Acta Crystallographica Section F Structural Biology and Crystallization Communications

ISSN 1744-3091

# Cong Liu,<sup>a,b</sup> Dan Li,<sup>a,b</sup> Lars Hederstedt,<sup>c</sup> Lanfen Li,<sup>a,b</sup> Yu-He Liang<sup>a,b</sup>\* and Xiao-Dong Su<sup>a,b</sup>

<sup>a</sup>National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, People's Republic of China, <sup>b</sup>Department of Biochemistry and Molecular Biology, College of Life Sciences, Peking University, Beijing 100871, People's Republic of China, and <sup>c</sup>Department of Cell and Organism Biology, Microbiology, Lund University, Lund SE-22362, Sweden

Correspondence e-mail: liangyh@pku.edu.cn

Received 27 June 2006 Accepted 14 August 2006



© 2006 International Union of Crystallography All rights reserved

# Preparation, crystallization and preliminary X-ray analysis of protein YtlP from *Bacillus subtilis*

*Bacillus subtilis* YtlP is a protein that is predicted to belong to the bacterial and archael 2'-5' RNA-ligase family. It contains 183 residues and two copies of the HXTX sequence motif conserved among proteins belonging to this family. In order to determine the structure of YtlP and to compare it with the paralogue YjcG and identified 2'-5' RNA ligases, the gene *ytlP* was amplified from *B. subtilis* genomic DNA and cloned into expression vector pET-21a. The soluble protein was produced in *Escherichia coli*, purified to homogeneity and crystals suitable for X-ray analysis were obtained. The crystal diffracted to 2.0 Å and belonged to space group  $P2_12_12_1$ , with unit-cell parameters a = 34.16, b = 48.54, c = 105.75 Å.

# 1. Introduction

The Bacillus subtilis ytlP gene encodes a 183-amino-acid residue protein of so far unknown physiological function. The primary sequence of YtlP is similar to that of proteins belonging to the 2'-5' RNA-ligase family (Fig. 1). The 2'-5' RNA-ligase activity, capable of joining Saccharomyces cerevisiae tRNA-splicing intermediates in the absence of ATP to form a 2'-5' phosphodiester linkage, was first detected in extracts from Escherichia coli (Greer et al., 1983). The E. coli enzyme was subsequently identified and similar enzymes were detected in Methanococcus jannaschii and B. stearothermophilus (Arn & Abelson, 1996). The primary sequence identity among members of the 2'-5' RNA-ligase family is relatively low, but they all contain two copies of the HXTX motif (where X is a hydrophobic residue). These highly conserved motifs are characteristic of the bacterial and archeal 2'-5' RNA-ligase family (Kato et al., 2003). The first structure of this family was that from Thermus thermophilus HB8, which showed a high structural similarity to Arabidopsis thaliana cyclic phophodiesterase (Kato et al., 2003). The crystal structure of 2'-5' RNA ligase from another archeal species, Pyrococcus horikoshii, was determined and appeared to have a substrate preference that probably differed from that of T. thermophilus (Rehse & Tahirov, 2005). A putative integral membrane-transport protein from P. furiosus also seems to belong to the 2'-5' RNA-ligase family. The <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N resonance assignments of a soluble part of this protein have been reported and the structure has been deposited in the PDB (Okada et al., 2006). The putative 2'-5' RNA ligase YjcG from B. subtilis is a member of the YjcG-like subgroup of the 2H phosphoesterase superfamily (Mazumder et al., 2002). Proteins belonging to this group have been suggested to react on unknown molecules with 2'-3' cyclic phosphoester linkages. To date, the structures of the enzymes A. thaliana CPDase (Hofmann et al., 2000), rat CNPase (Kozlov et al., 2003) and human CNPase (Sakamoto et al., 2005) from this family have been reported. All the structures showed conserved bilobal arrangement of two  $\alpha\alpha\beta$  modules. In *B. subtilis*, only YjcG and YtlP have two copies of the conserved HXTX motif. The preliminary X-ray analysis of YjcG has been published and the structure is to be released in the PDB (Li et al., 2005). The structures of YtlP and YjcG should help us to understand the function of YjcGlike proteins in bacteria.

### 2. Materials and methods

# 2.1. Gene cloning and expression

The *ytlP* gene was amplified by polymerase chain reaction (PCR) from B. subtilis strain 168 genomic DNA. The primers used were 5'-CGCGGATCCATGCCAGATATACGGCCTCATT-3' and 5'-CCGCCTGAGAGATAATTGAAATTTGAATATTTCC-3', which contain sequences for a BamHI and an XhoI restriction site, respectively. The PCR product was cloned into the corresponding sites in pET-21a (Novagen). The resulting plasmid, pLPL4, was transformed into E. coli strain BL21(DE3). Cells were grown at 310 K in Luria–Bertani medium containing 50  $\mu$ g ml<sup>-1</sup> kanamycin until mid-exponential growth phase. At this point, ytlP expression was induced by addition of isopropyl  $\beta$ -D-thiogalactopyranoside to a final concentration of 0.5 mM and the culture was grown at 291 K for 20 h. Cells were harvested by centrifuguation at 6700g for 10 min at 277 K and were then suspended in lysis buffer containing 20 mM Tris-HCl, 0.5 M NaCl, 1 mM PMSF, 1  $\mu$ g ml<sup>-1</sup> DNAse, 10  $\mu$ g ml<sup>-1</sup> RNAse pH 7.5.

### 2.2. Protein purification

The cells were lysed by sonication. The lysate was centrifuged at 34700g for 50 min at 277 K. The supernatant was applied onto a 1 ml HiTrap Ni column (GE Healthcare) equilibrated with buffer A containing 20 mM Tris-HCl, 0.5 M NaCl pH 7.5. Unbound proteins

were eluted using buffer *A* and loosely bound proteins were eluted using 5% buffer *B* (20 m*M* Tris–HCl, 0.5 *M* NaCl, 0.5 *M* imidazole pH 7.5) in buffer *A*. Tightly bound proteins were then eluted with buffer *B*. Fractions containing YtlP were further purified using a HiLoad Superdex 75 column (GE Healthcare) with 150 m*M* NaCl in 20 m*M* Tris–HCl pH 7.5 as eluant. The purity of the protein was examined by SDS–PAGE.

# 2.3. Crystallization

Crystallization was performed at 289 K using the hanging-drop vapour-diffusion method. Purified YtlP was concentrated to 16.6 mg ml<sup>-1</sup> by ultrafiltration (Millipore Amicon) and Crystal Screen I, Crystal Screen II and Index kits (Hampton Research) were used for initial screening. A mixture of 1  $\mu$ l protein solution and 1  $\mu$ l reservoir solution was equilibrated against 500  $\mu$ l reservoir solution.

# 2.4. Data collection

X-ray diffraction data were collected on a Bruker–Nonius FR591 rotating-anode generator with Cu  $K\alpha$  radiation and a Bruker SMART 6000 CCD detector. The crystal was flash-cooled in the presence of glycerol and maintained at 100 K in a nitrogen-gas stream during data collection. The data were processed with the Bruker *PROTEUM* online software.

	: : *:** :	
Bs YtlP 100%	MPDIRPHYFIGVPIPEGIANPIYOAAKNEPI-LTFOKWVHPLDYHITIIFLGAADETOIKKLEGSLAEIASEIDP	74
B.lichen 44.8%	MS-VKPHYFIGIPIPAELSDPLYOTVSSRPE-FSFHRWYHPLDYHITUVFLGSAEKEOLAMLAGSLKLISDETAP	73
B. cereus 33.8%	MEPHYEVATTI, PNHIKEVI, SNYREEMOEKI, PERSWYHKEDYHITTI, SFLGSATEEOI, EGIONGLOTI, T-ETKE	71
B anthr 33 3%	MEDHYFYA TTI, DNHIKEVI, SNYREEMOEFI, DERSWYHKEDYH TTI, SFLGSATEFOLEGI, ONGLOTI, I. ETKE	71
B thur 33 39	MUNDELVENA TTI DNUTKEVI CNVKEDTNEKI DED CW/WEDTUTTI CEL COATEGOI KCTKNCI OTI I	73
P woihon 32 7%	MUDUVEVA UT DOUTVETI CHVVEPNODI DED CHVVEPDVUTTI CELCCATECA TOCATO	71
C hours 21 6%		71
G.Kaus SI.0%		75
E.COII 20.2%	-MSEPORTFAIDLEARTREOTINWRATHFP-PERGRPVAADNLHLITAFLGEVSAEKEKALSLLAGKIRQ	69
1.thermo 22.4%	MRLFIAVFLPEEVRAALVEAQIKVRP-FRGWRPVPPPQLHLIILFLGERPEELPDILALGHRLARLEAP	69
P.norik 22.4%	MRAFIAIDVNESVRDSLVRAODYIGSKEAKIKFVERENLHITIKFLGEITEEOAEEIKNILKKIAEKYKK	70
P.turi 22.4%	MRAFIAIDVSESVRDALVRAQDYIGSKEAKIKFVERENFHITLKFLGEITEEQAEEIKKILEKIAKKYKK	70
Bs_YjcG 17.4%	MKYGIVLFPSKKLQDLANSYRKRYDPSY-SLIPPHLTLRASFECAEEKADQLVSHLRNIAKESHP	64
ruler	110203040 <u></u> 506070	
	: : : : : : : : : : : : : : : : : :	
Bs YtlP 100%	FSIKFGKIDVFGDRRKPRVLHLEPKKNKTLDRLREHTKQAVLQAGFQVEKRPYHPHMTLARKWTGEDGFP	144
B.lichen 44.8%	FOLAIKRFGTFGKPSEPRILHVEPEYSAALFHVREOVKHAADAAGFEIEKRPFHPHVTVARKWNGSASFN	143
B.cereus 33.8%	LSFKLOGFSTFGMEDRPRIFWSKVSENODLFOLOKOVHAICEGNGFSLETRPYHPHITVARKWVGEEKFN	141
B.anthr 33.3%	LSFTLOGFSTFGMEDRPRIFWAKVSENODLFPLOKOVHALCEGNGFSLETRPYHPHITVARKWVGEEKFN	141
B. thur 33.3%	LSETLOGESTEGREDOPRILWAKVSENHDLEOLOKOVHALCEENGESLETRPYHPHITVARKWIGEEKEH	143
B weihen 32.7%	LSETLOGI.STEGOVDOPRTEWAGTSDNODLESLOKOVHTICEENGESLDTRPYHPHITVARKWYGEKVED	141
G kaug 31 6%	FDIVIACICTECEPTADDIEWOCVEME	143
E coli 20 2%		130
T thermo 22 4%		144
D howit 22.4%		120
P.HOLIK 22.4%		120
P.IUII 22.46		107
BS_IJCG 17.4%	LVLKMTKISSFAPVNNVITIKAEPTBELKTINEKLYTGVLAGEQEINFVPHVTVGQNLSDDEH	127
ruler	$\dots 80$	
Bs_YtlP#100%	AHVPFESGEVSMMAERFSLFQIHLNQ-SPKYEEIFKFQLS 183	
B.lichen#44.8%	GRLEPLEDAVSFTARHFTIFQTHLDR-LPKYEHQAVFHFGERPAAEKQDSEMNADGTDY 201	
B.cereus#33.8%	LEHIKEVPEISFQADTITLYESHVKE-TPKYKAIAEIKLQK181	
B.anthr#33.3%	LEHIKEVPEISFQADTITLYESHVKE-TPKYKAIAEIKLQK 181	
B.thur#33.3%	LESIKEVPEISFQADTITLYESHVKE-TPKYKSIAEIKLQK 183	
B.weihen#32.7%	LTNVKELPVTSFOADTITLYESHVKE-TPKYKSITEIKLOK 181	
G.kaus#31.6%	PDRLRSLAAAKAVFŠVPEIVLYRTNMER-TPKYEAIAVFPLLGAPMNGRK 192	
E.coli#20.2%	PPGFNWSYAVTEFTLYASSFARGRTRYTPLKRWALTO 176	
T.thermo#22.4%	PVLFGLEWPVEGFALVRSELKPKGPVYTVLEKFSLRGEHGREOAOGPGERPEGD- 198	
P.horik#22.4%	LTMKLKELANEDFGSFVVDATELKKSTLTPKGPTVETLARFELSE	
P.furi#22.4%	LAMKLKELANEDEGSETVEATELKKSTLTPKGPTVETLAREELSE	
Be VicG#17 4%	SDVLGOLK - MORVSHEETVDEEHLLVOLE NGSWTVVETELLGEGE 171	
20_1]00#1/.1% rular		

#### Figure 1

Multiple alignment of amino-acid sequences related to YtlP. The two conserved HXTX motifs in the polypeptides are highlighted by red boxes. The percentages indicate the sequence identity of the paralogues to YtlP. The alignment was performed using the program *ClustalX* (Thompson *et al.*, 1997). Bs\_YtlP, *B. subtilis* YtlP; B.lichen, *B. licheniformis*; B.cereus, *B. cereus*; B.anthr, *B. anthracis*; B.thur, *B. thuringiensis*; B.weihen, *B. weihenstephanensis*; G.kaus, *Geobacillus kaustophilus*; E.coli, *E. coli*; T.thermo, *T. thermophilus*; Phorik, *Pyrococcus horikoshii*; P.furi, *P. furiosus*; Bs\_YjcG, *B. subtilis* YjcG.



### Figure 2

Crystals of *B. subtilis* YtlP. The dimensions of the crystal are approximately  $0.1 \times 0.2 \times 0.05$  mm.

#### 3. Results

The sequence of the cloned *ytlP* gene in plasmid pLPL4 was confirmed by DNA sequencing (AuGCT Biotechnology). YtlP produced in *E. coli* strain BL21 (DE3) containing pLPL4 was water-soluble and could be purified to homogeneity in two steps. SDS-PAGE showed that at least 95% of the protein in the preparations was YtlP, with an apparent molecular weight of about 25.0 kDa. This weight matches the theoretical molecular weight of the recombinant YtlP (21.0 kDa) plus about 4.0 kDa fusion peptide including the hexahistidyl tag.

Microcrystals were obtained under several conditions. After optimizing the concentration of the precipitant and the pH value of the buffer of the initial crystallization conditions, large single crystals were obtained in 0.1 *M* HEPES pH 8.0, 30%( $\nu/\nu$ ) Jeffamine ED-2001 pH 7.0. The crystals grew to dimensions of  $0.1 \times 0.2 \times 0.05$  mm in 3 d (Fig. 2). The final crystallization was carried out by the hanging-drop vapour-diffusion method; the volumes of precipitant, protein solution and reservoir solution were kept the same as in the initial screens. The crystal diffracted to a resolution of 2.0 Å and belonged to space group  $P2_12_12_1$ , with unit-cell parameters a = 34.16, b = 48.54, c = 105.75 Å. For one molecule per asymmetric unit, the Matthews coefficient is 2.1 Å<sup>3</sup> Da<sup>-1</sup> and the solvent content is 40.6% (Matthews, 1968). Data statistics are listed in Table 1. The completeness of the diffraction data is relatively low, which is a consequence of the square shape of the detector. Although the similarities between 2'-5' RNA ligases are

# Table 1

Data-collection statistics for YtlP.

Values in parentheses are for the highest resolution shell.

Resolution (Å)	50-2.0 (2.11-2.0)
Completeness (%)	92.3 (79.0)
$R_{\rm sym}$ † (%)	7.7 (15.8)
Mean $I/\sigma(I)$	9.1 (4.4)
Space group	$P2_{1}2_{1}2_{1}$
Unit-cell parameters (Å)	a = 34.16, b = 48.54, c = 105.75
No. of observed reflections	89925
No. of unique reflections	11565
Molecules per ASU	1
$V_{\rm M}$ (Å <sup>3</sup> Da <sup>-1</sup> )	2.1

†  $R_{\rm sym} = \sum |I_{\rm obs} - I_{\rm avg}| / \sum I_{\rm obs}$ , where the summation is over all reflections.

low, several initial phasing trials using molecular replacement were carried out. However, no positive results were obtained. Since there are four methionine residues in the 183-amino-acid YtlP, phase determination using the anomalous signal from SeMet is in progress.

This work was supported by a grant from the National High Technology and Development Program of China (863 program 2002BA711A13) and grant 621-2004-2762 from the Swedish Science Research Council to LH. Grants 985 and 211 from Peking University are greatly acknowledged. YHL is the recipient of a Fok Ying Tong Education Foundation grant (94017).

### References

- Arn, E. A. & Abelson, J. N. (1996). J. Biol. Chem. 271, 31145-31153.
- Greer, C. L., Javor, B. & Abelson, J. (1983). Cell, 33, 899-906.
- Hofmann, A., Zdanov, A., Genschik, P., Ruvinov, S., Filipowicz, W. & Wlodawer, A. (2000). *EMBO J.* **19**, 6207–6217.
- Kato, M., Shirouzu, M., Terada, T., Yamaguchi, H., Murayama, K., Sakai, H., Kuramitsu, S. & Yokoyama, S. (2003). J. Mol. Biol. 329, 903–911.
- Kozlov, G., Lee, J., Elias, D., Gravel, M., Gutierrez, P., Ekiel, I., Braun, P. E. & Gehring, K. (2003). J. Biol. Chem. 278, 46021–46028.
- Li, D., Chan, C., Liang, Y.-H., Zheng, X., Li, L. & Su, X.-D. (2005). Acta Cryst. F61, 496–498.
- Matthews, B. W. (1968). J. Mol. Biol. 33, 491-497.
- Mazumder, R., Iyer, L. M., Vasudevan, S. & Aravind, L. (2002). Nucleic Acids Res. 30, 5229–5243.
- Okada, K., Matsuda, T., Sakamoto, T., Muto, Y., Yokoyama, S., Kanai, A. & Kawai, G. (2006). In the press.
- Rehse, P. H. & Tahirov, T. H. (2005). Acta Cryst. D61, 1207-1212.
- Sakamoto, Y., Tanaka, N., Ichimiya, T., Kurihara, T. & Nakamura, K. T. (2005). J. Mol. Biol. 346, 789–800.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). Nucleic Acids Res. 25, 4876–4882.