain thermodynamic parameters from the experimental results.

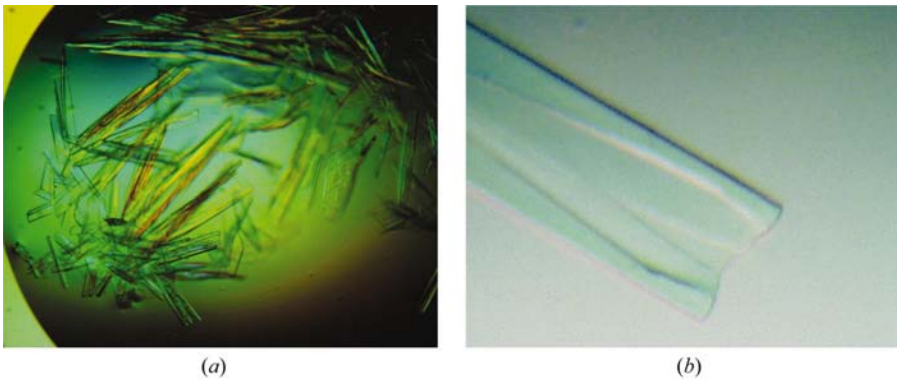

Figure 1
Crystals of SycE (*a*) obtained by screening and (*b*) optimized by the addition of imidazole.

2.4. Crystallization and data collection

All crystals were grown by vapor diffusion in VDX 24-well plates containing 1 ml of precipitant solution per well. Crystallization trials were performed with sparse-matrix kits obtained from Hampton Research and Emerald Biostructures. SycE crystallized under several neutral conditions with polyethylene glycol 8000 (PEG 8000) as the precipitant (Fig. 1*a*). These crystals were much improved by the addition of 0.4–0.6 *M* imidazole acetate pH 7.3 (Fig. 1*b*). The optimum conditions were 20% PEG 8000, 500 m*M* imidazole acetate pH 7.3, with a 3 µl:3 µl drop:reservoir ratio. Initially, multiple crystals appeared after 3–5 d. However, with careful seeding within 4 h of the setup time, a few large crystals could be grown over a period of 2–7 d. Crystals of selenomethionine-substituted SycE were grown in exactly the same manner, using the native protein for seeding.

Cryoprotection with artificial mother liquor containing glycerol or ethylene glycol was unsuccessful; freezing in oil was found to be much more useful in this case. Unfortunately, SycE crystals almost invariably grew as rectagonal prisms with large invaginations on one or both ends (Fig. 1*b*). The solvent trapped in these cavities made the crystals difficult to freeze in oil. A satisfactory solution to this problem was found by immersing the crystals in a fine suspension of a small quantity of ethylene glycol in Paratone-N. The suspended ethylene glycol quickly mixed with the mother liquor trapped in the fissures, which prevented the latter from crystallizing while not being present in sufficient quantity to adversely affect the bulk of the crystal. SycE crystals dipped in such a suspension were mounted immediately in a monofilament loop and flash-frozen in a cryogenic nitrogen stream (Oxford Cryosystems Cryostream) at 100 K. X-ray diffraction data were recorded using a Quantum 4 CCD (ADSC) at the National Synchrotron Light Source X-ray beamline X9B. The crystals were determined to belong to the space group $P2_1$. Data were collected at three X-ray wavelengths selected on the basis of the fluorescence scan. Diffraction was recorded over a φ rotation range of 180° for each wavelength, in 1° oscillations with an exposure time of 2 min $^{\circ -1}$, using a crystal-to-detector distance of 120 mm. The data were reduced and scaled with *HKL* (Otwinowski & Minor, 1997). The parameters for data collection are summarized in Table 1.

2.5. Structure determination

The high quality of the selenium-peak data made it possible to locate four Se atoms by direct methods, with a resulting Patterson correlation coefficient of 0.77 (*SHELXD*; Usón & Sheldrick, 1999). The presence of four methionines corresponded to two mole-

Table 1
Data collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Unit-cell parameters (Å, °)	a = 51.99, b = 56.163, c = 55.109, β = 113.94		
Space group	$P2_1$		
Mosaicity (°)	0.45		
Wavelength (Å)	0.9796 (inflection)	0.9795 (peak)	0.9400 (remote)
Resolution (Å)	30–1.95	30–2.1	30–2.0
Completeness (%)	99.9 (99.7)	99.2 (94.1)	99.0 (94.9)
Redundancy	1.90	1.92	1.7
Unique reflections	21311	17062	19444
$I/\sigma(I)$	14.6 (2.3)	17.97 (2.5)	13.97 (2.1)
$R_{\rm merge}$ (%)	5.0 (27.2)	4.2 (28.3)	5.4 (35.0)
Anomalous differences (%)	9.0	8.2	7.1
Dispersive differences (%)	4.74 (peak/remote); 7.05 (remote/inflection); 5.85 (remote/peak)		
Figure of merit†	0.61 (0.84)		
$R_{\rm cryst}$ (%)	19.4		
$R_{\rm free}$‡ (%)	24.8		
No. of parameters	8607		
No. of restraints	7982		
Parameter to data ratio	3.5		
Average B factor (Å^2)	31.1		
Ramachandran plot (%)			
Residues in preferred regions	93		
Residues in allowed regions	7		
R.m.s.d.s			
Bond (Å)	0.006		
Angle (°)	1.1		
Dihedral (°)	18.5		
Planarity (Å)	0.02		

† FOM after density modification with non-crystallographic symmetry averaging is shown in parentheses. ‡ Randomly selected 5% of the reflections.

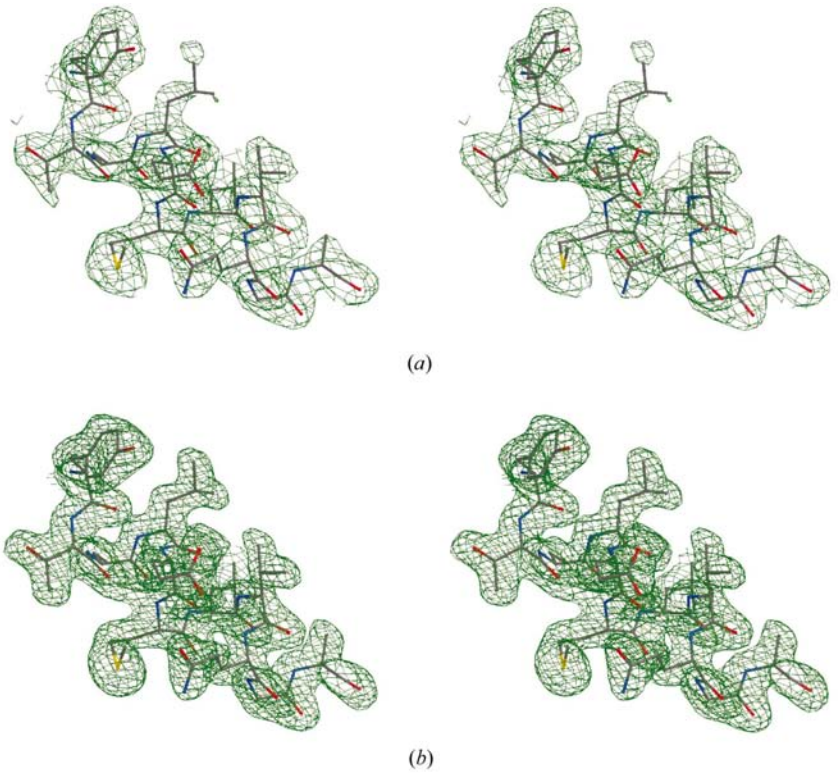

Figure 2
Stereoviews of electron-density maps. (*a*) Experimental map contoured at 1.2σ. (*b*) Final $|3F_o - 2F_c|$ map contoured at 1.5σ in the same area of the protein. The refined protein model is shown together with the maps.

cules of SycE in the asymmetric unit (AU), as expected from the theoretical V_M of 2.6 Å^3 Da^{-1} for a dimer in the AU. Phasing with *SHARP* (de La Fortelle & Bricogne, 1997) followed by density modification in *SOLOMON* (Abrahams & Leslie, 1996) resulted in an exceptionally clear electron-density map (Fig. 2*a*), which was manually traced using the program *O* (Jones *et al.*, 1991). The total time necessary to trace the protein backbone and put in all the amino-acid side chains was approximately 1.5 h, which allowed us to complete the first model while the data for the inflection and high-energy remote wavelengths were still being collected. Ultimately, the data for all three wavelengths were incorporated into the phasing solution. The essential details of the structure solution are summarized in Table 1.

2.6. Refinement

Residues 6–115 of the two SycE monomers were initially built into the experimental electron density. After several cycles of conjugate-gradient least-squares refinement with *SHELXL*97 (Sheldrick & Schneider, 1997) using the selenium-peak data in the resolution range 30–2.0 Å, two contiguous patches of electron density were interpreted as residues 128–130 followed by the two histidines of the C-terminal hexahistidine tag. The remainder of the His tag as well as most of the loop 116–127 appear to be disordered in the crystal. Additional N-terminal residues 3–5 were built after several refinement cycles; the final model included residues 3–132 for each monomer.

After iterative cycles of manual rebuilding followed by least-squares minimization in *SHELX*97, 226 water molecules and two imidazole molecules were added to the structure on the basis of difference density and standard proximity criteria. The final refinement (using data in the range 30–1.95 Å) was performed using 50 conjugate-gradient least-squares cycles with isotropic temperature-displacement factors for all of the atoms. Validation of the structure was performed with *WHATIF* (Vriend, 1990). Molecular graphics were generated with *GRASP* (Nicholls *et al.*, 1991), *BOBSCRIPT* (Esnouf, 1997), *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merrit & Murphy, 1994). The essential details of the refinement are provided in Table 1.

3. Results and discussion

3.1. Overall description of the structure

The fold of SycE (Fig. 3) does not closely resemble that of any known structure, according to the results of a structure-similarity search performed by the DALI server (Holm & Sander, 1993). However, it bears a vague resemblance to the 2α/6β 'half-sandwich' found in the cyanobacterial protein CyaY (PDB code 1ew4; DALI score 3.4). The SycE monomer is composed of a curved six-stranded antiparallel β-sheet that is wrapped around a central α-helix and flanked by two smaller α-helices (Fig. 3). The polypeptide begins with an N-terminal helix (α1). The backbone then adopts an extended conformation for a short distance before forming three strands of the antiparallel β-sheet (β1–β3) in succession. After β3, the backbone meanders to the opposite side of the protein in an extended conformation whereupon it forms helix α2. A loop connects α2 with the next β-strand (β4). β4 is followed immediately by β5, after which the backbone forms a single helical turn in the midst of the loop between β5 and helix α3. The ten residues that follow α3 are disordered in the crystal, but the last three residues of SycE are clearly discernible as the sixth strand of the β-sheet (β6). However, the biological significance of β6 is dubious because the ortholog of SycE in *Pseudomonas aeruginosa* lacks this part of the sequence (see below).

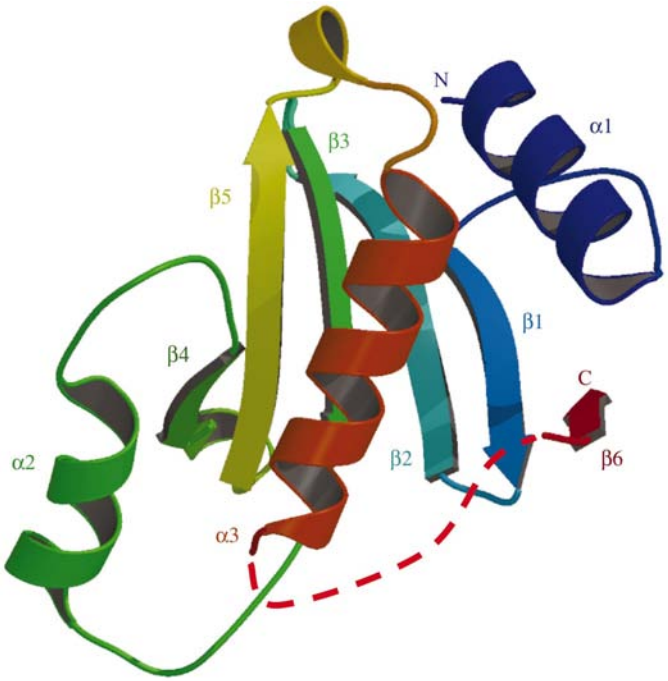

Figure 3
Schematic representation of a SycE monomer colored by secondary-structure procession. The disordered loop (residues 118–127) is represented by the broken red line.

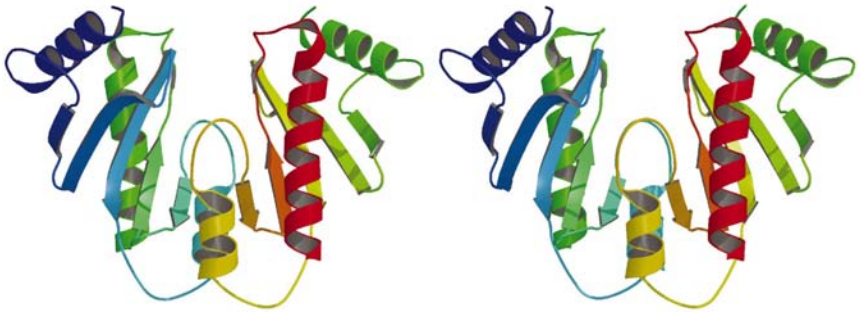

Figure 4
Stereoview of the SycE dimer colored by secondary-structure procession.

It has been suggested that many if not all of the type III secretion chaperones possess a C-terminal amphipathic α-helix (Wattiau *et al.*, 1996). Such a feature is indeed present in the structure of SycE (α3, residues 102–115). Most of the hydrophobic amino acids in α3 interact with the interior of the protein, whereas all of the polar and charged amino acids, with the exception of the buried Gln106, are exposed to the solvent.

3.2. The dimer interface

There is strong evidence that SycE exists as a stable homodimer in solution (Cheng & Schneewind, 1999). Accordingly, we believe that the tight dimer observed in the AU of the crystal (Fig. 4) is very likely to be biologically relevant. The SycE dimer interface is formed by several clusters of residues in the central part of the polypeptide (Fig. 5). Dimerization occludes $\sim$900 Å^2 per monomer, or about 10% of the solvent-accessible protein surface. Approximately 85% of the occluded surface, which includes residues His41, Pro42, Gln45, Leu47, Lys60, Glu61, Leu64, Ser65, Asn67, Ile68, Trp80, Val88, Trp90 and Arg92, is hydrophobic. A score of hydrogen bonds also link the monomers: Lys60 $N_\zeta\cdots O$ *Asn56*, Lys60 $N_\zeta\cdots O$ *Asp58*, Leu64 $N\cdots O$ *Trp80*, Ser65 $O_\gamma\cdots O$ *Trp80*, Ser65 $O_\gamma\cdots O_\varepsilon$ *Glu82*, Asn67 $O\cdots N_\varepsilon$ *Trp90*, Asn67 $O_\delta\cdots N_\delta$ *Asn67* and Arg92 $N_\eta\cdots O$ *Ser70* (residues from the other monomer are in italics). The compact fit of the interface suggests that the dimerization constant is relatively high, which correlates well with observations made during the purification of SycE.

3.3. The cavity

The structure of the SycE dimer reveals a groove formed by pairs of β-strands 93–99, loops 70–74, turns 41–42 and residues 102–106 of helix α3. The bottom of this groove opens into a deep bifurcated cavity formed by residues emanating from helix α2 and strand β4 (Fig. 6). The dimensions of the groove and the cavity are depicted schematically in Fig. 6(*b*); the total volume of the cavity is approximately 240 Å^3 (*VOIDOO*; Kleywegt & Jones, 1994). The walls of the groove and the cavity expose a variety of amino-acid side chains and some main-chain atoms, creating a surface that is roughly 55% hydrophobic. The cavity itself is positively charged owing to the unneutralized Arg92 residues in the main body of the fissure and the Lys75 residues located near the ends of its branches (Fig. 6*a*).

Although it is conceivable that the cavity is merely an evolutionary relic, it seems likely that selective pressure would

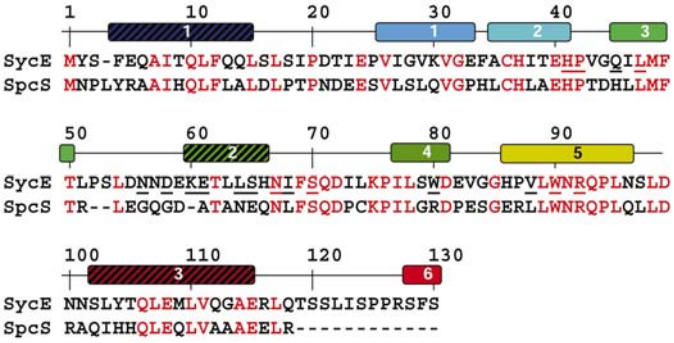

Figure 5
Sequence alignment of *Y. pestis* SycE and *P. aeruginosa* SpcS. Amino-acid sequences (in single-letter code) were aligned with the program *PILEUP* (University of Wisconsin Genetics Computer Group) using default parameters. The numbering scheme corresponds to the SycE sequence. Conserved residues are colored red and dimer interface residues are underlined. The locations of α-helices and β-strands in SycE are indicated above the sequence and colored in accord with Fig. 3.

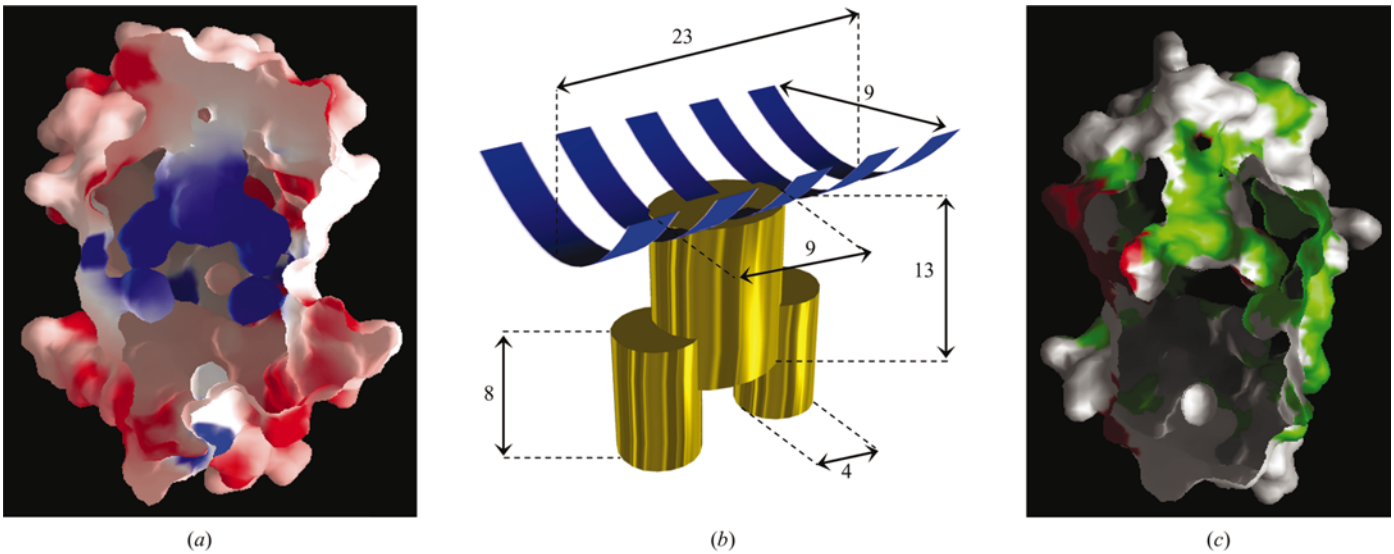

Figure 6
The cavity at the dimer interface. (*a*) A slice through the dimer, colored according to electrostatic potential; blue corresponds to positive potential and red to negative. (*b*) Three-dimensional diagram of the cavity, depicting the groove in the protein surface as blue ribbons and the cavity itself as brass cylinders. The approximate dimensions of every major element are indicated in angstroms. (*c*) A slice through the cavity, colored to indicate the residues in SycE that are conserved in SpcS. The conserved residues are colored red on one monomer and green on the other.

have eliminated such an energetically unfavorable feature as a large cleft on the dimer interface containing two unneutralized arginine residues if it served no useful purpose. Thus, we believe that the cavity must have some functional significance; either it plays a role in the interaction of SycE with YopE, as discussed below, or it is related to some hitherto undiscovered function of SycE.

3.4. Putative structural similarity between SycE and other type III secretion chaperones

The opportunistic pathogen *P. aeruginosa* also harbors a type III secretion system (Winstanley & Hart, 2001). One of the effector proteins that it exports, called exoenzyme S (ExoS), is a bifunctional GTPase activator (GAP) and ADP-ribosyltransferase (Barbieri, 2000). The GAP domain of ExoS, which selectively activates mammalian Rho-family GTPases, is structurally very similar to the GAP domain of YopE (Würtele *et al.*, 2001; Evdokimov *et al.*, 2002). The open reading frame immediately adjacent to ExoS, which was designated ORF1 by Yahr *et al.* (1995), encodes a small acidic protein whose amino-acid sequence shares 44% identity with that of SycE (Fig. 5). For convenience, we will refer to this putative secretion chaperone as SpcS (specific *Pseudomonas* chaperone S).

A high-confidence model of SpcS was constructed on the basis of the SycE coordinates using the SWISS-MODEL homology threading server (Guex & Peitsch, 1997). Most of the SpcS residues could be threaded onto the SycE structure without conflicts. SpcS loop 52–57 clearly adopts a different conformation than its SycE counterpart, but in general all of the structural elements of SycE also seem to be present in SpcS, with the exception of the disordered loop 118–127 and the last β-strand (β6), which are completely missing from the SpcS sequence. It is very likely that SpcS is also a homodimeric protein because most of the residues that comprise the dimer interface in SycE are conserved. The residues that form the positively charged cavity in SycE are also conserved in SpcS (Fig. 6*c*), underscoring its potential importance. Considering also that the amino-acid sequence near the N-terminus of ExoS bears some resemblance to that of the SycE-binding site in YopE, it seems highly probable that SpcS serves the same function for ExoS as SycE does for YopE.

We next asked whether the amino-acid sequences of the other known bacterial type III secretion chaperones are compatible with the fold of SycE. The specific proteins examined were SycT, SycH, SycN, YscB and LcrH from *Y. pestis*, SicA and SicP from *Salmonella typhimurium*, CesD and CesT from enteropathogenic *E. coli*, SpcU from *P. aeruginosa*, IpgC from *Shigella flexneri* and the *E. coli* flagellar chaperones FlgN, FliT, FliS and FliJ (Bennett & Hughes, 2000; Aizawa, 2001). The results of these threading experiments suggested that the C-terminal residues of SycH, SycT, YscB and SicP are also likely to form amphipathic α-helices. However, with the possible exception of SycT, it is not evident that these proteins adopt the same general fold as SycE and SpcS. Virtually no structural homology could be detected between SycE and any of the other type III secretion chaperones examined. Moreover, although the amino-acid residues involved in dimerization and the formation of the cavity in SycE are conserved in SpcS, neither of these signature motifs could be detected in any of the other type III secretion chaperones.

Table 2
Thermodynamic parameters describing the interaction of SycE with peptides derived from the N-terminus of YopE.

Peptide	pI_{theor}	K (10^5 M^{-1})	N	ΔH (kJ)
SP3487	10.9	1.6 (2)	1.00 (1)	−5.1 (2)
SP3485	8.5	1.2 (8)	0.80 (5)	−0.7 (2)
SP3461	5.8	0.8 (3)	0.88 (2)	−1.6 (2)
SP3486	4.9	0.7 (3)	0.63 (8)	−0.4 (1)

3.5. The interaction of SycE with YopE

The SycE-binding site in YopE has been localized by deletion analysis to amino acids 15–50; the GAP domain (residues 90–219) is dispensable for the interaction (Woestyn *et al.*, 1996). However, a fragment of YopE consisting of residues 1–90 accumulated in an insoluble form when it was overproduced in *E. coli* (data not shown) and the full-length effector is also poorly soluble (Cheng & Schneewind, 1999), making thermodynamic studies of the YopE–SycE interaction difficult. We therefore attempted to define the YopE determinant of the YopE–SycE interaction more precisely by studying the binding of synthetic peptides derived from the N-terminal 90 residues of YopE to SycE in solution using isothermal titration calorimetry. To avoid potential artifacts, wild-type SycE (*i.e.* with no affinity tags) was used for these experiments. Peptides corresponding to residues 1–33, 34–61, 62–90 and 15–50 of YopE were all readily soluble in aqueous buffers. The thermodynamic parameters describing the interaction of these peptides with SycE are summarized in Table 2. The peptide binding constants are roughly proportional to the theoretical isoelectric points of the individual peptides, which is not surprising in view of the fact that SycE is very acidic. Curiously, however, none of these peptides exhibited a high affinity for SycE in solution; their binding constants were observed to be in the range 10^4–10^5 M^{-1}, whereas the constant of YopE binding to SycE is reported to be of the order of 10^{10} M^{-1} (Cheng & Schneewind, 1999). It is particularly puzzling that the peptide corresponding to residues 15–50 of YopE failed to bind SycE, as Woestyn *et al.* (1996) detected a stable interaction between SycE and this fragment of YopE when it was fused to the adenylate cyclase domain of *Bordetella pertussis* cyclolysin. The reason for this discrepancy remains to be determined.

If each SycE dimer binds only one molecule of YopE, as suggested by the results of Cheng & Schneewind (1999), then the interaction of YopE with SycE must involve both SycE monomers in a fashion that precludes the binding of a second YopE molecule in a symmetrical manner. This, in turn, means that either the N-terminus of YopE wraps around the SycE dimer so that it contacts the same areas of both monomers

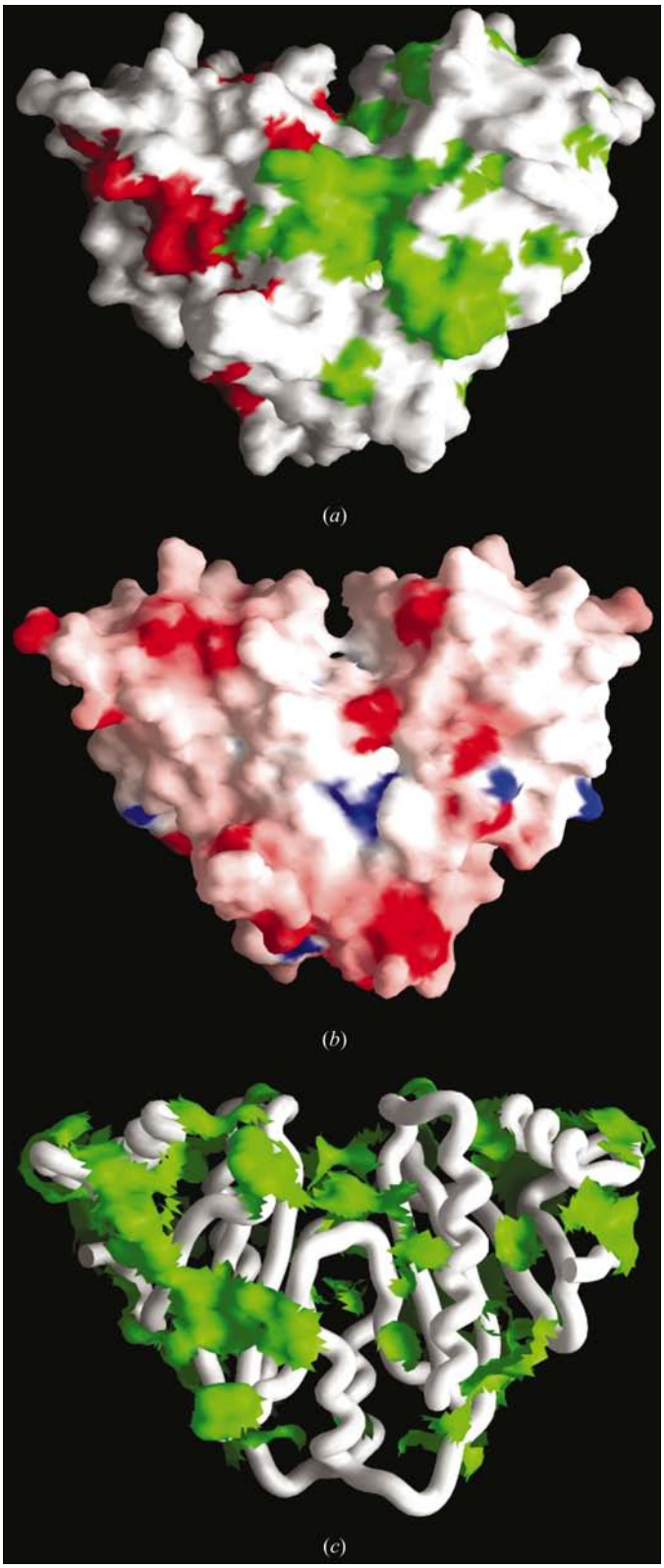

Figure 7
Characteristics of the SycE dimer surface. (*a*) Residues on the surface of SycE that are conserved in SpcS are colored red on one monomer and green on the other. (*b*) Molecular surface colored according to electrostatic potential; blue corresponds to positive potential and red to negative. (*c*) The solvent-exposed hydrophobic molecular surface is displayed in green, overlaying the C^{α} trace in white.

and/or that the binding interface involves the cavity, which is a unique interaction site owing to its location on the dyad axis of the SycE dimer.

Charge and hydrophobicity patterns on the dimer surface (Fig. 7*b* and 7*c*) do not immediately suggest any obvious interaction site(s). We reasoned that some insight into the potential binding site(s) on SycE might be gained by comparing its surface with that of SpcS, the counterpart of SycE in *P. aeruginosa*. The pattern of conserved residues outlines a broad stripe on both the front and the back of the dimer (Fig. 7*a*) and the residues that line the cavity are also highly conserved (Fig. 6*c*). The cavity is not likely to be the exclusive determinant of the YopE–SycE interaction because it is not large enough to accommodate more than just a few links of the polypeptide chain. However, it is attractive to imagine that a portion of the YopE polypeptide may be threaded through the groove at the opening of the SycE cavity in the YopE–SycE complex, perhaps with an acidic side penetrating into the positively charged cavity, because this would create the asymmetry that is needed to account for the observed stoichiometry of the interaction.

4. Conclusions

In this article, we present the crystal structure of the bacterial type III secretion chaperone SycE from *Y. pestis*. SycE possesses a novel fold with an unusual dimerization motif and an intriguing basic cavity that may participate in its interaction with YopE. The crystal structure of SycE allowed us to draw clear structural parallels between this chaperone from *Y. pestis* and its *P. aeruginosa* ortholog SpcS, but it is uncertain whether the other members of the extended family of type III secretion chaperones adopt the same general fold.

Note added in proof. After this manuscript was submitted for publication, the crystal structure of *Y. pseudotuberculosis* SycE was reported (Birtalan & Ghosh, 2001). Although the crystallization conditions, space group and unit-cell parameters are different, the two structures are virtually identical.

We wish to acknowledge Dr Zbigniew Dauter for providing the synchrotron beam time (beamline X9B of the National Synchrotron Light Source) used to determine the structure of SycE. Additionally, we are grateful to Suzanne Specht for assistance with peptide synthesis and purification, Dr Sergey Tarasov (Biophysics Resource) for performing the calorimetric experiments and for granting us access to the mass spectrometer and Dr Alex Wlodawer for critical reading of the manuscript.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Aizawa, S.-I. (2001). *FEMS Microbiol. Lett.* **202**, 157–164.
Anderson, D. M. & Schneewind, O. (1997). *Science*, **278**, 1140–1143.
Barbieri, J. T. (2000). *Int. J. Med. Microbiol.* **290**, 381–387.
Bennett, J. C. Q. & Hughes, C. (2000). *Trends Microbiol.* **8**, 202–204.
Birtalan, S. & Ghosh, P. (2001). *Nature Struct. Biol.* **8**, 974–978.

Cheng, L. W., Anderson, D. M. & Schneewind, O. (1997). *Mol. Microbiol.* **24**, 757–765.
Cheng, L. W. & Schneewind, O. (1999). *J. Biol. Chem.* **274**, 22102–22108.
Cheng, L. W. & Schneewind, O. (2000). *Trends Microbiol.* **8**, 214–220.
Cornelis, G. R. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 8778–8783.
Cornelis, G. R., Boland, A., Boyd, A. P., Geuijen, C., Iriarte, M., Neyt, C., Sory, M.-P. & Stainier, I. (1998). *Microbiol. Mol. Biol. Rev.* **62**, 1315–1352.
Cornelis, G. R. & Van Gijsegem, F. (2000). *Annu. Rev. Microbiol.* **54**, 735–774.
Donnenberg, M. S. (2000). *Nature (London)*, **406**, 768–774.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
Evdokimov, A. G., Anderson, D. E., Routzahn, K. M. & Waugh, D. S. (2000). *Acta Cryst.* D**56**, 1676–1679.
Evdokimov, A. G., Tropea, J. E., Routzahn, K. M. & Waugh, D. S. (2002). *Protein Sci.* **11**, 401–408.
Forsberg, A. & Wolf-Watz, H. (1990). *J. Bacteriol.* **172**, 1547–1555.
Guex, N. & Peitsch, M. C. (1997). *Electrophoresis*, **18**, 2714–2723.
Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kapust, R. B. & Waugh, D. S. (2000). *Protein Expr. Purif.* **19**, 312–318.
Kleywegt, G. J. & Jones, T. A. (1994). *Acta Cryst.* D**50**, 178–185.
Koster, M., Bitter, W. & Tommassen, J. (2000). *Int. J. Med. Microbiol.* **290**, 325–331.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Lee, V. T. & Schneewind, O. (2001). *Genes Dev.* **15**, 1725–1752.
Lloyd, S. A., Forsberg, A., Wolf-Watz, H. & Francis, M. S. (2001). *Trends Microbiol.* **9**, 367–371.
Merrit, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perry, R. D., Straley, S. C., Fetherston, J. D., Rose, D. J., Gregor, J. & Blattner, F. R. (1998). *Infect. Immun.* **66**, 4611–4623.
Schesser, K., Frithz-Lindsten, E. & Wolf-Watz, H. (1996). *J. Bacteriol.* **178**, 7227–7233.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Sory, M.-P., Boland, A., Lambermont, I. & Cornelis, G. R. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 11998–12002.
Usón, I. & Sheldrick, G. M. (1999). *Curr. Opin. Struct. Biol.* **9**, 643–648.
Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.
Wattiau, P., Bernier, B., Deslée, P., Michelis, T. & Cornelis, G. R. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 10493–10497.
Wattiau, P. & Cornelis, G. R. (1993). *Mol. Microbiol.* **8**, 123–131.
Wattiau, P., Woestyn, S. & Cornelis, G. R. (1996). *Mol. Microbiol.* **20**, 255–262.
Winstanley, C. & Hart, C. A. (2001). *J. Med. Microbiol.* **50**, 116–126.
Woestyn, S., Sory, M.-P., Boland, A., Lequenne, O. & Cornelis, G. R. (1996). *Mol. Microbiol.* **20**, 1261–1271.
Würtele, M., Renault, L., Barbieri, J. T., Wittinghofer, A. & Wolf, E. (2001). *FEBS Lett.* **491**, 26–29.
Yahr, T. L., Hovey, A. K., Kulich, S. M. & Frank, D. W. (1995). *J. Bacteriol.* **177**, 1169–1178.