Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Characterization of four orthologs of stringent starvation protein A

Michelle Andrykovitch,[a] Karen M. Routzahn,[a] Mi Li,[b] Yijun Gu,[a] David S. Waugh[a] and Xinhua Ji[a]*

[a]Macromolecular Crystallography Laboratory, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA, and [b]Intramural Research Support Program, SAIC-Frederick, MD 21702, USA

Correspondence e-mail: jix@ncifcrf.gov

Received 9 December 2002
Accepted 11 March 2003

Orthologous proteins can be beneficial for X-ray crystallographic studies when a protein from an organism of choice fails to crystallize or the crystals are not suitable for structure determination. Their amino-acid sequences should be similar enough that they will share the same fold, but different enough so that they may crystallize under alternative conditions and diffract to higher resolution. This multi-species approach was employed to obtain diffraction-quality crystals of the RNA polymerase (RNAP) associated stringent starvation protein A (SspA). Although *Escherichia coli* SspA could be crystallized, the crystals failed to diffract well enough for structure determination. Therefore, SspA proteins from *Yersinia pestis*, *Vibrio cholerae* and *Pseudomonas aeruginosa* were cloned, expressed, purified and subjected to crystallization trials. The *V. cholerae* SspA protein failed to crystallize under any conditions tested and the *P. aeruginosa* SspA protein did not form crystals suitable for data collection. On the other hand, *Y. pestis* SspA crystallized readily and the crystals diffracted to 2.0 Å.

1. Introduction

Orthologs are genes in different organisms that are direct evolutionary counterparts of each other (Fitch, 1970, 2000; Mirny & Gelfand, 2002). Orthologous proteins have similar or identical amino-acid sequences and are presumed to have the same function (Makarova *et al.*, 1999; Gelfand *et al.*, 2000; Tatusov *et al.*, 2000). Certain amino-acid residues in the polypeptide chain may be conserved in all orthologous proteins, while in other places they may differ. Invariant amino-acid residues commonly occur at crucial positions along the polypeptide chain that are important for the function of the protein, including turns or bends in the chain, the hydrophobic core residues, cross-linking points between loops in the tertiary structure and residues that comprise the catalytic sites of enzymes or binding sites for prosthetic groups (Lehninger *et al.*, 1983). The degree of similarity between amino-acid sequences of orthologous proteins from different species is correlated with the evolutionary relationship between them. Orthologs with similar amino-acid sequences are presumed to also have similar tertiary structures.

Transcription is the initial and crucial control point in the process of gene expression. During transcription, one strand of double-stranded DNA serves as a template on which a growing RNA strand is manufactured one nucleotide at a time. Transcription errors have been linked to defects in cell-cycle control, apoptosis and immune response. Clearly, a complete understanding of transcription-control mechanisms is critical to our understanding of genetic diseases and for the control of pathogens. RNA molecules are synthesized by RNA

polymerase (RNAP), which is activated and guided by a variety of transcription factors. Interactions of core RNAP with nucleic acids and accessory factors govern its course through a progression of structurally distinct intermediates in a transcription cycle (Fu *et al.*, 1999; Mooney & Landick, 1999; Zhang *et al.*, 1999).

Stringent starvation protein A (SspA) is a bacterial RNAP-associated protein (Reeh *et al.*, 1976; Ishihama & Saitoh, 1979) that is induced during stationary phase growth, owing to nutrient starvation or upon phage λ infection (Williams *et al.*, 1994). SspA is involved in the expression of at least 11 proteins in *Escherichia coli* (Williams *et al.*, 1994). In addition, SspA has been shown to be essential for the lytic growth cycle of bacteriophage P1 (Williams *et al.*, 1994). SspA orthologs in *Neisseria gonorrhoeae* and *Francisella novicida* have also been shown to affect the expression of genes implicated in pathogenesis (De Reuse & Taha, 1997; Baron & Nano, 1998). However, the specific role of SspA in association with RNAP and its cellular function have yet to be elucidated. The goal of this research was to obtain diffraction-quality crystals of SspA for the purpose of determining its three-dimensional structure and gaining insight into its function.

Obtaining diffraction-quality crystals is the rate-limiting step in macromolecular X-ray crystallography. Some proteins never crystallize, or if they do crystallize the crystals are not suitable for crystal structure determination. Under these conditions, it can be beneficial to attempt to crystallize an orthologous protein instead. Often, an orthologous protein may crystallize under different conditions, thus offering another prospect of solving the protein structure. The structure of an ortholog can then be used to predict the structure of the protein of interest or to guide protein-engineering experiments. An engineered protein of interest may crystallize and the crystals may diffract to desired resolution.

In this study, we endeavored to overexpress, purify and crystallize orthologous SspA proteins from *E. coli*, *Yersinia pestis*, *Vibrio cholerae* and *Pseudomonas aeruginosa*. The *E. coli* protein was the first choice for protein production. However, although crystals of *E. coli* SspA could be obtained, they did not diffract X-rays well. SspA proteins from *Y. pestis*, *V. cholerae* and *P. aeruginosa* were chosen because their amino-acid sequences are 83, 72 and 53% identical and 90, 86 and 70% similar to the *E. coli* protein, respectively (Fig. 1). This level of sequence identity ensures that all four proteins will share the same tertiary structure. On the other hand, their amino-acid sequences are sufficiently different that they may crystallize under different conditions and diffract to higher resolution than the *E. coli* SspA.

```
           10         20         30         40         50         60
Ec  MAVAANKRSV MTLFSGPTDI YSHQVRIVLA EKGVSFEIEH VEKDNPPQDL IDLNPNQSVP
Yp  MAVAANKRSV MTLFSGPTDI FSHQVRIVLA EKGVSVEIEQ VEADNLPQDL IDLNPYRTVP
Vc  MAVAANKRSV MTLFSSASDM YSHQVRIVLA EKGVSFEVEL VDENNLPAEL IELNPYKTVP
Pa  MA-SINK--- LTCYSDPADH YCHRIRLVLA EKNVAVEVID IEAGRCPPML AEVNPYGSVP

           70         80         90        100        110        120
Ec  TLVDRELTLW ESRIIMEYLD ERFPHPPLMP VYPVARGESR LYMHRIEKDW YTLMNTIING
Yp  TLVDRELTLY ESRIIMEYLD ERFPHPPLMP VYPVARGSSR LMMHRIEHDW YSLLYKIEQG
Vc  TLVDRELALY DSKIIMEYLD ERFPHPPLMP VYPVARGNSR LMIYRIERNW YSLAEKVVNG
Pa  TLVDRDLALY EPTVVMEYLD ERYPHPPLLP DYPVKRANSR LLIHRIQRDW CSLVDRILDA

           130        140        150        160        170        180
Ec  SASEAD--AA RKQLREELLA IAPVFGQKPY FLSDEFSLVD CYLAPLLWRL PQLGIEFSGP
Yp  NAQEAE--AA RKQLREELLS IAPVFNETPF FMSEEFSLVD CYLAPLLWRL PVLGIEFTGA
Vc  SPEVAE--NA RNKLRNDLLT LGPVFAEFEY FMSEEFSLVD CYLAPLLWRL PVLGIDLIGP
Pa  RSKEAERVQA RKELRESLTG VSPLFADKAY FLSEDFSLVD CCLLPILWRL PLLGIELPRQ

           190        200        210
Ec  GAKELKGYMT RVFERDSFLA SLTEAEREMR LG-RS
Yp  GSKELKGYMT RVFERDAFLA SLTEAEREMH LKTRS
Vc  GSKELKVYMN RVFERDSFLA SLTEAEREMR LAR--
Pa  -AKPLLDYME RGFARESFRA SLSSVERDMR -----
```

Figure 1
Manual alignment of *E. coli* (Ec), *Y. pestis* (Yp), *V. cholerae* (Vc) and *P. aeruginosa* (Pa) SspA amino-acid sequences. The strictly conserved residues are colored red.

2. Materials and methods

2.1. Cloning, expression, purification and crystallization of *E. coli* SspA

The DNA fragment encoding *sspAB* was amplified by the polymerase chain reaction (PCR) from *E. coli* chromosomal DNA using the DNA oligomers DJ259b (5′ TAG CAG **GGA TCC** ATG GCT GTC GCT GCC AAC AAA CGT TCG 3′) and DJ260 (5′ TTT ACT ACT **AAG CTT** TTA CTT CAC AAC GCG TAA TGC CGG TCG A 3′). The plasmid pDJ706 was used to express the recombinant *E. coli* SspA protein, which encodes *sspAB* cloned under the control of a *tac* promoter in the *Bam*HI–*Hin*dIII sites of pQE30 (Qiagen). The resulting plasmid produces SspA with an N-terminal polyhistidine tag.

E. coli SspA was expressed in strain DJ706, which consists of *E. coli* MG1655 Δ*lacX-74 mal::lacI*^{q} cells harboring pDJ706. An overnight culture of DJ706, grown at 303 K in Luria broth (LB) medium containing 100 µg ml^{-1} ampicillin, was diluted 1:100 in 1 l LB and grown at 303 K with shaking. The 1 l cultures were harvested at an OD_{600} of 1.0. No induction of SspA expression was required. The cell pellet was stored at 193 K.

The cell pellet was thawed on ice. The cell-lysis buffer consisted of 20 m*M* Tris–HCl pH 7.4, 0.2 *M* NaCl and the 'complete' protease-inhibitor cocktail (Roche). 10 ml of lysis buffer per gram of cells was stirred at 277 K. After 10 min, the mixture was lysed with an APV Model G1000 Gaulin homogenizer. The supernatant was collected after centrifugation at 37 000*g* for 15 min at 277 K. The supernatant was collected and stirred at 277 K for 5 min with a final concentration of 0.1% polyethelenimine (PEI). The precipitate was removed by centrifugation at 37 000*g* for 15 min at 277 K.

A BioCAD/Sprint chromatography workstation (Applied Biosystems) and an Advantec fraction collector were used to carry out chromatographic purification of the SspA proteins at room temperature. Taking advantage of the N-terminal polyhistidine tag, immobilized metal-affinity chromatography (IMAC) on Ni–NTA resin was employed as the initial purification step. After PEI precipitation, the supernatant was dialyzed overnight against Ni–NTA buffer (50 m*M* NaH_2PO_4 pH 8.0, 300 m*M* NaCl) before being loaded onto an XK 26/20 column (Amersham Biosciences) containing 37 ml of Ni–NTA

Superflow resin (Qiagen). The column was run according to the manufacturer's specifications. After elution, the sample was adjusted to 2 m*M* DTT and 2 m*M* EDTA and then precipitated with a final concentration of 75% ammonium sulfate. The precipitate was pelleted by centrifugation at 37 000*g* for 15 min at 277 K and then resuspended in the smallest possible volume of 20 m*M* Tris–HCl pH 8.0, 0.2 *M* NaCl and 5 m*M* DTT. Approximately 3 ml at a time was loaded onto a Sephacryl S-100 26/60 sizing column (Amersham Biosciences), equilibrated with the same buffer, at a flow rate of 1 ml min^{-1}. The sample was then concentrated in the same buffer with a stirred cell (Amicon) to 10 mg ml^{-1} for crystallization.

Approximately 1300 different crystallization conditions were explored using the hanging-drop vapor-diffusion technique in an effort to crystallize *E. coli* SspA.

2.2. Cloning, expression, purification and crystallization of SspA proteins from *Y. pestis*, *V. cholerae* and *P. aeruginosa*

ORFs encoding the *Y. pestis*, *V. cholerae* and *P. aeruginosa* SspA proteins were amplified from genomic DNA by PCR and cloned as maltose-binding protein (MBP) fusion partners using the Gateway cloning system (Fox & Waugh, 2003). A recognition site for tobacco etch virus (TEV) protease was added to the N-terminus of the SspA ORFs to facilitate their separation from MBP. The nucleotide sequences of the plasmids containing the insert were confirmed before proceeding further. After the sequences had been verified, the DNA was transformed by electroporation into BL21 cells containing the RIL plasmid (Stratagene), which makes tRNAs for rare arginine, isoleucine and leucine codons. The transformed cells were spread on LB agar plates that contained 125 µg ml^{-1} ampicillin and 30 µg ml^{-1} chloramphenicol. Single colonies from these plates were used to inoculate liquid cultures.

10 ml of a 100 ml overnight culture grown in LB containing 100 µg ml^{-1} ampicillin and 30 µg ml^{-1} chloramphenicol was used to inoculate six 1 l cultures of the same medium. Both the 100 ml overnight and 1 l cultures were incubated at 310 K with shaking. The 1 l cultures were induced with 1 m*M* IPTG when the cells reached an optical density OD_{600} of 0.3. Approximately 3–4 h later the cells were pelleted by centrifugation and stored at 193 K. Cell lysis and PEI precipitation were performed as described above.

Following PEI precipitation, the supernatant was loaded onto an XK 50 column (Amersham Biosciences) containing 98 ml amylose resin (New England BioLabs). Amylose-affinity chromatography was performed according to the instructions provided by New England Biolabs. After elution, the sample was concentrated by precipitation with a final concentration of 75% ammonium sulfate. The precipitate was pelleted by centrifugation at 37 000*g* for 15 min at 277 K and then resuspended in 100 ml of 20 m*M* Tris–HCl pH 7.4, 0.2 *M* NaCl, 10 m*M* β-mercaptoethanol, 1 m*M* EDTA.

Approximately 1 mg of TEV protease was used to cleave 150 mg of fusion protein (Kapust *et al.*, 2002). The TEV protease was added directly to the sample along with 1 m*M* DTT and dialyzed overnight against 20 m*M* Tris–HCl pH 7.4, 0.2 *M* NaCl, 10 m*M* β-mercaptoethanol and 1 m*M* EDTA at room temperature. After dialysis, the sample was reloaded on the amylose column as before. However, this time the flow-through was collected. Next, the sample was precipitated with a final concentration of 75% ammonium sulfate. The precipitate was pelleted by centrifugation at 37 000*g* for 15 min at 277 K and then resuspended in the smallest possible volume of 20 m*M* Tris–HCl pH 8.0, 0.2 *M* NaCl and 5 m*M* DTT. Approximately 3 ml at a time was loaded onto a Sephacryl S-100 26/60 sizing column (Amersham Biosciences), equilibrated with the same buffer, at a flow rate of 1 ml min^{-1}. The sample was then concentrated in the same buffer with a stirred cell (Amicon) to 10 mg ml^{-1}.

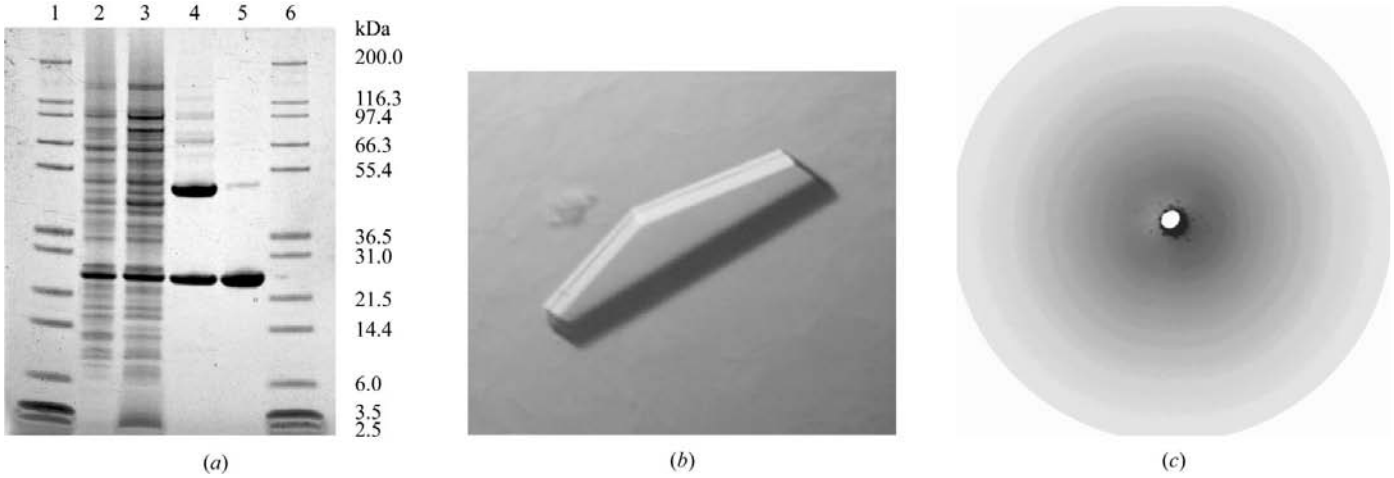

Figure 2
Purification, crystallization and X-ray diffraction of *E. coli* SspA. (*a*) Purification steps on a NuPAGE 4–12% bis-tris gel stained with SimplyBlue (Invitrogen). Lanes 1 and 6 are Mark 12 markers with molecular weights (kDa) indicated on the right-hand side. Lanes 2–5 show the soluble components of the cell lysate after homogenization, PEI precipitation, Ni–NTA (IMAC) chromatography and S-100 size-exclusion chromatography, respectively. (*b*) A typical crystal with dimensions of 0.3 × 0.2 × 0.05 mm. (*c*) The X-ray diffraction pattern of the crystal at ∼17 Å resolution.

Crystallization trials were performed at room temperature with both commercial and custom screening kits. Each kit generally contains 48 conditions based on many successful crystallization experiments (Gilliland *et al.*, 1994). Approximately 600, 1500 and 1000 different crystallization conditions were explored using the hanging-drop vapor-diffusion technique to crystallize *Y. pestis*, *V. cholerae* and *P. aeruginosa* SspA proteins, respectively.

2.3. X-ray diffraction of *E. coli* and *Y. pestis* SspA crystals

The *Y. pestis* SspA and *E. coli* SspA crystals were tested for X-ray diffraction with an in-house MAR 345 image plate mounted on a Rigaku rotating-anode generator operated at 50 kV and 100 mA. The *Y. pestis* crystals were also tested and a native data set at ∼2.0 Å was collected with an ADSC Quantum 4 CCD detector at beamline X9B of the National Synchrotron Light Source, Brookhaven National Laboratory. The crystals were flash-frozen and maintained at 100 K for experiments carried out with both in-house and synchrotron facilities. The cryoprotectant for the *E. coli* SspA crystals was composed of 0.1 *M* Na HEPES pH 7.5, 0.8 *M* potassium/sodium tartrate and 10% MPD. The mother liquor plus 10% MPD was used as a cryoprotectant for the *Y. pestis* SspA crystals.

3. Results and discussion

3.1. *E. coli SspA*

After an initial purification step by IMAC (Fig. 2*a*), most contaminants were removed. However, in addition to monomeric SspA, another prominent band with the approximate mobility expected for dimeric SspA was also observed (Fig. 2*a*, lane 4). The identity of this protein was not determined, but it is unlikely to correspond to oxidatively cross-linked dimers of SspA because the samples were heated in the presence of the reducing agent 2-mercaptoethanol prior to SDS–PAGE and the single cysteine in SspA is not solvent-accessible (data not shown). A translational read-through product can be ruled out because there is another in-frame termination codon located just a short distance downstream from the end of the SspA ORF. A third possibility is that the larger band results from a translational frameshifting event that gives rise to an SspA–SspB fusion protein, although this phenomenon has not been noted previously. In any case, it is curious that this larger protein appears to be captured with much greater efficiency on

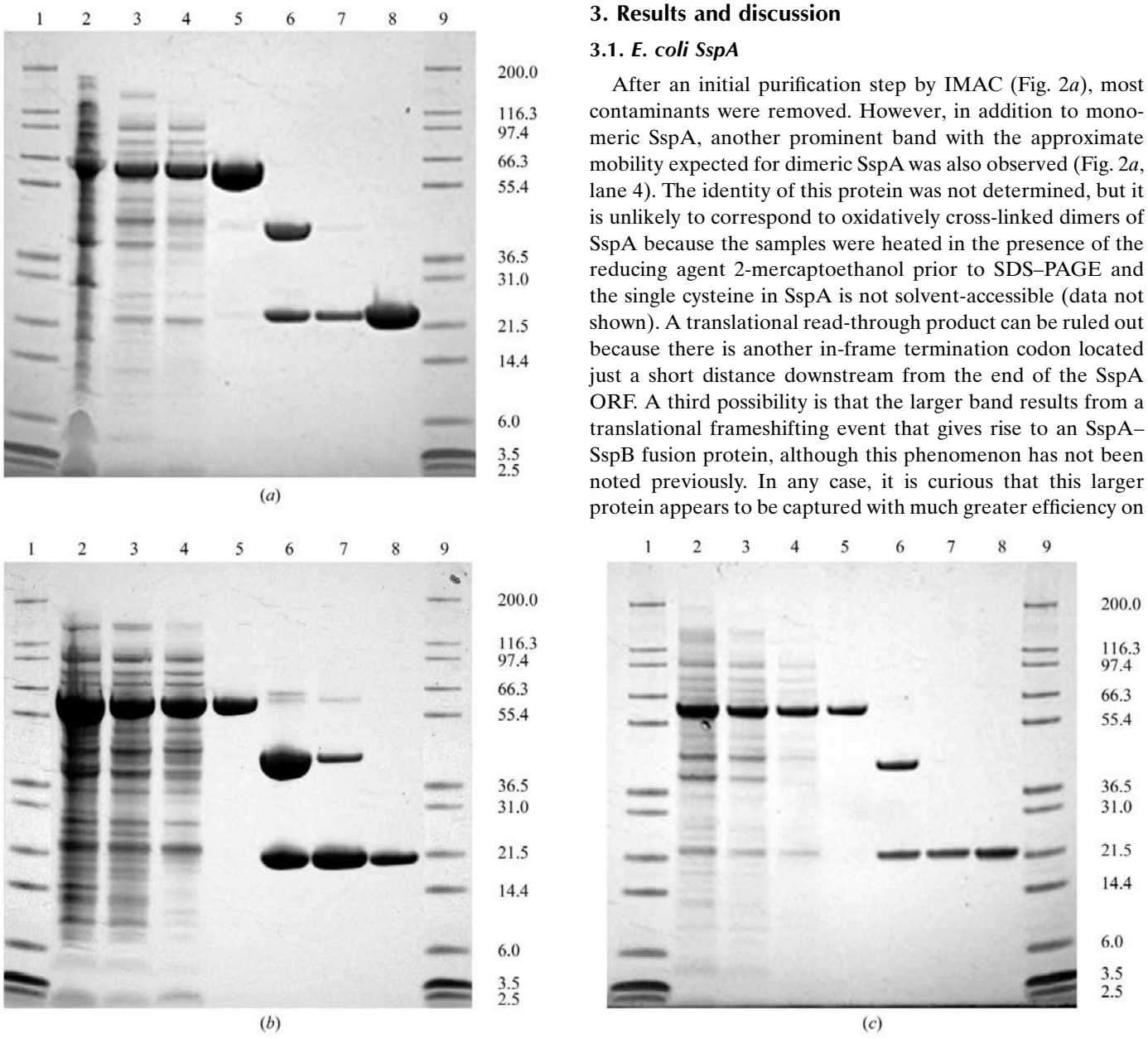

Figure 3
Purification of SspA proteins from *Y. pestis* (*a*), *V. cholerae* (*b*) and *P. aeruginosa* (*c*). Lanes 1 and 9 are the Mark 12 markers (Invitrogen) with molecular weights (kDa) indicated on the right-hand side. Lane 2 shows the total intracellular protein after homogenization. Lanes 3–8 show the soluble components after homogenization, PEI precipitation, the first amylose column, cleavage with TEV protease, the second amylose column and the S-100 column, respectively. Two bands are visible in lane 6, an MBP band at approximately 42 kDa and an SspA band at 22.5 kDa.

the Ni–NTA resin than is the monomeric His-tagged SspA protein. Following IMAC, the *E. coli* SspA was further purified by size-exclusion chromatography to separate the monomeric protein from the remaining contaminants (lane 5). Only a trace amount of the unidentified protein (above) was detected by SDS–PAGE in the final preparation of SspA.

After size-exclusion chromatography, the pure SspA was concentrated to 10 mg ml^{-1} for crystallization trials. Crystals of *E. coli* SspA were grown by the hanging-drop vapor-diffusion technique. The best crystallization condition identified was 0.1 *M* Na HEPES pH 7.5, 0.4 *M* potassium/sodium tartrate. This condition was obtained from Crystal Screen Lite reagent No. 29 (Hampton Research). The crystals grew over 4–10 d from a 1:1 mixture of crystallization solution and concentrated protein, ultimately reaching dimensions of approximately 0.3 × 0.2 × 0.005 mm (Fig. 2*b*). These crystals were frozen in 0.1 *M* Na HEPES pH 7.5, 0.8 *M* potassium/sodium tartrate with 10% 2-methyl-2,4-pentanediol (MPD) as a cryoprotectant and tested for X-ray diffraction. Unfortunately, the best diffraction observed was in the range of 17 Å resolution (Fig. 2*c*).

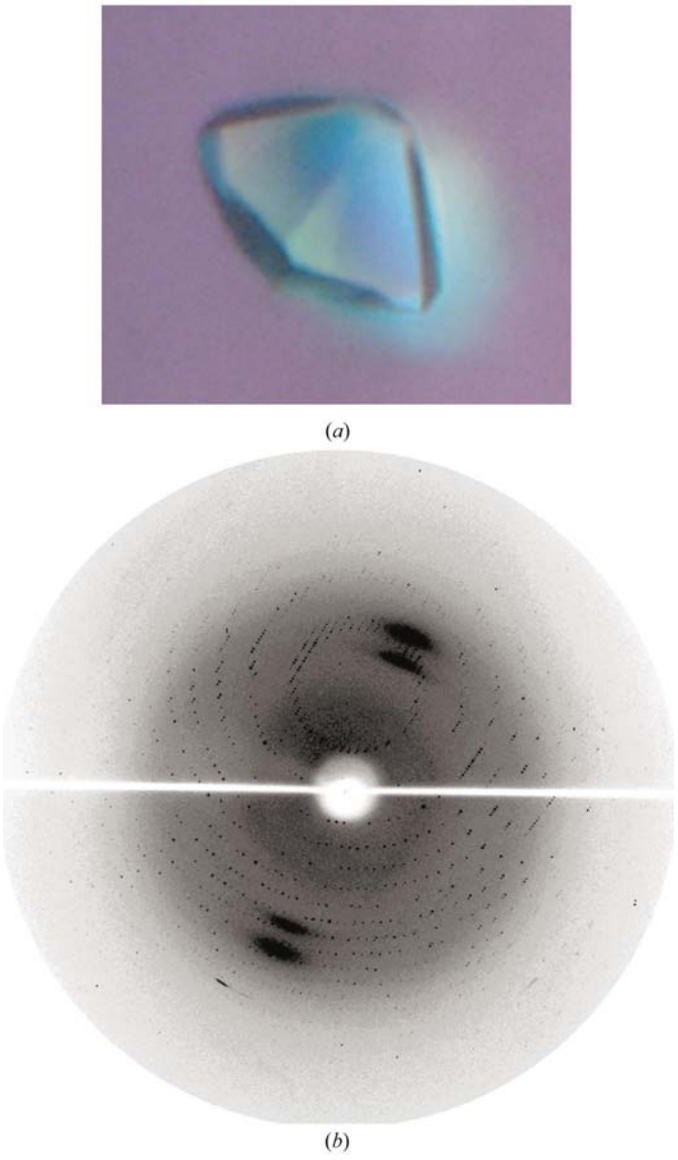

Figure 4
X-ray diffraction-quality crystals. (*a*) A typical crystal of *Y. pestis* SspA with dimensions of 0.1 × 0.1 × 0.1 mm. (*b*) X-ray diffraction at ∼2.0 Å of the crystal at the National Synchrotron Light Source.

3.2. *Y. pestis*, *V. cholerae* and *P. aeruginosa* SspA orthologs

ORFs encoding the *Y. pestis*, *V. cholerae* and *P. aeruginosa* SspA proteins were first cloned and expressed in the same manner as the *E. coli* SspA. However, the yield of all three proteins was extremely poor. Therefore, they were subsequently produced as MBP-fusion proteins instead, using the Gateway recombinational cloning system (Invitrogen). In addition to its utility as an affinity tag, MBP has also been observed to increase the yield and solubility of its fusion partners (Pryor & Leiting, 1997; Kapust & Waugh, 2000). The MBP–SspA fusion proteins were engineered so that they could be cleaved by TEV protease to yield SspA proteins with just a single non-native glycine residue at their N-termini. It is possible that the presence of the non-removable polyhistidine tag on *E. coli* SspA contributed to the poor diffraction quality of the crystals. By using a removable MBP tag instead for the production of the other SspA orthologs, we hoped to avoid this potential problem.

The same protocol was used to purify the *Y. pestis*, *V. cholerae* and *P. aeruginosa* SspA proteins (Fig. 3). The first step was amylose-affinity chromatography. All three of the MBP–SspA fusion proteins were nearly pure after they were eluted from the amylose column (Fig. 3, lane 5). Next, the fusion proteins were digested with TEV protease (Kapust *et al.*, 2002) to separate the MBP from the SspA proteins. The reactions proceeded essentially to completion in all cases and no unexpected digestion products were observed (Fig. 3, lane 6). Although the MBP probably could have been separated from the monomeric SspA proteins by size-exclusion chromatography at this stage, to ensure optimum resolution another amylose column was first used to absorb most of the free MBP in the sample (Fig. 3, lane 7). Following size-exclusion chromatography (Fig. 3, lane 8), the samples were concentrated to approximately 10 mg ml^{-1} for crystallization trials.

Crystals of *Y. pestis* SspA were grown by the hanging-drop vapor-diffusion method at 291 K in 0.2 *M* diammonium hydrogen citrate, 20%(*w*/*v*) PEG 3350 (Hampton Research PEG/Ion Screen Reagent No. 48). The protein concentration in the drop was approximately 5 mg ml^{-1}. The crystals grew to dimensions of 0.1 × 0.1 × 0.1 mm in 1 d (Fig. 4*a*). A large amount of precipitation also formed in the drop (not shown). The cryoprotectant for the crystals contained the mother liquor and 10% MPD. The highest resolution achieved was 2.0 Å with synchrotron radiation (Fig. 4*b*). No crystals were obtained for *V. cholerae* SspA under any of the conditions tested and *P. aeruginosa* SspA only yielded clusters of microcrystals.

3.3. Obtaining diffraction-quality crystals of SspA

While the cloning, expression and purification procedures were identical for *Y. pestis*, *V. cholerae* and *P. aeruginosa* SspA orthologs, the proteins behaved very differently in crystallization trials. The *Y. pestis* SspA protein produced beautiful crystals (Fig. 4*a*) that were easily manipulated and diffracted to ∼2.4 Å on an in-house X-ray diffractometer and to ∼2.0 Å (Fig. 4*b*) at synchrotron beamline X9B of the National Synchrotron Light Source, Brookhaven National Laboratory. Native data were collected at the latter facility and the crystal structure determination is in progress. *P. aeruginosa* SspA did form some crystals, but they were very small, irregular in shape and grew in clusters, making them useless for single-crystal X-ray diffraction. The *V. cholerae* SspA failed to crystallize under any of the conditions tested. The *E. coli* SspA was cloned and purified with a different protocol. Although *E. coli* SspA crystallized, the crystals diffracted very poorly (Fig. 2).

At this time, it is not possible to determine exactly why the *V. cholerae* SspA did not crystallize, why the *E. coli* SspA crystals diffracted so poorly and why the *P. aeruginosa* SspA crystals did not grow well. It is believed that crystallization requires the protein molecules to make critical protein–protein interactions for nucleation. Many factors can negatively influence nucleation. Replacing the N-terminal polyhistidine tag with the MBP tag may result in better quality crystals for *E. coli* SspA, which remains to be tested. By using a multi-species approach (four targets from the SspA family), we ultimately obtained diffraction-quality crystals of one SspA ortholog. Additionally, we have four purified SspA proteins for further functional studies. Once the structure of *Y. pestis* SspA has been determined, the three-dimensional structures of the other SspA proteins can be modeled as discussed above.

We thank Ms Anne-Marie Hansen and Dr Ding J. Jin for the *E. coli* SspA clone, insightful discussions and critical reading of the manuscript. We also thank Dr Zbigniew Dauter for assistance during the X-ray diffraction experiment at the National Synchrotron Light Source.

References

Baron, G. S. & Nano, F. E. (1998). *Mol. Microbiol.* **29**, 247–259.
De Reuse, H. & Taha, M. K. (1997). *Res. Microbiol.* **148**, 289–303.
Fitch, W. M. (1970). *Syst. Zool.* **19**, 99–113.
Fitch, W. M. (2000). *Trends Genet.* **16**, 227–231.
Fox, J. D. & Waugh, D. S. (2003). *Methods Mol. Biol.* **205**, 99–117.
Fu, J., Gnatt, A. L., Bushnell, D. A., Jensen, G. J., Thompson, N. E., Burgess, R. R., David, P. R. & Kornberg, R. D. (1999). *Cell*, **98**, 799–810.
Gelfand, M. S., Koonin, E. V. & Mironov, A. A. (2000). *Nucleic Acids Res.* **28**, 695–705.
Gilliland, G. L., Tung, M., Blakeslee, D. M. & Ladner, J. (1994). *Acta Cryst.* D**50**, 408–413.
Ishihama, A. & Saitoh, T. (1979). *J. Mol. Biol.* **129**, 517–530.
Kapust, R. B., Tözsér, J., Copeland, T. D. & Waugh, D. S. (2002). *Biochem. Biophys. Res. Commun.* **294**, 949–955.
Kapust, R. B. & Waugh, D. S. (2000). *Protein Expr. Purif.* **19**, 312–318.
Lehninger, A. L., Nelson, D. L. & Cox, M. M. (1983). *Principles of Biochemistry*, edited by V. Neal, pp. 134–159. New York: Worth Publishers.
Makarova, K. S., Aravind, L., Galperin, M. Y., Grishin, N. V., Tatusov, R. L., Wolf, Y. I. & Koonin, E. V. (1999). *Genome Res.* **9**, 608–628.
Mirny, L. A. & Gelfand, M. S. (2002). *J. Mol. Biol.* **321**, 7–20.
Mooney, R. A. & Landick, R. (1999). *Cell*, **98**, 687–690.
Pryor, K. D. & Leiting, B. (1997). *Protein Expr. Purif.* **10**, 309–319.
Reeh, S., Pedersen, S. & Friesen, J. D. (1976). *Mol. Gen. Genet.* **149**, 279–289.
Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V. (2000). *Nucleic Acids Res.* **28**, 33–36.
Williams, M. D., Ouyang, T. X. & Flickinger, M. C. (1994). *Mol. Microbiol.* **11**, 1029–1043.
Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. & Darst, S. A. (1999). *Cell*, **98**, 811–824.