Crystal structure of *Bacillus subtilis* YabJ, a purine regulatory protein and member of the highly conserved YjgF family

J.L. Smith[*§], S. Sinha[*], P. Rappu[†], S.C. Lange[*], P. Mäntsälä[†], and H. Zalkin[‡]
[*]*Department of Biological Sciences, Purdue University, West Lafayette, IN 47907 USA*
[‡]*Department of Biochemistry, Purdue University, West Lafayette, IN 47907 USA*
[†]*Department of Biochemistry and Food Chemistry, University of Turku, Vatselankatu 2, FIN-20014 Turku, Finland*

Introduction

Purine biosynthesis in *Bacillus subtilis* is regulated by the purine operon repressor gene, *purR*. The *purR* operon consists of *purR* and *yabJ*, an open reading frame of unknown function. The repressor PurR binds to control regions upstream of the transcription start sites and regulates transcription of three purine operons. YabJ, the protein product of *yabJ*, affects PurR function *in vivo* by an unknown mechanism.

YabJ also belongs to a widely distributed family of proteins of unknown function. Members of the YER057c/YjgF/UK114 family of proteins have 20–98% pairwise identity and have been reported to influence a variety of biological processes. There is no certain biochemical function for any of them. Homologs are broadly distributed in bacteria, animals and fungi.

Methods and Materials

Multiwavelength anomalous diffraction (MAD) data used for structure determination and refinement were collected from a single, Hg-derivatized, frozen crystal at the BioCARS beamline 14-BM-D. The positions of three Hg atoms were determined by inspection of a Bijvoet difference Patterson map. MAD phasing was performed by the pseudo-MIR (multiple isomorphous replacement) approach. Phases were refined and extended to 2.0 Å by three-fold averaging, histogram matching, and solvent flattening to yield a figure of merit of 0.63 for all reflections to 2.0 Å. The refined model has R_{work} = 0.173 and R_{free} = 0.201 for all data between 30.0 Å and 1.7 Å and is deposited in the Protein Data Bank with accession code 1qd9.

Results and Discussion

The YabJ monomer is a single domain having a six-stranded, mostly antiparallel β-sheet with two α-helices packed against one face of the β-sheet. YabJ is a symmetric trimer in which the β-sheets form a triangular barrel with the six α-helices on the outside. The trimer interface is constructed of alternating hydrophobic and polar regions with a central water-filled cavity. The total monomer accessible surface buried in the trimer is very large, approximately 28% of the total, and includes two extensive hydrophobic contact regions. Residues in the subunit interface are very well conserved among the YabJ homologs, suggesting that all are trimers. Neighboring subunits of the trimer create three deep, narrow clefts on the outer surface of YabJ. The predominance of acidic residues on the surface of YabJ suggests that it does not mediate repression of purine genes by direct interaction with DNA.

YabJ has an unexpected fold and aggregation state in common with chorismate mutase from *Bacillus subtilis*. However, based on the structure alignment, the sequences of YabJ and chorismate mutase are only 8% identical. The three active sites of the chorismate mutase trimer are located in clefts between subunits, analogous to the clefts at the subunit interfaces of the YabJ trimer. However, residues in the active site of chorismate mutase implicated in catalysis do not map to common functional groups in YabJ. YabJ and chorismate mutase may have a common ancestor, but are evolved to perform different functions.

YabJ has a total of 37 homologs with 20–53% sequence identity, indicating a common ancestor, a common fold, and possibly a common function. Among the 24 most similar homologs, there are nine invariant residues, which all map to the narrow, deep cleft between subunits of the trimer, although they are spread throughout the sequence. The clustering of invariant residues in the cleft strongly suggests that the variety of biological functions attributed to YabJ and its homologs is due to a common molecular function of catalysis or binding. This implies recognition of an appropriately shaped small-molecule ligand or a single amino acid residue in a protein ligand. The structure of YabJ will be an important guide for identification of this ligand. This work also illustrates the challenges of deducing protein function from sequence and three-dimensional structure.

Acknowledgments

Supported by National Iinstitute of Health grants DK-42303 to J.L. Smith and GM-24658 to H. Zalkin, and by the Finnish Ministry of Education to P. Mäntsälä. The authors thank the staff of BioCARS at the Advanced Photon Source (APS), Argonne National Laboratory, for their advice during the data collection, and Carol Greski for expert preparation of the report. Use of the APS was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-Eng-38.

Reference

[1] This work has been published: S. Sinha, P. Rappu, S.C. Lange, P. Mäntsälä, H. Zalkin, and J.L. Smith, *Proc. Natl. Acad. Sci. USA* **96**, 13074–13079 (1999).