Cloning, expression, purification, crystallization and initial crystallographic analysis of the preprotein translocation ATPase SecA from *Thermus thermophilus*

Marina N. Vassylyeva,[a,b] **Hiroyuki Mori,**[c] **Tomoya Tsukazaki,**[c]‡ **Shigeyuki Yokoyama,**[b,d] **Tahir H. Tahirov,**[e] **Koreaki Ito**[c] **and Dmitry G. Vassylyev**[a,b]*

[a]Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, 402B Kaul Genetics Building, 720 20th Street South, Birmingham, AL 35294, USA, [b]RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan, [c]Institute for Virus Research, Kyoto University, Sakyo-ku, Kyoto 606-8507, Japan, [d]RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, USA, and [e]Eppley Institute for Research in Cancer and Allied Diseases, 986805 Nebraska Medical Center, Omaha, NE 68198, USA

‡ Present address: Department of Biological Information, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Kanagawa 226-8501, Japan.

Correspondence e-mail: dmitry@uab.edu

Received 4 June 2006
Accepted 6 August 2006

The *Thermus thermophilus* gene encoding the preprotein translocation ATPase SecA was cloned and expressed and the purified protein was crystallized by the hanging-drop vapour-diffusion technique in two different space groups $P3_{1(2)}21$ ($a = b = 168.6$, $c = 149.8$ Å) and $P6_{1(5)}22$ ($a = b = 130.9$, $c = 564.6$ Å). The crystals, improved by macroseeding, diffracted to beyond 2.8 and 3.5 Å resolution for the trigonal and hexagonal crystal forms, respectively. Structure determination using the multiple isomorphous replacement method is in progress.

1. Introduction

Protein translocation across the bacterial plasma membrane is mediated by the SecYEG translocon, which forms a transmembrane channel, and by the SecA ATPase, which drives preprotein translocation utilizing the energy from ATP hydrolysis (Mori & Ito, 2001*b*; Van den Berg *et al.*, 2004; Veenendaal *et al.*, 2004; Vrontou & Economou, 2004). Specific interactions between these two components are essential for the efficiency of the translocation processes (Matsumoto *et al.*, 1997; Mori & Ito, 2001*a*, 2003; Taura *et al.*, 1997).

SecA (molecular weight ∼100 kDa), which plays a key role in preprotein translocation, has been subjected to extensive biochemical and structural studies in recent years. The primary structure (Schmidt *et al.*, 1988) and the reported crystal structures of SecA (Hunt *et al.*, 2002; Sharma *et al.*, 2003; Osborne *et al.*, 2004) show that it contains two N-terminal nucleotide-binding folds: NBF1 and NBF2. A preprotein cross-linking domain (PPXD) inserted into the NBF1 domain protrudes from the N-terminal region of the protein and binds the C-terminal translocation domain, which consists of the helical wing domain (HWD) and the helical scaffold domain (HSD). The C-terminal part of SecA also includes an ATPase regulatory domain (IRA1). Finally, the N-terminal and C-terminal (which includes the PPXD domain) regions of the protein are connected by a characteristic long *α*-helix of the HSD domain.

The important question which remains to be addressed concerns the oligomeric state of SecA in the active translocation complex. The previously determined structures of the *Bacillus subtilis* and *Mycobacterium tuberculosis* SecA proteins (Hunt *et al.*, 2002; Sharma *et al.*, 2003) revealed two distinct types of antiparallel dimers. Each of them was proposed to be an active unit of SecA during translocation. It is worth noting that in both works the crystals were obtained under non-physiological high salt concentrations. Most recently, structural studies of *B. subtilis* SecA crystallized under distinct, more physiological conditions resulted in a monomeric configuration, providing therefore a third alternative oligomeric state of SecA (Osborne *et al.*, 2004).

The relevant biochemical studies also failed to resolve the ambiguity concerning the functionally active oligomeric state of SecA. Whereas SecA is thought to predominantly form homodimers in solution (Akita *et al.*, 1991; Woodbury *et al.*, 2002; Or *et al.*, 2002), the experiments carried out to elucidate its oligomeric state on the membrane provided controversial results suggesting that SecA adopts either a monomeric or dimeric configuration (Driessen, 1993; Or *et al.*, 2002, 2005; Benach *et al.*, 2003; Jilaveanu *et al.*, 2005; Randall *et al.*, 2005; Jilaveanu & Oliver, 2006).
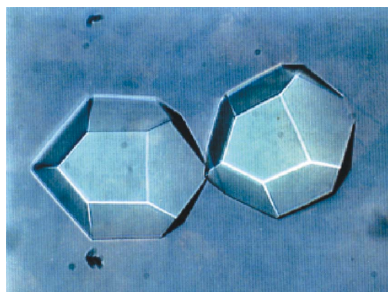
In order to shed additional light on the conformation of SecA and its oligomerization mode, we have cloned, expressed, purified and crystallized the SecA protein from *Thermus thermophilus* (TtSecA). To this end, we have obtained crystals of two crystal forms of the TtSecA protein grown under distinct but very mild (nearly physiological) conditions. The crystals belong to different space groups, each of which are likely to contain at least two SecA molecules in the asymmetric unit. We therefore presume that the structures will reveal a physiologically relevant oligimeric state(s) of the TtSecA protein.

2. Experimental procedures and results

2.1. Cloning, expression, and purification

A 3 kbp *Nde*I–*Eco*RV fragment containing the entire *SecA* gene (TTHA1251) from pET-TtsecA (Mikako Shirouzu, RIKEN Genomic Sciences Center, unpublished results) was inserted into a pUC118 derivative (with an introduced *Nde*I restriction site at the first codon of *lacZ*) after *Nde*I–*Sma*I digestion of the vector. The resulting plasmid was named pHM451.

We used a $\Delta rpoH$ strain, R40-3 (Kusukawa & Yura, 1988), in which $\sigma32$-controlled proteases are down-regulated, as a host for TtSecA overproduction. Plasmids encoding TtSecA and its variants were introduced into the R40-3 strain and plasmid-bearing cells were cultivated at 303 K until mid-log phase. *lac* transcription was then induced with 1 m*M* IPTG and 2 m*M* cAMP for 2 h. Cells were collected by centrifugation and disrupted at 277 K by passage
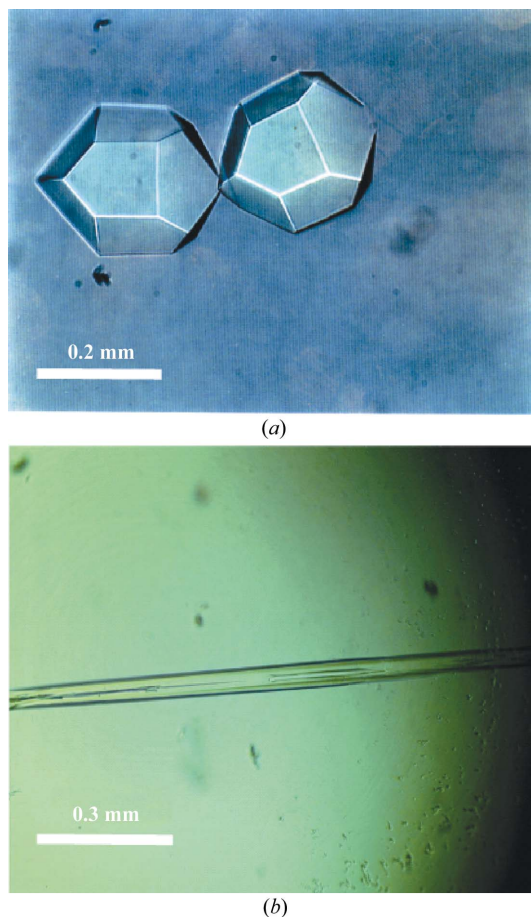
Figure 1
Crystals of the TtSecA protein. (*a*) $P3_{1(2)}22$ crystal form and (*b*) $P6_{1(5}\_22$ crystal form.

Table 1
Data-collection statistics for the TtSecA crystals.

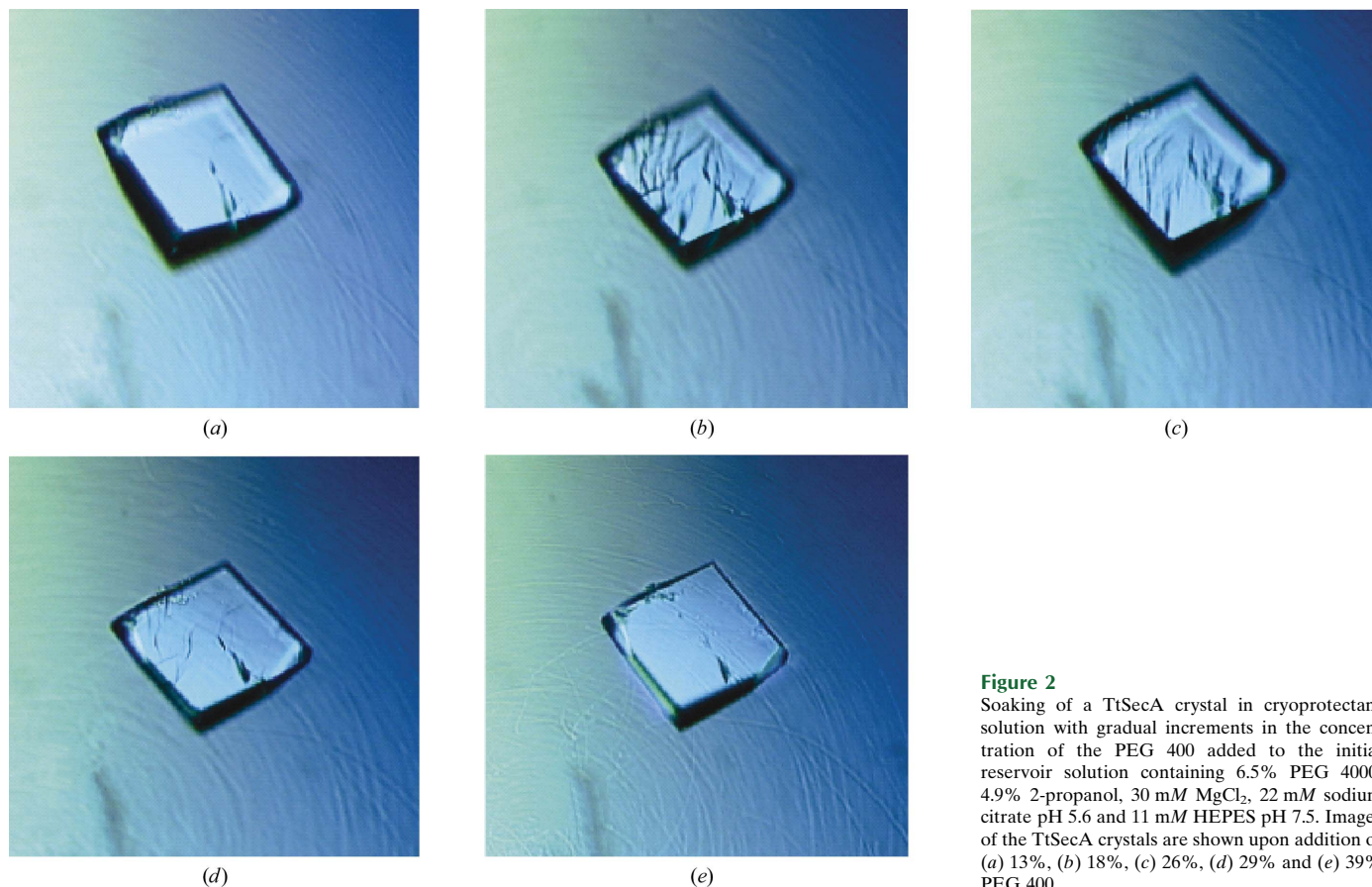Values in parentheses are for the highest resolution shell.

Space group	$P3_{(1)}21$	$P6_{1(5)}22$
Unit-cell parameters		
a (Å)	168.6	130.9
b (Å)	168.6	130.9
c (Å)	149.8	564.6
Source	Beamline BL45, SPring8	Beamline BL45, SPring8
Wavelength (Å)	1.0	1.0
Temperature (K)	100	100
Molecules in ASU	2–3	2–3
Solvent content (%)	45–63	46–64
Resolution (Å)	40.0–2.8 (2.9–2.8)	40.0–3.5 (3.65–3.5)
Observations	189610	141174
Unique reflections	59253 (5469)	33613 (3076)
Multiplicity	3.2 (2.4)	4.2 (3.4)
R_{merge}†	0.058 (0.409)	0.075 (0.424)
Completeness (%)	97.4 (92.3)	95.4 (91.5)
$I/\sigma(I)$	17.9 (4.2)	14.2 (3.0)

† $\sum |I_j - \langle I_j \rangle| / \sum \langle I_j \rangle$, where I_j is the intensity of reflection j and $\langle I_j \rangle$ is the average intensity of reflection j.

through a French pressure cell. After ultracentrifugation, the clarified soluble fraction was incubated at 343 K for 10 min to denature and to precipitate the bulk of the *Escherichia coli* proteins, which were removed by an additional centrifugation. The resulting supernatant was applied onto a Hi-Trap Blue column (Amersham Biosciences; two 1 ml columns connected in tandem), which was eluted with a linear 0.2–2.0 *M* KCl gradient in 50 m*M* Tris–HCl pH 7.5. TtSecA was eluted as a major peak at ~1.5 *M* KCl. The peak fractions were combined and concentrated by Centriprep 10 (Amicon) ultrafiltration. The sample was desalted with a Hi-Trap desalting column (Amersham Biosciences; two 5 ml columns connected in tandem) and subjected to Hi-Trap Q (Amersham Biosciences; 1 ml) column chromatography with elution with a 100–400 m*M* NaCl gradient in 50 m*M* Tris–HCl pH 7.5. TtSecA, which eluted as a sharp peak at NaCl concentration of about 200 m*M*, was further purified by gel filtration using a Hi-Load 16/60 Superdex 200pg column (Amersham Biosciences) equilibrated with 50 m*M* Tris–HCl pH 7.5, 300 m*M* NaCl. We obtained at least 10 mg of highly purified TtSecA preparation from a 4 l culture, which was used for crystallization. The purity was estimated to be 99% or higher. Note that TtSecA tends to aggregate under low-salt conditions and hence all the chromatographic media after the Hi-Trap Blue step contained at least 100 m*M* NaCl.

2.2. Crystallization and data collection

The Hampton Crystal Screen method (Jancarik & Kim, 1991) was used to determine the initial crystallization conditions for the TtSecA protein. Crystallization was carried out by the sitting-drop vapour-diffusion technique at 293 K. Drops containing 0.5 µl of the original Crystal Screen solutions and 2 µl protein solution were equilibrated against 0.5 ml of the precipitant solutions diluted threefold with 0.6 *M* NaCl solution. After 2 d equilibration, small hexagonal crystals of TtSecA protein appeared using Crystal Screen solution No. 17. The crystals were subjected to macroseeding using drops prepared under the same conditions. After a few days of equilibration, the crystals grew to typical dimensions of 0.3 × 0.3 × 0.3 mm (Fig. 1*a*) and diffracted to beyond 3.2 Å resolution at the in-house X-ray generator. They belong to space group $P3_{1(2)}21$, with unit-cell parameters $a = b = 168.6$, $c = 149.8$ Å. The best TtSecA crystals were obtained from drops in which 2.5 µl protein solution at 6 mg ml^{-1} was

(a)

(b)

(c)

(d)

(e)

mixed with 0.5 µl 26.3% PEG 4000, 175 mM lithium sulfate, 87.5 mM Tris–HCl pH 8.5.

The crystals of the other crystal form appeared after ∼2 d equilibration of a 2 µl drop in which solution No. 40 of Crystal Screen was diluted threefold before making up a crystallization drop by mixing with a protein solution containing TtSecA at a concentration of 3 mg ml^{-1} in a 1:1 ratio. After optimization of crystallization conditions followed by macroseeding, the crystals were grown to final dimensions of ∼0.08 × 0.08 × 1.0 mm (Fig. 1b) in ∼10 d. The final precipitant solution used for crystallization contained 6.5% PEG 4000, 4.9% 2-propanol, 30 mM MgCl$_2$, 22 mM sodim citrate pH 5.6 and 11 mM HEPES pH 7.5. The crystals diffracted beyond 3.5 Å resolution at the in-house X-ray generator and belong to space group $P6_{1(5)}22$, with unit-cell parameters $a = b = 130.9$, $c = 564.6$ Å.

It is worth noting that the crystals of the trigonal crystal form appeared to be highly sensitive to cryoprotectant solutions. Many cryoprotectant conditions were tested and practically all of them resulted in irreversible damage to the crystals. We finally stabilized the crystals in a solution containing PEG 400 using an approach in which the concentration of PEG 400 was increased stepwise from an initial 0% to a final 39%. Interestingly, at the beginning of this procedure the crystals began to crack and looked most damaged at a PEG 400 concentration of 18%. However, further increase of PEG 400 concentration in the drop resulted in a gradual restoration of the crystal quality, providing practically intact crystals at the final concentration of 39% (Fig. 2).

Diffraction data were collected at 100 K at 2.8 and 3.5 Å resolution for the $P3_{1(2)}21$ and $P6_{1(5)}22$ crystal forms, respectively, using synchrotron radiation (beamline BL45, SPring-8, Japan) as an X-ray source and an R-AXIS V (Rigaku) detector (Table 1). The data were processed using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Multiple attempts to solve the structures of the TtSecA protein by the molecular-replacement technique using the previously determined SecA structures from another bacterium (Hunt *et al.*, 2002; Sharma *et al.*, 2003; Osborne *et al.*, 2004) as search models were not successful. Therefore, the TtSecA structure is now in being determined using the multiple isomorphous replacement method.

References

Akita, M., Shinkai, A., Matsuyama, S. & Mizushima, S. (1991). *Biochem. Biophys. Res. Commun.* **174**, 211–216.

Benach, J., Chou, Y.-T., Fak, J. J., Itkin, A., Nicolae, D. D., Smith, P. C., Wittrock, G., Floyd, D. L., Golsaz, C. M., Gierasch, L. M. & Hunt, J. F. (2003). *J. Biol. Chem.* **278**, 3628–3638.

Driessen, A. J. (1993). *Biochemistry*, **32**, 13190–13197.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Jilaveanu, L. B. & Oliver, D. (2006). *J. Bacteriol.* **188**, 335–338.

Jilaveanu, L. B., Zito, C. R. & Oliver, D. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 7511–7516.

Hunt, J. F., Weinkauf, S., Henry, L., Fak, J. J., McNicholas, P., Oliver, D. B. & Deisenhofer, J. (2002). *Science*, **297**, 2018–2026.

Kusukawa, N. & Yura, T. (1988). *Genes Dev.* **2**, 874–882.

Matsumoto, G., Yoshihisa, T. & Ito, K. (1997). *EMBO J.* **16**, 6384–6393.

Mori, H. & Ito, K. (2001*a*). *Proc. Natl Acad. Sci. USA*, **98**, 5128–5133.

Mori, H. & Ito, K. (2001*b*). *Trends Microbiol.* **9**, 494–500.

Mori, H. & Ito, K. (2003). *J. Bacteriol.* **185**, 405–412.

Or, E., Boyd, D., Gon, S., Beckwith, J. & Rapoport, T. A. (2005). *J. Biol. Chem.* **280**, 9097–9105.

Or, E., Navon, A. & Rapoport, T. (2002). *EMBO J.* **21**, 4470–4479.

Osborne, A. R., Clemons, W. M. Jr & Rapoport, T. A. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 10937–10942.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Randall, L. L., Crane, J. M., Lilly, A. A., Liu, G., Mao, C., Patel, C. N. & Hardy, S. J. S. (2005). *J. Mol. Biol.* **348**, 479–489.

Schmidt, M. G., Rollo, E. E., Grodberg, J. & Oliver, D. B. (1988). *J. Bacteriol.* **170**, 3404–3414.

Sharma, V., Arockiasamy, A., Ronning, D. R., Savva, C. G., Holzemburg, A., Braunstein, M., Jacobs, W. R. Jr & Sacchettini, J. C. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 2243–2248.

Taura, T., Yoshihisa, T. & Ito, K. (1997). *Biochimie*, **79**, 512–517.

Van den Berg, B., Clemons, W. M. Jr, Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C. & Rapoport, T. A. (2004). *Nature (London)*, **427**, 36–44.

Veenendaal, A. K. J., van der Does, C. & Driessen, A. J. M. (2004). *Biochim. Biophys. Acta*, **1694**, 81–95.

Vrontou, E. & Economou, A. (2004). *Biochim. Biophys. Acta*, **1694**, 67–80.

Woodbury, R. L., Hardy, S. J. & Randall, L. L. (2002). *Protein Sci.* **11**, 875–882.