Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of ST0929, a putative glycosyl transferase from *Sulfolobus tokodaii*

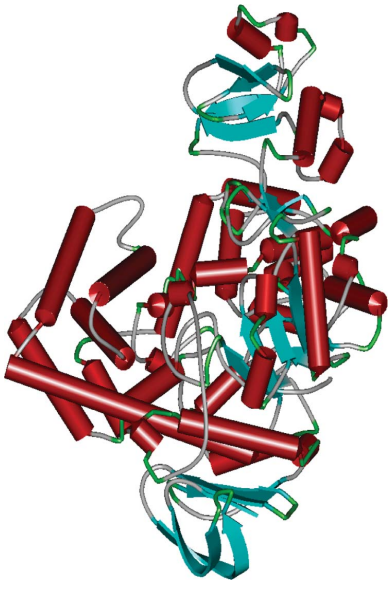
Charles B. C. Cielo,[a] Seiji Okazaki,[a] Atsuo Suzuki,[a] Tsunehiro Mizushima,[a] Ryoji Masui,[b] Seiki Kuramitsu[b] and Takashi Yamane[a]*

[a]Department of Biotechnology, School of Engineering, Nagoya University, Chikusa-ku, Nagoya 464-8603, Japan, and [b]Department of Biological Sciences, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan

Correspondence e-mail: yamane@nubio.nagoya-u.ac.jp

Received 18 December 2009
Accepted 18 February 2010

PDB Reference: hypothetical maltooligosyl trehalose synthase, 3hje.



The *Sulfolobus tokodaii* protein ST0929 shares close structural homology with *S. acidocaldarius* maltooligosyl trehalose synthase (SaMTSase), suggesting that the two enzymes share a common enzymatic mechanism. MTSase is one of a pair of enzymes that catalyze trehalose biosynthesis. The relative geometries of the ST0929 and SaMTSase active sites were found to be essentially identical. ST0929 also includes the unique tyrosine cluster that encloses the reducing-end glucose subunit in *Sulfolobus* sp. MTSases. The current structure provides insight into the structural basis of the increase in the hydrolase side reaction that is observed for mutants in which a phenylalanine residue is replaced by a tyrosine residue in the subsite +1 tyrosine cluster of *Sulfolobus* sp.

1. Introduction

Trehalose (α-D-glucopyranosyl α-D-glucopyranoside) is a chemically stable naturally occurring industrially important nonreducing disaccharide that is catabolized by various organisms exposed to extremes in environmental conditions (Higashiyama, 2002; Schiraldi *et al.*, 2002). A novel two-enzyme system obtained from cultures of *Arthrobacter* sp. Q36 has been reported to catalyze the efficient production of trehalose from starch. The system involves the combined action of maltooligosyl trehalose synthase (MTSase; EC 5.4.99.15) and maltooligosyl trehalose hydrolase (MTHase; EC 3.2.1.141) (Nakada, Maruta, Mitsuzumi *et al.*, 1995; Nakada, Maruta, Tsusaki *et al.*, 1995). MTSase catalyzes the transglycosylation of the reducing-end maltose α-1,4-glucosidic bond of its oligosaccharide substrate to a trehalose α-1,1-glucosidic bond. The MTSase active site is organized into a series of subsites (+1, −1, ..., −9), with the reducing-end glucose subunit of the substrate positioned in subsite +1 and the remaining glucose subunits interacting with subsites −1 to −9 sequentially. The α-1,4-glucosidic bond between the reducing-end glucose and its adjacent glucose subunit is cleaved by the catalytic triad, which is followed by reorientation of the cleaved subsite +1 glucose subunit and the subsequent formation of an α-1,1-glucosidic bond to the oligoside substrate. A weak hydrolytic side-reaction has also been reported in which the formation of the α-1,1-glucosidic bond is bypassed by the interaction of the cleaved reducing-end glucose unit with a water molecule, resulting in free glucose and a shortened maltooligoside (Kobayashi *et al.*, 2003; Kato *et al.*, 2000; Davies *et al.*, 1997)

Similar glycosyl transferases have subsequently been discovered in extremophilic archaebacteria such as *Sulfolobus acidocaldarius* (Nakada *et al.*, 1996*a*,*b*), *S. solfataricus* (Kato, 1999) and *S. shibatae* (Chen *et al.*, 2000).

The 1.9 Å resolution structure of lysine-methylated *S. acidocaldarius* MTSase (SaMTSase; PDB code 1iv8; Kobayashi *et al.*, 1999, 2003) is currently the only structure of an MTSase to have been solved.

Recently, ST0929, a hypothetical glycosyl transferase that shares 52% primary structure homology with SaMTSase, was isolated from the archaeon *S. tokodaii*. In this paper, we report the crystal structure

of ST0929 at 1.9 Å resolution and discuss the unique subsite architecture of *Sulfolobus* sp. MTSases.

2. Materials and methods

The ST0929 gene from *S. tokodaii* strain 7 genomic DNA was cloned by the PCR method using the forward primer 5′-ATATATGCATG-AAATTGTTGAGCACCTATAGGCTTCAACCA-3′ and the reverse primer 5′-ATATGGATCCTTATTATTTAACAAGAATTAAAGG-3′. These primers include *Nsi*I and *Bam*HI sites, respectively. The PCR product digested with *Nsi*I was processed using 3′-exonuclease to make a blunt end and then digested with *Bam*HI. This fragment was inserted into the expression vector pET-11a, which was digested with *Nde*I, filled-in with a Klenow fragment and then digested with *Bam*HI. (The 5′-site was ligated with blunt ends and the 3′-site was ligated with cohesive ends.)

After expression in BL21 cells and purification using Resource ISO (Amersham Biosciences), Resource Q (Amersham Biosciences) and HiLoad 16/60 Superdex 200 (Amersham Biosciences) columns, the recombinant protein was concentrated to 25.6 mg ml^{-1} prior to crystallization.

Crystallization of ST0929 was carried out at 293 K *via* the hanging-drop vapour-diffusion method. A drop consisting of 2 µl concentrated protein solution was mixed with 2 µl reservoir solution [0.2 *M* potassium sodium tartrate tetrahydrate, 20%(*w*/*v*) polyethylene

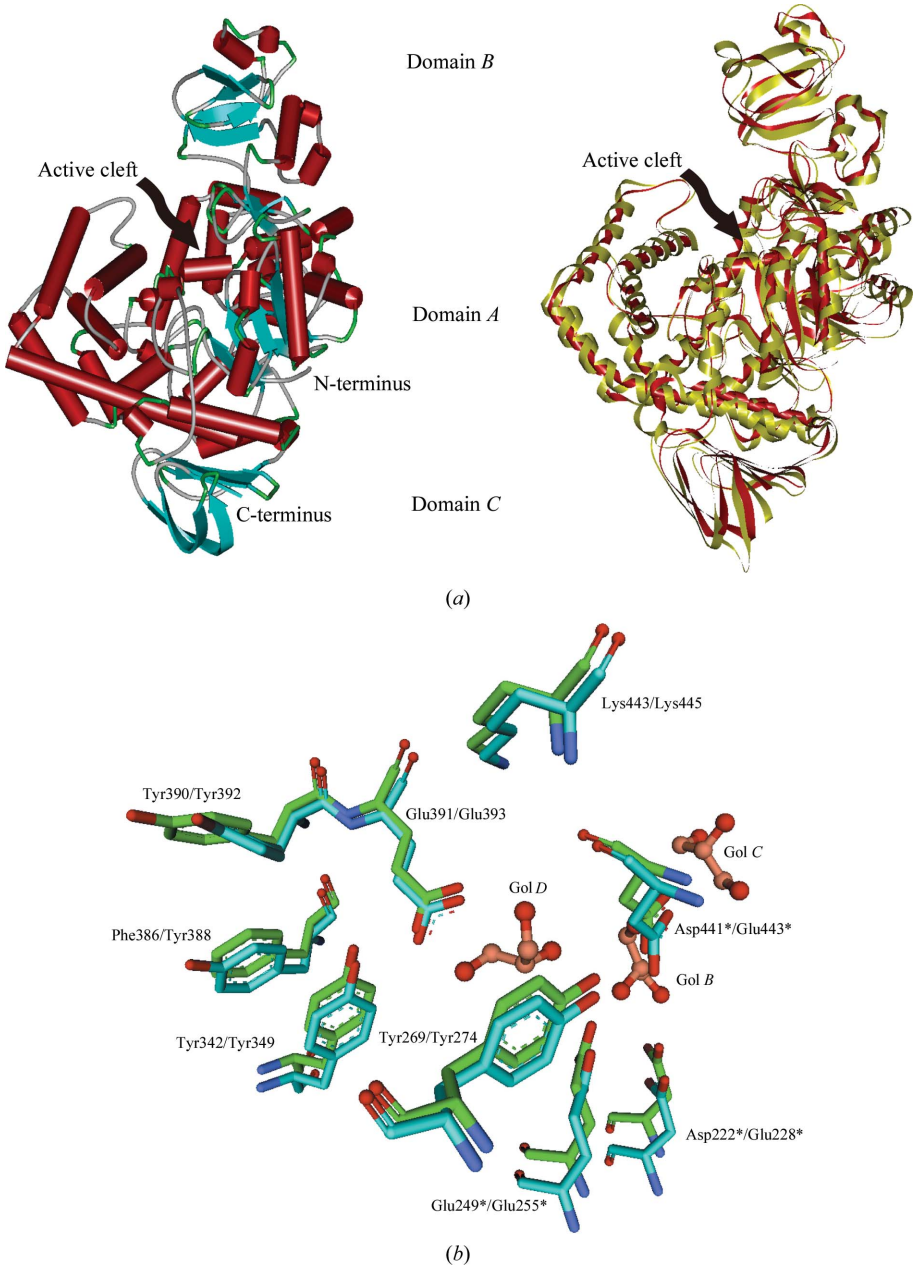


Figure 1
(*a*) Left, the overall structure of ST0929. The N- and C-termini are indicated. α-Helices are depicted in red and β-strands are depicted in cyan. Right, the superimposed structures of ST0929 (red) and SaMTSase (yellow). The black arrow shows the location of the active-site cleft. (*b*) The SaMTSase subsite +1 tyrosine pocket (cyan) superimposed with its corresponding region in ST0929 (green). The glycerol molecules (peach) found in the ST0929 model and the catalytic triad residues (*) of SaMTSase and ST0929 are also shown.

glycol 3350] and equilibrated against 350 µl reservoir solution. Hexagonal crystals grew in 3–5 d. 0.1 *M* glycerol in reservoir solution was used as a cryoprotectant. Diffraction data were collected to 1.9 Å resolution using a Rigaku FR-E Cu *Kα* X-ray source equipped with a Rigaku R-AXIS VII detector in the High-Intensity X-ray Diffraction Laboratory of Nagoya University. Diffraction data were processed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

The structure was solved by molecular replacement using *MOLREP* (Vagin & Teplyakov, 1997) with the structure of SaMTSase (PDB code 1iv8) as a search template. Subsequent refinement was performed using *REFMAC*5 (Murshudov *et al.*, 1997) and *ARP*/*wARP* (Morris *et al.*, 2003) was used to model water molecules. Fine-tuning of the model was carried out using *Coot* (Emsley & Cowtan, 2004). Model validation was carried out using *PROCHECK* (Laskowski *et al.*, 1993; Morris *et al.*, 1992). The figures were prepared using *CCP*4*MG* (Potterton *et al.*, 2004) and *Accelrys Discovery Studio* 2.5. *MOLREP*, *REFMAC*5, *ARP*/*wARP*, *PROCHECK* and the *CCP*4*mg* program are included in the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Data-refinement statistics are listed in Table 1.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 63.99, b = 86.10, c = 70.13, β = 95.43
X-ray source	Cu $K\alpha$
Temperature (K)	100
Wavelength (Å)	1.5418
Resolution range (Å)	100–1.9 (1.97–1.9)
Total No. of observations	116816 (11652)
Completeness (%)	99.1 (99.2)
R_{merge} (%)	8.1 (39.5)
$I/\sigma(I)$	22.5 (4.1)
Refinement (19.94–1.90 Å)	
Completeness (%)	99.8
No. of reflections	56580
Test set (5.3%)	3014
No. of protein atoms	6075
No. of water atoms	858
No. of glycerol atoms	24
Average B factors (Å^2)	23.5
R_{cryst}	0.171
R_{free}	0.206
R.m.s.d. bond lenths (Å)	0.007
R.m.s.d. bond angles (°)	0.98
Ramachandran plot statistics (%)	
Residues in most favoured regions	90
Residues in additional allowed regions	9.7
Residues in generously allowed regions	0.2
Residues in disallowed regions	0.1

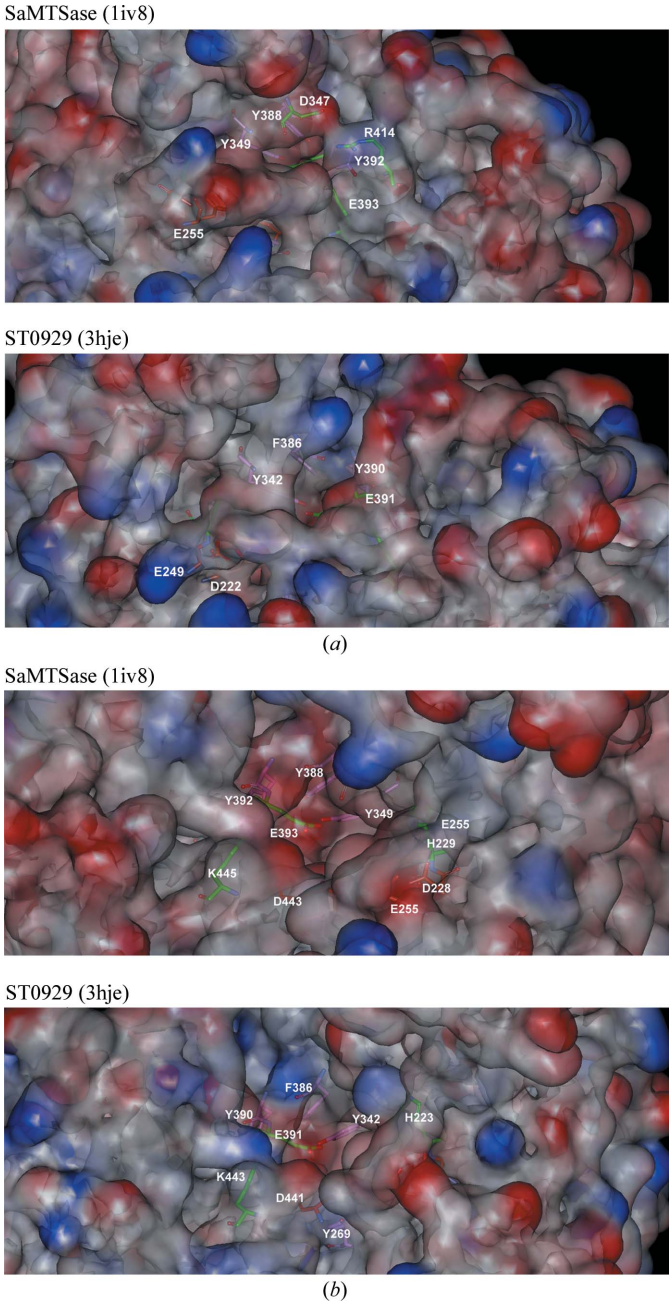


Figure 2
Representation of the unsolvated surfaces of the ST0929 and SaMTSase subsite +1 tyrosine clusters. Electronegative areas are depicted in red, while positive areas are depicted in blue. (*a*) External view of subsite +1. The tyrosines comprising the tyrosine cluster are depicted as pink sticks. Other visible subsite +1 residues are depicted as green sticks. The catalytic residues Asp222 and Glu249 are also shown. The hydroxyl group of Tyr388 and the salt bridge formed between Asp347 and Arg414 in SaMTSase decrease the solvent-accessibility of subsite +1. (*b*) View of subsite +1 as seen from the active cleft. The tyrosines comprising the tyrosine cluster are depicted as pink sticks. Other visible subsite +1 residues are depicted as green sticks. The catalytic residue Asp441 is also shown. The hydroxyl group of the SaMTSase Tyr388 contributes to the electronegative patch found at the entrance of the subsite.

3. Results and discussion

ST0929 folds into three domains that exhibit the same domain organization as SaMTSase, with an overall r.m.s.d. of 1.1 Å for 667 of 704 ST0929 C$^\alpha$ atoms (Fig. 1*a*). The N-terminal domain (domain *A*; Met1–Pro86 and Val196–Gly641) contains an incomplete α/β-barrel core structure comprised of seven β-strands similar to the barrel core that is characteristic of family 13 glucosyl hydrolases (Gueguen *et al.*, 2001; Henrissat, 1991). Domain *B* (Asn87–Ala195) forms a slightly distorted domain of seven β-strands arranged into two antiparallel β-sheets flanked by two α-helices. Domain *C* (Glu642–Lys704) folds into six antiparallel β-strands on the opposite side of the putative active cleft.

The residues that comprise the active site (in particular subsites +1, −1 and −2) in ST0929 are almost identical in location and orientation to those of SaMTSase (Fig. 1*b* and Supplementary Fig. 1[1]). The close

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: TB5022).

resemblance of the active sites of the two proteins strongly suggests identical functionality.

Subsite +1 is of particular interest as it appears to be a structural feature that is unique to MTSase. Like that of SaMTSase, subsite +1 of ST0929 has sufficient dimensions to accommodate a single glucose subunit, as could be inferred from the two glycerol molecules found in the proximity of subsites −1 and +1 (Fig. 1*b* and Supplementary Fig. 1). It also includes a hydrophobic pocket lined by a tyrosine cluster which is conserved among *Sulfolobus* sp. MTSases (Supplementary Fig. 2) and is structurally equivalent to the corresponding SaMTSase residues (Fig. 1*b*).

Several attempts have been made to improve the yield of trehalose by minimizing the hydrolytic side-reaction *via* the mutation of subsite residues (Fang *et al.*, 2006; Maruta *et al.*, 2006), providing valuable insights into the roles of specific subsite +1 residues with respect to the equilibrium between the main transglycosylation reaction and the hydrolysis side reaction. The primary function of the tyrosine cluster appears to be to facilitate the reorientation of the cleaved reducing-end glucose subunit by shielding it from exposure to solvent molecules that may interfere with the transglycosylation process as well as to provide hydrogen-bonding points to assist in its reorientation (Kobayashi *et al.*, 1999).

Like ST0929, wild-type *S. solfataricus* MTSase contains a phenylalanine residue at the position of Tyr388 in SaMTSase (Figs. 1*b*, 2*a* and 2*b*; Fang *et al.*, 2006). Mutation of this phenylalanine residue to tyrosine in *S. solfataricus* MTSase was observed to slightly decrease the hydrolysis side-reaction (Fang *et al.*, 2006, 2007). The structure of ST0929 provides insight into the structural basis of this shift. In SaMTSase, the hydroxyl group of Tyr388 helps to occlude the cavity leading from subsite +1. The ST0929 subsite +1 contains Phe386 at this position, resulting in a larger aperture and conceivably increased solvent access to the subsite (Fig. 2*a*). Furthermore, in SaMTSase Asp347 and Arg414 form a salt bridge that contributes to shielding the subsite from solvent intrusion (Fig. 2*a*).

Fig. 2(*b*) depicts the surface of the tyrosine cluster as seen from the active cleft. The hydroxyl group of Tyr388 in SaMTSase appears to contribute to the electronegative patch formed by the tyrosine cluster and is thought to engage in hydrogen bonding to the reducing-end glucose subunit to assist in the reorientation of the cleaved glucose molecule (Fang *et al.*, 2006, 2007; Kobayashi *et al.*, 2003).

The slight shift towards increased hydrolase activity could be explained by increased solvent access to subsite +1 of ST0929, as well as by the loss of a hydroxyl group from the tyrosine that is replaced by a phenylalanine in ST0929.

Thus, we predict that ST0929 is a less effective glycosyl transferase than SaMTSase.

This work was partly supported by a research grant from the National Project on Protein Structural and Functional Analysis from the Ministry of Education, Culture, Sports, Science and Technology of Japan. The authors thank Mr Hikage, Nagoya University for his invaluable help during data collection.

References

Chen, W., Liu, L., Sun, P. & Jin, C. (2000). *Acta Microbiol. Sin.* **40**, 57–61.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Davies, G. J., Wilson, K. S. & Henrissat, B. (1997). *Biochem. J.* **321**, 557–559.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fang, T.-Y., Tseng, W.-C., Chung, Y.-T. & Pan, C.-H. (2006). *J. Agric. Food Chem.* **54**, 3585–3590.
Fang, T.-Y., Tseng, W.-C., Pan, C.-H., Chun, Y.-T. & Wang, M.-Y. (2007). *J. Agric. Food Chem.* **55**, 5588–5594.
Gueguen, Y., Rolland, J. L., Schroeck, S., Flament, D., Defretin, S., Saniez, M. H. & Dietrich, J. (2001). *FEMS Microbiol. Lett.* **194**, 201–206.
Henrissat, B. (1991). *Biochem. J.* **280**, 309–316.
Higashiyama, T. (2002). *Pure Appl. Chem.* **74**, 1263–1269.
Kato, M. (1999). *J. Mol. Catal. B Enzym.* **6**, 223–233.
Kato, M., Takehara, K., Kettoku, M., Kobayashi, K. & Shimizu, T. (2000). *Biosci. Biotech. Biochem.* **64**, 319–326.
Kobayashi, M., Kubota, M. & Matsuura, Y. (1999). *Acta Cryst.* D**55**, 931–933.
Kobayashi, M., Kubota, M. & Matsuura, Y. (2003). *J. Appl. Glycosci.* **50**, 1–8.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Maruta, K., Kubota, M., Yamashita, H., Nishimoto, T., Chaen, H. & Fukuda, S. (2006). *J. Appl. Glycosci.* **53**, 199–203.
Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). *Proteins*, **12**, 345–364.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nakada, T., Ikegami, S., Chaen, C., Kubota, M., Fukuda, S., Sugimoto, T., Kurimoto, M. & Tsujisaka, Y. (1996*a*). *Biosci. Biotechnol. Biochem.* **60**, 263–266.
Nakada, T., Ikegami, S., Chaen, C., Kubota, M., Fukuda, S., Sugimoto, T., Kurimoto, M. & Tsujisaka, Y. (1996*b*). *Biosci. Biotechnol. Biochem.* **60**, 267–270.
Nakada, T., Maruta, K., Mitsuzumi, H., Chaen, C., Kubota, M., Sugimoto, T., Kurimoto, M. & Tsujisaka, Y. (1995). *Biosci. Biotechnol. Biochem.* **59**, 2215–2218.
Nakada, T., Maruta, K., Tsusaki, K., Chaen, C., Kubota, M., Fukuda, S., Sugimoto, T., Kurimoto, M. & Tsujisaka, Y. (1995). *Biosci. Biotechnol. Biochem.* **59**, 2210–2214.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G. N., Cohen, S., Perrakis, A. & Noble, M. (2004). *Acta Cryst.* D**60**, 2288–2294.
Schiraldi, C., Di Lernia, I. & De Rosa, M. (2002). *Trends Biotechnol.* **20**, 420–425.
Vagin, A. A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.