structural communications

**Kwok-Ho Chan and Kam-Bo
Wong***

School of Life Sciences, Centre for Protein
Science and Crystallography, The Chinese
University of Hong Kong, Hong Kong,
People's Republic of China

Correspondence e-mail: kbwong@cuhk.edu.hk

# Structure of an essential GTPase, YsxC, from *Thermotoga maritima*

YsxC belongs to the YihA family of TRAFAC class GTPases. The protein is involved in the biogenesis of ribosomes and is essential for the survival of a wide range of bacteria. Here, crystal structures of YsxC from *Thermotoga maritima* and its complex with GDP were determined at maximal resolutions of 2.3 and 1.9 Å, respectively. Major structural differences are observed in the switch I region, which is disordered in the apo form but exists in both an 'open' and a 'closed' conformation in the GDP-bound state. A comparison with the structure of the GMPPNP–YsxC complex from *Bacillus subtilis* provides insights into the mechanism of conformational change in the switch I and II regions upon hydrolysis of GTP.

1. Introduction

YsxC belongs to the YihA protein family within the 'TrmE-Era-EngA-YihA-Septin-like' superfamily of the translation factor-related (TRAFAC) class of GTPases (Leipe *et al.*, 2002). Systematic knockout of genes from the *Escherichia coli* genome showed that *yihA*, a homologue of *ysxC*, is essential for survival of the bacterium (Arigoni *et al.*, 1998). It was later shown that YsxC is also an essential gene for the growth of *Bacillus subtilis* (Prágai & Harwood, 2000), *Haemophilus influenzae*, *Staphylococcus aureus* and *Streptococcus pneumoniae* (Lehoux *et al.*, 2003). Because of this feature, YsxC may be a novel target for antibacterial drugs (Lehoux *et al.*, 2003).

It has been shown that YsxC is involved in the biogenesis of the ribosome in *B. subtilis* and *S. aureus* (Schaefer *et al.*, 2006; Wicker-Planquart *et al.*, 2008; Cooper *et al.*, 2009). However, how YsxC plays such a role is poorly understood. Depletion of YsxC resulted in the accumulation of a 45S intermediate lacking the L16 and L36 components instead of the mature 50S subunit (Schaefer *et al.*, 2006). Recombinant YsxC was shown to interact mainly with the 50S subunit. A number of potential interacting partners for YsxC have been proposed, including ribosomal proteins L1, L3, L6 and L7/L12 (Wicker-Planquart *et al.*, 2008). In the presence of GDP/GTP/GMPPNP, YsxC could also interact, to a lesser extent, with the 30S subunit (Wicker-Planquart *et al.*, 2008).

GTPases are molecular switches that regulate a variety of cellular processes such as signal transduction and protein synthesis (Bourne *et al.*, 1991). The structure of GTPase adopts an α/β-fold with a mixed six-stranded β-sheet sandwiched by five α-helices. GTP hydrolysis changes the conformation of GTPases from a GTP-bound 'on' state to a GDP-bound 'off' state. Such conformational changes often localize in two regions, namely the switch I and II regions. Crystal structures of YsxC from *B. subtilis* (BS-YsxC) in the presence of GDP or the nonhydrolyzable GTP analogue 5′-guanylyl-imidodi-phosphate (GMPPNP) and in the absence of nucleotides have been reported by Ruzheinikov *et al.* (2004). They showed that the switch I loop of BS-YsxC adopts a 'closed' conformation in the GMPPNP-bound state and an 'open' conformation in the apo form. In contrast, switch I was disordered in the GDP-bound state. Here, we have determined the crystal structures of the apo and GDP-bound forms
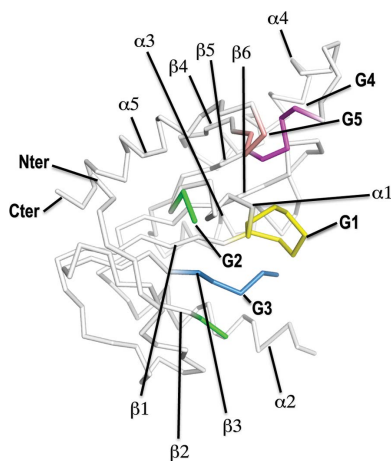
of YsxC from *Thermotoga maritima* (TM-YsxC). The structure of GDP–TM-YsxC reported here showed that the switch I loop can exist in both 'closed' and 'open' conformations, which provides novel insights into the structural mechanism of the conformational changes upon GTP hydrolysis.

2. Materials and methods

2.1. Cloning, expression and purification

The gene TM1466 encoding *T. maritima* YsxC (TM-YsxC) was amplified by polymerase chain reaction from the genomic DNA of *T. maritima* (ATCC) using the forward primer 5′-TATGCAG-GATCCATTATCAGAGATGTAGAA-3′ and the reverse primer 5′-TATGCAGAATTCTCAATTTTCTTTCAGTAA-3′. The amplified product was cloned into the pRSETA-MBP vector using the *Bam*HI and *Eco*RI restriction sites for the expression of TM-YsxC as a C-terminal fusion with maltose-binding protein. The expression vector was transformed into *E. coli* strain BL21 (DE3) pLysS (Novagen). Transformed cells were grown in LB medium containing 100 mg l^{-1} ampicillin and 50 mg l^{-1} chloramphenicol at 310 K until an OD_{600} of 0.4–0.6 was reached, when protein expression was induced using 0.4 mM isopropyl β-D-1-thiogalactopyranoside. Cells were harvested by centrifugation at 9000g for 10 min. Cells from 1 l culture were resuspended in 50 ml buffer *A* (20 mM sodium phos-

phate pH 7.4) and disrupted by sonication. The cell lysate was centrifuged at 12 000g for 20 min and the supernatant was loaded onto a 5 ml HiTrap SP ion-exchange column (GE Healthcare) equilibrated with buffer *A* and eluted with a 150 ml linear gradient of 0–0.8 M NaCl in buffer *A*. The maltose-binding protein fusion tag was then cleaved by overnight digestion at 310 K with 50 units of thrombin (Sigma). The digestion product was dialysed in buffer *A*, loaded onto a 5 ml HiTrap Heparin column (GE Healthcare) equilibrated with buffer *A* and eluted with a 150 ml linear gradient of 0–1 M NaCl in buffer *A*. The eluted protein was further purified by size-exclusion chromatography using a HiLoad 26/60 Superdex 75 gel-filtration column (GE Healthcare) equilibrated with 0.1 M NaCl, 10 mM MgCl$_2$, 20 mM HEPES buffer pH 7.5. TM-YsxC eluted at approximately 200 ml.

2.2. Sample preparation for the apo form of TM-YsxC

We initially crystallized TM-YsxC without the addition of any nucleotide. Inspection of the electron-density map after structure solution revealed the presence of GDP in the nucleotide-binding site of TM-YsxC, suggesting that GDP from the expression host was retained after the purification protocol. To remove the bound GDP, 1 mg ml^{-1} TM-YsxC protein sample was dialysed extensively in unfolding buffer (7 M guanidine hydrochloride in 20 mM Tris buffer pH 7.5). The protein was then refolded by dialysis in 0.1 M NaCl, 20 mM Tris buffer pH 7.5 to obtain the apo form of TM-YsxC. The
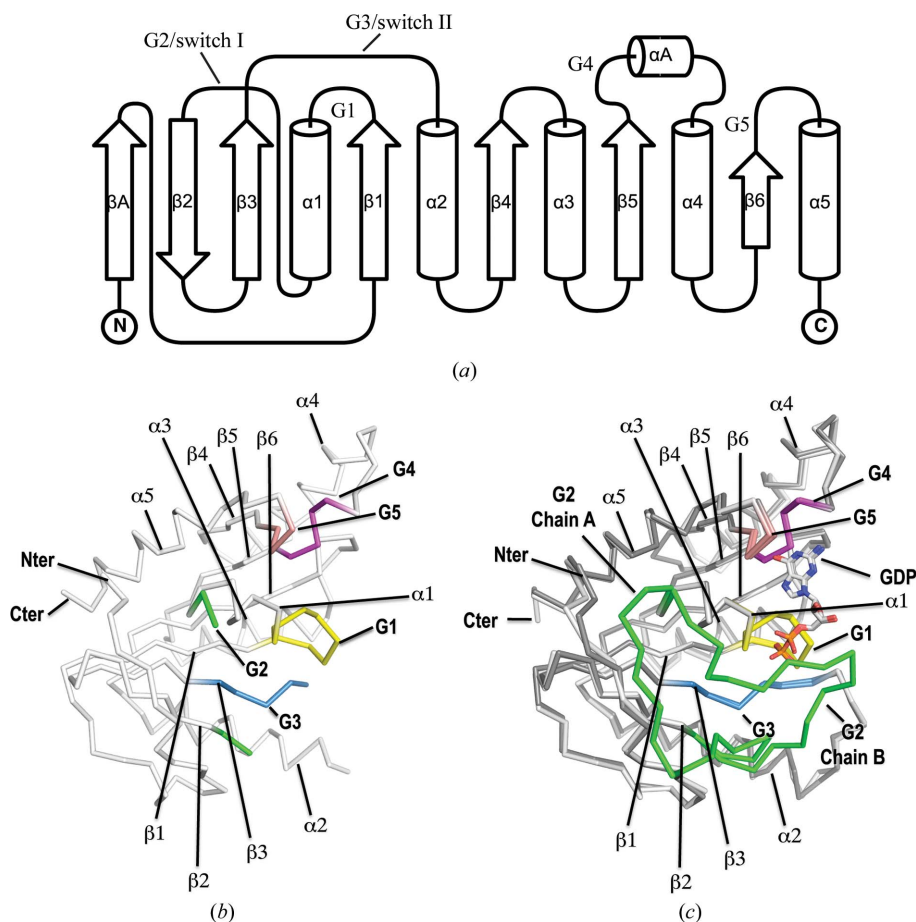

Figure 1
Topology diagram (*a*) and cartoon representation of TM-YsxC in (*b*) the apo form and (*c*) the GDP-bound form. The canonical nomenclature is used to number the secondary-structure elements of TM-YsxC. The conserved regions G1 (yellow), G2 (green), G3 (sky blue), G4 (magenta) and G5 (salmon) are indicated. The structure of apo TM-YsxC is very similar to the structure of the GDP-bound form except that the G2 and G3 loops are disordered in the apo form. It is notable that in GDP-bound TM-YsxC the conformation of the G2 (switch I) loops (green) is different in chains *A* and *B*. The other regions in the two chains (*A*, dark grey; *B*, light grey) are almost identical.

protein sample was subsequently used to obtain crystals of both the apo and the GDP-bound forms of TM-YsxC.

2.3. Crystallization, data collection and structure determination

Crystals of the apo form of TM-YsxC were grown using the sitting-drop vapour-diffusion method by mixing 1 µl 10 mg ml^{-1} protein sample with 1 µl 0.1 M Bis-Tris buffer pH 5.5 at 289 K and were cryoprotected by soaking them in 15%(v/v) PEG 400 in mother liquor. Crystals of the GDP–TM-YsxC complex were obtained by cocrystallization of TM-YsxC with GDP. Crystals were grown using the sitting-drop vapour-diffusion method by mixing 1 µl 10 mg ml^{-1}

protein sample supplemented with 1 mM GDP with 1 µl 30% PEG 4000, 0.2 M ammonium sulfate, 0.2 M acetate buffer pH 5.5 at 289 K and were cryoprotected by soaking them in 30%(v/v) glycerol in mother liquor.

Cryoprotected crystals were flash-cooled in a nitrogen stream and diffraction data were collected at 100 K using an in-house Cu $K\alpha$ source (Rigaku MicroMax-007) and an R-AXIS IV image-plate detector. For the GDP-bound form, a total of 180 images were collected using 0.5° oscillation with an exposure time of 10 min and a crystal-to-detector distance of 100 mm. For the apo form, a total of 360 images were collected using 0.5° oscillation with an exposure time of 10 min and a crystal-to-detector distance of 150 mm. Diffraction
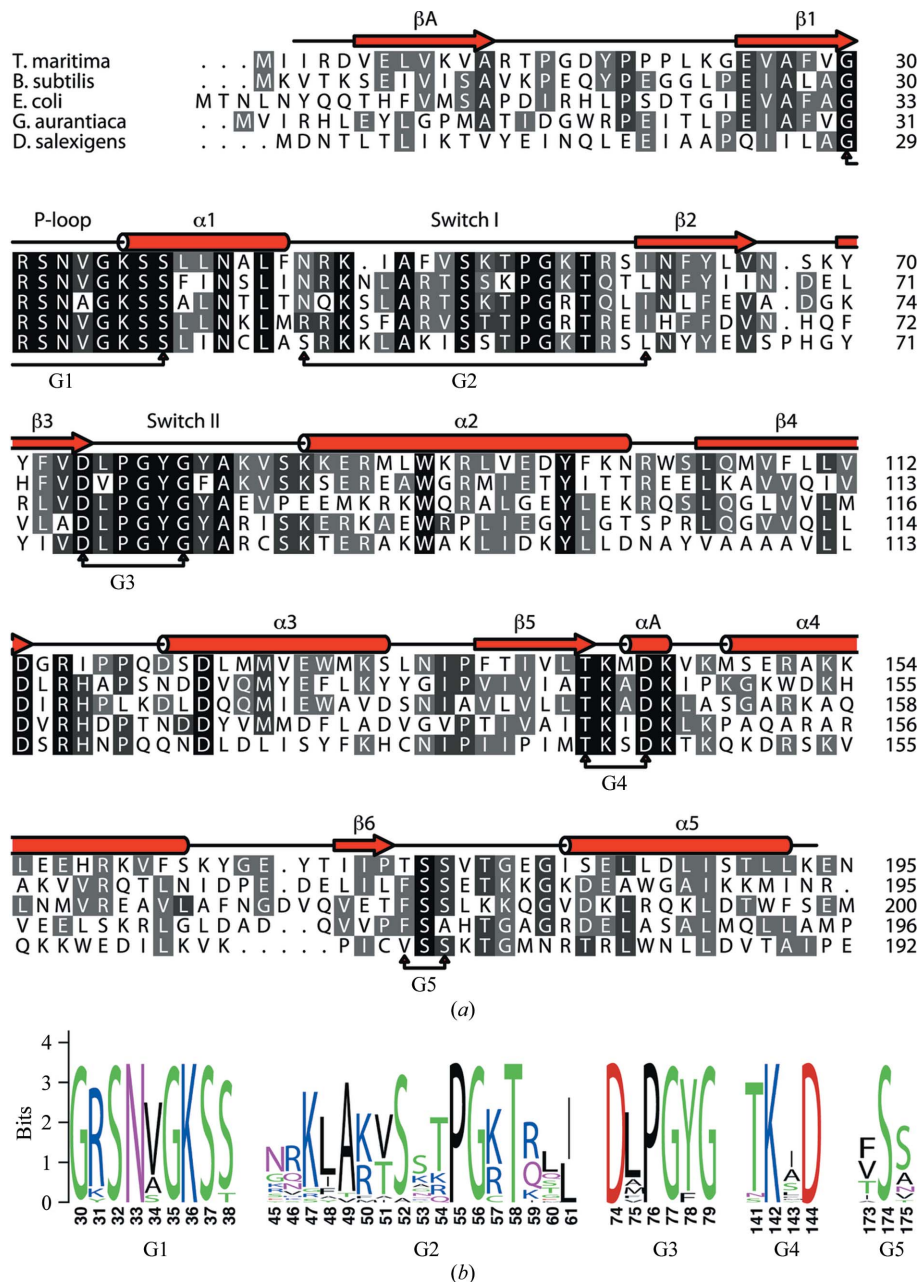

Figure 2

(a) Amino-acid sequences of YsxC and its homologues from *Thermogota maritima*, *Bacillus subtilis*, *Escherichia coli*, *Gemmatimonas aurantiaca* and *Desulfovibrio salexigens* were aligned using the program *MUSCLE* (Edgar, 2004). These homologous sequences share 39.5% (*B. subtilis*), 35.0% (*E. coli*), 38.0% (*G. aurantiaca*) and 35.2% (*D. salexigens*) sequence identity to the *T. maritima* sequence. The secondary-structure elements of TM-YsxC and the conserved regions G1–G5 are indicated. (b) Sequence-logo representation of the consensus sequence of the G1–G5 regions in the YsxC protein family. The sequence numbers of TM-YsxC are shown below the logos. After removal of redundant sequences at 80% sequence identity, 23 YsxC sequences from the SWISS-PROT database were aligned with the program *MUSCLE*. The sequence-logo representation was created by the program *WebLogo* (Crooks *et al.*, 2004).

Table 1
Diffraction data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Apo TM-YsxC	GDP-bound TM-YsxC
Data collection		
X-ray source	Cu $K\alpha$	Cu $K\alpha$
Space group	$C222_1$	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 69.71$, $b = 97.99$, $c = 54.85$	$a = 62.67$, $b = 69.71$, $c = 94.84$
Resolution (Å)	48.97–2.30 (2.42–2.30)	23.71–1.90 (2.00–1.90)
Unique reflections	8326 (1236)	33459 (4808)
Completeness (%)	96.0 (100)	99.9 (100)
Multiplicity	6.8 (6.9)	6.8 (6.2)
$\langle I/\sigma(I)\rangle$	19.4 (4.8)	15.9 (3.2)
R_{merge}	0.102 (0.269)	0.070 (0.366)
No. of protein molecules per asymmetric unit	1	2
Matthews coefficient (Å^3 Da^{-1})	2.09	2.31
Solvent content (%)	41.1	46.7
Refinement		
R_{cryst}/R_{free}	0.218/0.257	0.183/0.232
Model		
Protein atoms	1403	3133
Heterogen atoms	5	70
Water atoms	71	377
Missing residues	1, 47–59, 79–85, 194–195	Chain A, 1, 194–195; chain B, 1
R.m.s.d. from ideal values		
Bond lengths (Å)	0.006	0.005
Bond angles (°)	0.951	0.947
Ramachandran analysis		
Favoured region (%)	97.0	96.1
Outliers (%)	0	0.3

data were indexed, integrated, scaled and merged using *MOSFLM*, *SCALA* and *TRUNCATE* from the *CCP*4 suite (Winn *et al.*, 2011). The structures were solved by molecular replacement using *MOLREP* (Vagin & Teplyakov, 2010) with the structure of BS-YsxC (PDB entry 1svi; Ruzheinikov *et al.*, 2004) as the search template. The search model was prepared using *Align*4*mr* (Cheung *et al.*, 2004).

Models were built iteratively using *Coot* (Emsley & Cowtan, 2004) and refined in *PHENIX* (Afonine *et al.*, 2010). The statistics for data collection and refinement are summarized in Table 1. Superimposition of structures and calculation of C$^\alpha$ r.m.s.d. were performed using *PyMOL* (Schrödinger LLC).

3. Results

3.1. Overall structure of apo TM-YsxC and GDP–TM-YsxC

We solved the structure of TM-YsxC in the apo form and in the GDP-bound form (Fig. 1). The apo form crystallized in space group $C222_1$ with one protein molecule in the asymmetric unit, while the GDP-bound form crystallized in space group $P2_12_12_1$ with two protein molecules in the asymmetric unit (Table 1). The fold of TM-YsxC belongs to the TRAFAC class of GTPases (Leipe *et al.*, 2002). TM-YsxC adopts an α/β-fold with five α-helices on both sides of a mixed seven-stranded β-sheet (Fig. 1*a*). In addition to the canonical six-stranded β-sheet of the GTPase fold, there is an extra β-strand (βA) at the N-terminus that runs antiparallel to strand 2.

The C$^\alpha$ atoms of the GDP-bound form can be superimposed with those of the apo form with an r.m.s.d. of 0.49–0.54 Å for chains A and B, respectively. TM-YsxC is a 195-residue protein with a molecular weight of 22.4 kDa. In apo TM-YsxC residues 47–59 (switch I) and 79–85 (switch II) are disordered, suggesting intrinsic flexibility of these two loops in the absence of bound nucleotide (Fig. 1*b*). In contrast, these residues are well defined in the GDP–TM-YsxC structure. It is notable that the conformation of the switch I loop of GDP–TM-YsxC is different in chains A and B. Otherwise, the two chains are very similar and can be superimposed with an r.m.s.d. of 0.30 Å for the C$^\alpha$ atoms (Fig. 1*c*).

3.2. Nucleotide-binding site

Structure–sequence comparisons of TM-YsxC with other homologues within the YsxC protein family allowed us to identify the five
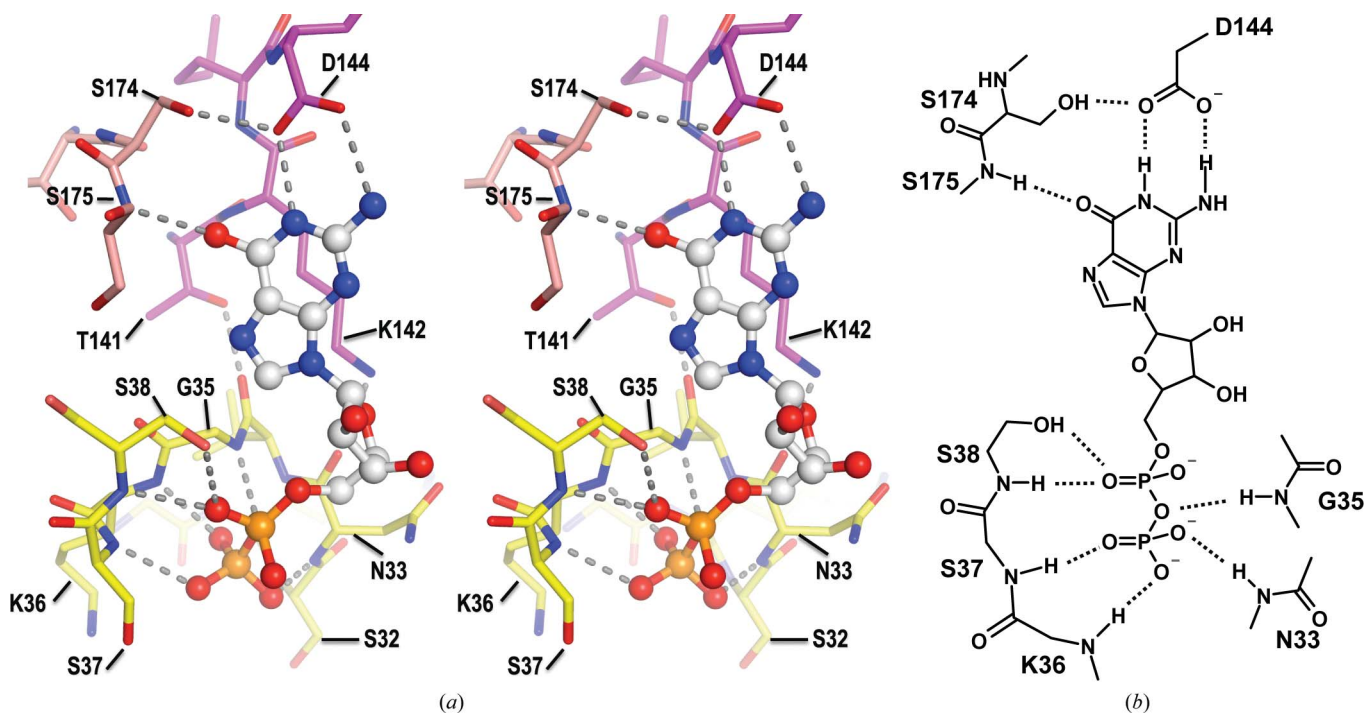
Figure 3
(*a*) Stereo diagram of the nucleotide-binding site of TM-YsxC. Residues in the conserved regions G1 (yellow), G4 (magenta) and G5 (salmon) are colour-coded. The bound GDP is shown as a ball-and-stick model. (*b*) Schematic diagram illustrating the interactions between TM-YsxC and the bound GDP.

conserved regions in GTPases, G1–G5, which are involved in binding to the guanine nucleotide (Fig. 2). The G1 region (G$xxxx$GKS) forms the P-loop that is responsible for binding the phosphate groups of the guanine nucleotide (Leipe *et al.*, 2002). In the YihA protein family there is an additional conserved serine (Ser38) next to the P-loop, resulting in the consensus sequence $G_{30}x$SNxGKS(S/T)$_{38}$ (where x can be any amino acid; Fig. 2). The backbone amide N atoms of Asn33, Gly35, Lys36 and Ser37 and the hydroxyl group of Ser38 form hydrogen bonds to the O atoms of the α- and β-phosphate groups of GDP (Fig. 3). The G2 region corresponds to the switch I loop, which contains the consensus motif PGxT (Fig. 2). The invariant threonine residue (Thr58) is responsible for binding a magnesium ion required for catalysis (Ruzheinikov *et al.*, 2004). The G3 region corresponds to the Walker B motif (DxxG), which forms part of switch II. The YihA protein family is characterized by an (F/Y)G consensus sequence next to the Walker B motif (Leipe *et al.*, 2002; Fig. 2). The G4 and G5 regions are responsible for guanine-nucleotide specificity. The

invariant Asp144 in the TKxD motif of the G4 region forms a double hydrogen bond to the N1 and N2 atoms of GDP (Fig. 3). In contrast, the invariant serine residue (Ser174) of the G5 region forms a hydrogen bond to Asp144 that anchors the backbone amide of Ser175 in a position to form a hydrogen bond to the O6 atom of GDP (Fig. 3). Thr141 and Lys142 of the G4 region form hydrogen bonds to the backbone O atoms of Asn33 and Val44 of the G1 region, stabilizing the conformation of the two regions in a position which allows the binding of guanine nucleotides.

3.3. The switch I loop of GDP-bound TM-YsxC exists in 'open' and 'closed' conformations

In the GDP–TM-YsxC complex the conformations of the switch I loops of chains A and B differ. The switch I loop of chain A is in a 'closed' conformation and that of chain B is in an 'open' conformation (Fig. 4a; Ruzheinikov *et al.*, 2004). In the 'closed' conformation
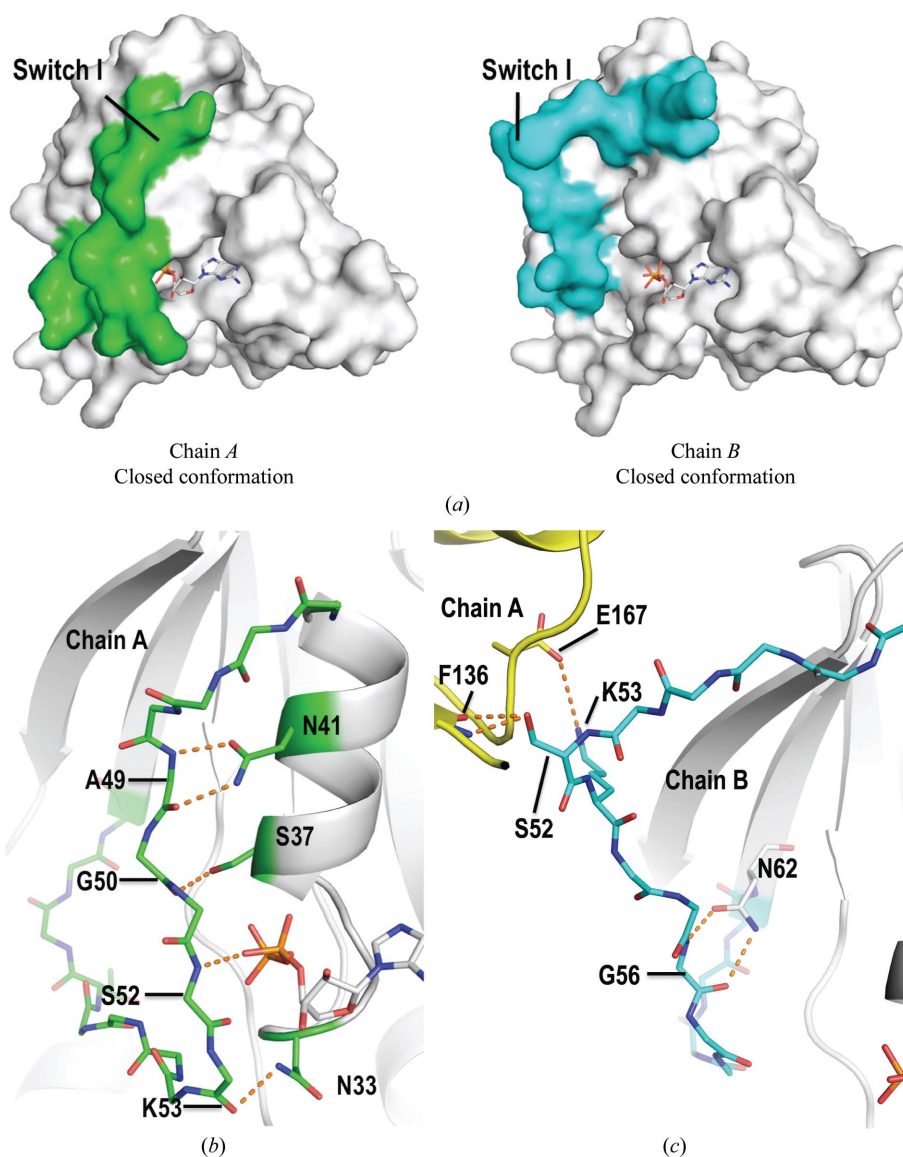

Figure 4
The switch I loop in the 'open' and 'closed' conformations of GDP–TM-YsxC. (a) In chain A (green), the switch I loop is flipped towards helix 1 and the P-loop covering the phosphate groups of the bound GDP. In contrast, the switch I loop of chain B (cyan) is positioned away from the nucleotide-binding site, exposing the bound GDP. (b) The 'closed' conformation of the switch I loop in chain A is stabilized by hydrogen bonds (orange dotted lines) to Asn33, Ser37, Asn41 and the α-phosphate group of GDP. (c) The 'open' conformation of the switch I loop in chain B is mainly stabilized by crystal packing. Ser52 and Lys53 of chain B form hydrogen bonds and a salt bridge to Phe136 and Glu167 of chain A.

the switch I loop is flipped towards helix 1 and the P-loop so that the backbone N atom of Ser52 forms a hydrogen bond to the O atom of the α-phosphate group of GDP (Fig. 4b). The main-chain conformation of the G2 loop is maintained by hydrogen bonding to the side chains of conserved residues in helix 1 and the P-loop (Fig. 4b). In contrast, the 'open' conformation in chain B is stabilized by crystal packing. Ser52 forms two hydrogen bonds to the backbone N and O atoms of Phe136 of chain A and Lys53 forms a salt bridge to Glu167 of chain A (Fig. 4c). In addition, the main chain of Gly56 forms hydrogen bonds to Asn62 of strand 3.

3.4. Structural comparison with BS-YsxC

Crystal structures of YsxC from *B. subtilis* (BS-YsxC) have been determined in apo, GDP-bound and GMPPNP-bound forms (PDB codes 1sul, 1svi and 1svw, respectively) by Ruzheinikov *et al.* (2004). The structures of TM-YsxC superimpose well with those of BS-YsxC (C^α r.m.s.d.s of 0.8–1.2 Å). Major structural changes are found in the switch I loop (Fig. 5a). The conformation of the 'closed' form in GDP–TM-YsxC closely resembles that in GMPPNP–BS-YsxC (Fig. 5). In both cases, the α-phosphate group and the conserved residues Asn33 and Asn41 form hydrogen bonds to the main-chain

atoms of the G2 loop (Fig. 5b). The major difference is that in the GMPPNP–BS-YsxC complex the invariant threonine residue of the PGxT motif moves towards the bound nucleotide and forms hydrogen bonds to the γ-phosphate group and the bound magnesium ion (Fig. 5b).

4. Discussion

In this study, we determined the crystal structure of the apo and GDP-bound forms of YsxC from *T. maritima*. Since the switch I loop is disordered in the GDP–BS-YsxC complex, the structures of GDP–TM-YsxC provide novel structural insights into the conformational changes upon the hydrolysis of GTP to GDP. The structures of GDP–TM-YsxC reported in this study suggest that the switch I loop can exist in both an 'open' conformation and a 'closed' conformation in the GDP-bound state (Fig. 4b). It is likely that the 'closed' conformation is in exchange with the 'open' conformation, which may represent a structural mechanism inherent in YsxC for the diffusion of GDP out of the nucleotide-binding site after GTP hydrolysis. The disorder of the switch I loop in GDP–BS-YsxC also suggests that the region is flexible and may undergo conformational exchange.
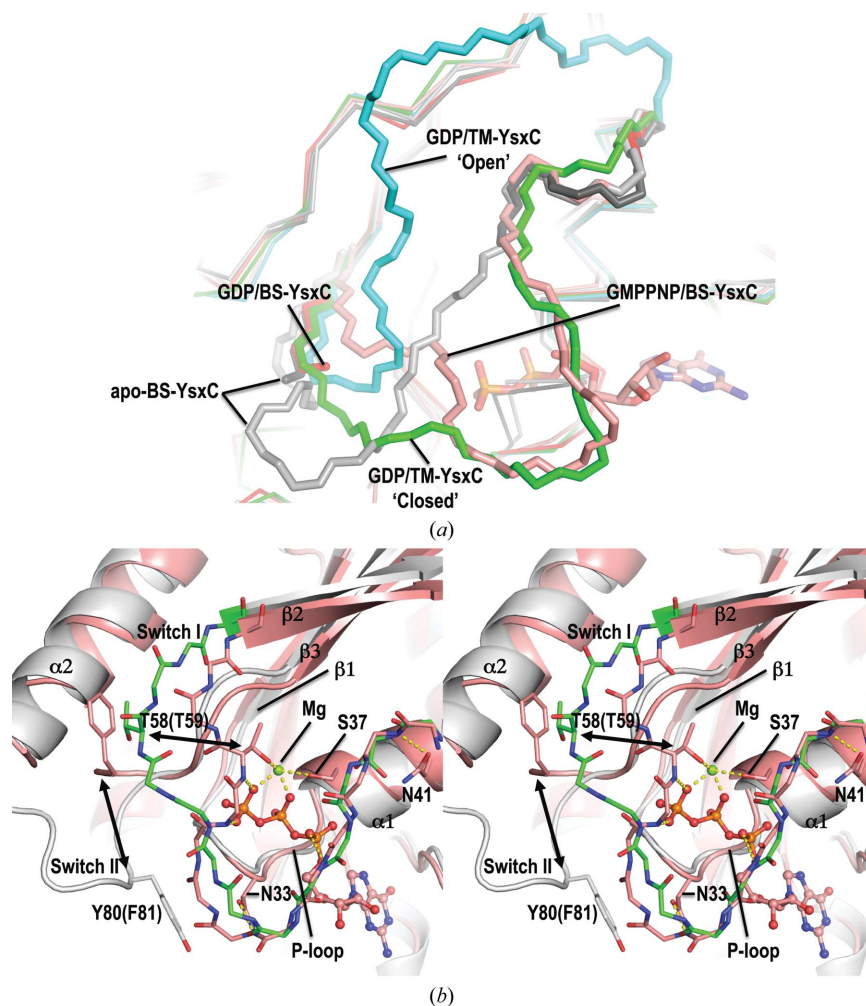

Figure 5
(a) Structural comparison of TM-YsxC and BS-YsxC reveals major conformational changes in the switch I loop. Structures of the GDP–TM-YsxC complex (chain A, green; chain B, cyan) were superimposed with structures of BS-YsxC (apo form, light and dark grey; GDP-bound form, red; GMPPNP-bound form, salmon). The switch I loop is disordered in GDP–BS-YsxC. (b) Stereo diagram showing a comparison of the 'closed' conformation of GDP-bound TM-YsxC (green) and GMPPNP-bound BS-YsxC (salmon). In the GMPPNP-bound form the invariant threonine residue in the switch I loop (Thr58 in TM-YsxC or Thr59 in BS-YsxC) moves towards the γ-phosphate group and interacts with the magnesium ion required for catalysis.

structural communications

However, our results do not exclude the possibility that an exchange factor accelerates the nucleotide-release process.

In GTPases, hydrolysis of GTP often triggers conformational switching from the 'on' state to the 'off' state. Ruzheinikov and coworkers observed conformational changes in the switch II loop between GMPPNP-bound and GDP-bound forms of BS-YsxC. The structure of GDP–TM-YsxC reported here illustrates some additional conformational changes in the switch I region (Fig. 5b). How these conformational changes regulate the biogenesis of ribosomes is not known. One possibility is that the conformational changes in the switch I and II regions upon GTP hydrolysis may modulate the binding of YsxC to the 30S and 50S subunits of ribosomes (Wicker-Planquart et al., 2008).

References

Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P. & Adams, P. D. (2010). *Acta Cryst.* D**66**, 1153–1163.

Arigoni, F., Talabot, F., Peitsch, M., Edgerton, M. D., Meldrum, E., Allet, E., Fish, R., Jamotte, T., Curchod, M. L. & Loferer, H. (1998). *Nature Biotechnol.* **16**, 851–856.

Bourne, H. R., Sanders, D. A. & McCormick, F. (1991). *Nature (London)*, **349**, 117–127.

Cheung, Y.-Y., Allen, M. D., Bycroft, M. & Wong, K.-B. (2004). *Acta Cryst.* D**60**, 1308–1310.

Cooper, E. L., García-Lara, J. & Foster, S. J. (2009). *BMC Microbiol.* **9**, 266.

Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. (2004). *Genome Res.* **14**, 1188–1190.

Edgar, R. C. (2004). *Nucleic Acids Res.* **32**, 1792–1797.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Lehoux, I. E., Mazzulla, M. J., Baker, A. & Petit, C. M. (2003). *Protein Expr. Purif.* **30**, 203–209.

Leipe, D. D., Wolf, Y. I., Koonin, E. V. & Aravind, L. (2002). *J. Mol. Biol.* **317**, 41–72.

Prágai, Z. & Harwood, C. R. (2000). *J. Bacteriol.* **182**, 6819–6823.

Ruzheinikov, S. N., Das, S. K., Sedelnikova, S. E., Baker, P. J., Artymiuk, P. J., García-Lara, J., Foster, S. J. & Rice, D. W. (2004). *J. Mol. Biol.* **339**, 265–278.

Schaefer, L., Uicker, W. C., Wicker-Planquart, C., Foucher, A. E., Jault, J. M. & Britton, R. A. (2006). *J. Bacteriol.* **188**, 8252–8258.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Wicker-Planquart, C., Foucher, A. E., Louwagie, M., Britton, R. A. & Jault, J. M. (2008). *J. Bacteriol.* **190**, 681–690.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.