Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 360 (2007) 784–790
BBRC
www.elsevier.com/locate/ybbrc

Crystal structure of YrrB: A TPR protein with an unusual peptide-binding site

Dohyun Han [a,1], Jongkil Oh [a,1], Kyunggon Kim [a], Hyosun Lim [a], Youngsoo Kim [a,b,*]

[a] College of Medicine, Seoul National University, Yongon-Dong, Seoul 110-799, Republic of Korea
[b] Cancer Research Institute, Seoul National University, Yongon-Dong, Seoul 110-799, Republic of Korea

Received 21 June 2007
Available online 5 July 2007

Abstract

YrrB is a hypothetical protein containing a tetratricopeptide repeat (TPR) domain from a Gram-positive bacterium, *Bacillus subtilis*. We determined YrrB structure in the C2 space group to 2.5 Å resolution, which is the first TPR structure of the Gram-positive bacterium *B. subtilis*. In contrast to other known TPR structures, the concave surface of the YrrB TPR domain is composed of the putative peptide-binding pocket lined with positively-charged residues. This unique charge distribution reveals that YrrB can interact with partner proteins via an unusual TPR-mediated interaction mode, compared to that of other TPR-containing structures. Functional annotation using genomics analysis suggested that YrrB may be an interacting mediator in the complex formation among RNA sulfuration components. No proteins containing a TPR domain have been identified in the biosynthesis of sulfur-containing biomolecules. Thus, YrrB could play a new role as a connecting module among those proteins in the conserved gene cluster for RNA sulfuration.

Keywords: YrrB; TPR; Tetratricopeptide repeat; Crystal structure; Protein interaction

A hypothetical protein from the genome of *Bacillus subtilis* 168M, YrrB, is a 23 kDa protein composed of 206 amino acids. The YrrB protein is predicted to contain the tetratricopeptide repeat (TPR) domain. TPR motif is found in almost all organisms, ranging from bacteria to humans [1]. A TPR unit consists of a 34-residue repeating motif, which adopts a helix-turn-helix formation known to be associated with protein–protein interaction [2]. The highly conserved positions of a TPR unit are located at the positions Trp4, Leu7, Gly8, Tyr11, Ala20, Phe24, Ala27, and Pro32. The consensus conserved residues of a TPR unit are mainly hydrophobic amino acids, which are located in the inner groove, and which accordingly are commonly involved in binding interactions that target molecules [1,3]. The TPR domain structure mainly forms a superhelix of repeating anti-parallel α-helices. This TPR superhelix creates concave and convex surfaces at the inner and outer faces, respectively [4]. These concave and convex surfaces impart flexibility and variety to the TPR domain, which makes possible the binding of diverse ligands [5]. YrrB is thought to play a role in the protein–protein interaction and complex formation, but exact functions are not known. No ligand interactions for YrrB have yet been annotated.

In order to further elucidate the function of YrrB and compare to other TPR proteins, we cloned, overexpressed, and determined the three-dimensional X-ray structure of YrrB. The X-ray structure allowed us to understand the unique binding site of YrrB among known TPR structures. In addition, we performed functional analysis to annotate one of the probable functions of YrrB through genetic analysis of the conserved gene cluster. Finally, we found evidence to suggest that YrrB is an interacting mediator in the complex formation among RNA sulfuration

* Corresponding author. Address: College of Medicine, Seoul National University, Yongon-Dong, Seoul 110-799, Republic of Korea. Fax: +82 2 741 0253.
E-mail address: biolab@snu.ac.kr (Y. Kim).
[1] These authors contributed equally to this work.

0006-291X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2007.06.129

components, RNA processing components, and amino-acyl-tRNA synthetases.

Materials and methods

Cloning YrrB constructs. The full DNA sequence of YrrB (amino acid residues 1–206) was obtained by direct PCR amplification of the genomic DNA of *B. subtilis*168 M (ATCC No. 27370), obtained from the American Type Culture Collection (ATTC, Manassas, VA, USA). The primers contained, in the forward primer, an EcoR1 restriction site (5′-AT A*GAATTC*GAAGGCGATTA-3′), and in the reverse primer, an Xho1 site (5′-ATC*CTCGAG*ATCAATCAGTTT-3′). The PCR reaction amplified the DNA sequence for residues 3–204 (the residues 1–2 and 205–206 of the full YrrB sequence were not included), which each contained restriction sites.

The prepared pET28a(+) vector containing the N-terminal residues MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEF and the C-terminal residues LGHHHHHH at both ends of the vector was ligated to the purified PCR product. The ligated pET28a(+)-yrrB vector was transformed into *Escherichia coli* DH5α (Novagen, Madison, WI, USA).

Expression and purification. To prepare selenomethionine (Se-Met)—substituted YrrB protein, methionine auxotroph *E. coli* B834 (DE3) (Novagen, Madison, WI, USA) containing the pET28 a(+)-yrrB vector were grown in M9 medium (20 ml) supplemented with kanamycin (50 μg/ml), 0.4% glucose, methionine (50 mg/ml), $MgSO_4$ (2 mM), $CaCl_2$ (0.1 mM), and Thiamin (1 mg/ml). The overnight seed culture (20 ml) was harvested and resuspended in the M9 medium mixture (500 ml) supplemented with seleno-L-methionine (60 mg/ml) (Sigma, St. Louis, Missouri, USA). The experimental conditions and procedures for induction and purification of Se-Met labeled YrrB were the same as described in the preparation of Se-Met labeled PILF [6]. Purified YrrB was then dialyzed in 20 mM Tris–HCl (pH 7.5), 100 mM NaCl, 3 mM β-mecaptoethanol, and 0.1 mM PMSF and concentrated to 10 mg/ml.

Crystallization. The initial YrrB crystals were grown by the sitting drop vapor diffusion method. The drop consisted of 2.0 μl of protein at 10 mg/ml mixed with 2.0 μl of reservoir solution (4% (w/v) PEG 6000, 20% glycerol, 40 mM KH_2PO_4). The droplets were equilibrated against 400 μl of the reservoir solution. The YrrB crystals were obtained, after one day at 21 °C, in the size of $0.1 \times 0.1 \times 0.1$ mm.

Data collection and phasing. X-ray diffraction data were collected under single-wavelength anomalous diffraction (SAD) for the Se-Met absorbance peak ($\lambda = 0.9794$ Å) in the Pohang Accelerator Laboratory at 4A MXW Wiggler Beam Line (PAL, Pohang, Kyungbuk, Korea). The YrrB crystals were diffracted to 2.5 Å resolution. The space group was C2, with unit-cell dimensions of $a = 100.9$, $b = 85.1$, $c = 69.0$ Å, $\alpha = \gamma = 90°$, and $\beta = 130.6°$, and with two monomers in an asymmetric unit and a Matthews coefficient of 2.1 $Å^3\ Da^{-1}$ (solvent content = 40.5%) [7]. Data from 50 to 2.5 Å were indexed and scaled using a HKL2000 software package (HKL, Charlottesville, VA). The program *solve/resolve* was used to find the selenium sites and to calculate initial phasing to 2.5 Å resolution [8].

Model building and refinement. Initial model building was mainly performed using COOT program [9]. The complete model contains residues 9–202 at 2.5 Å resolution, missing the N-terminal 36 vector residues (residues MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEF) for purification, the eight YrrB N-terminal residues (residues 1–8), and the four YrrB C-terminal residues (residues 203–206). Interestingly, the inserted residues from the cloning vector pET28a(+) vector were all disordered in the experimental electrondensity map, indicating that the YrrB model structure does not contain any of the vector residues.

The final model structure was refined with the REFMAC 5.0 program using the TLS option [10]. The final refined model contained the YrrB residues 9–202 for the chains A/B, with 54 water molecules in the asymmetric unit.

Analysis of structure. Sequence alignments were performed with ClustalW [11], and structural homologs were identified through the DALI sever (http://www.ebi.ac.uk/dali/) at the European Bioinformatics Institute, EMBL. The structure representations rendered were generated with PYMOL (DeLano Scientific LLC, South San Francisco, CA, USA).

Accession number. The coordinates of the YrrB structure were deposited in the Protein Data Bank under the accession code 2Q7F.

Results and discussion

Overall structure of YrrB

The crystal structure of YrrB contains residues 9–202 (full residues 1–206) and the data collection and refinement statistics are shown in supplement (Table S1). Two YrrB monomers in the asymmetric unit were packed diagonally. Each YrrB monomer was composed of 12 anti-parallel α-helices and turn motifs, forming a right-handed $70 \times 30 \times 25$ Å superhelix (Fig. 1A). The structure of YrrB monomer consists of an extra N-terminal helix (H1), TPR1 (H2, H3), TPR2 (H4, H5), TPR3 (H6, H7), TPR4 (H8, H9), TPR5 (H10, H11), and an extra C-terminal helix (H12) (Fig. 1A) (supplementary Fig. S1).

YrrB structure consists of five TPRs

The five TPR units of YrrB (TPR1–TPR5) are characterized by a pattern of hydrophobic residues conserved in the canonical TPR sequence motif (W4, L7, G8, Y11, A20, F24, A27, and P32), where the numbers indicate residue positions within the canonical TPR motif (Fig. 1A and B). The YrrB superhelix consists of two helical layers that are packed with an inner helical layer (H1, H2, H4, H6, H8, H10, and H12) and an outer helical layer (H3, H5, H7, H9, and H11) forming the inner (concave) and outer (convex) faces, respectively (Fig. 1A) (also see supplementary Fig. S2). The adjacent helices of the two layers are packed tightly with hydrophobic interactions. The interfaces between the TPR units mostly consist of hydrophobic residues forming a hydrophobic interaction, which is represented by the helical wheel of the inner and outer helical layers (supplementary Fig. S2). The packing angle between helices is ~22° within a TPR unit, and the TPR units are piled up, each at an angle of about 72°, which generates a right-handed superhelical shape where one superhelical turn is composed of five TPR units (Fig. 1A).

Structural similarity of YrrB to other TPR proteins

A DALI search (http://www.ebi.ac.uk/dali/) yielded significantly similar structures to YrrB in 418 PDB entries with Z scores larger than 2.0 [12]: O-linked GlcNAc transferase (OGT, PDB entry 1w3b, Z score = 19.4) [13], TPR1 domain of Hsp70/Hsp90 organizing protein (HOP) (PDB entry 1ELW, Z score = 17.5), TPR2A domain of HOP (PDB entry 1ELR, Z score = 16.8) [14], and U-box e3 ubiquitin ligase (CHIP) (PDB entry 2c2l, Z score = 16.6) [15].

Those four TPR domains were aligned using the COOT program [9] (Fig. 1C). The Root Mean Square Deviations