NMR structure of the hypothetical protein encoded by the YjbJ gene from Escherichia coli

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · JUNE 2002

Impact Factor: 2.63 · DOI: 10.1002/prot.10120 · Source: PubMed

8 AUTHORS, INCLUDING:

Antonio Pineda-Lucena
Centro de Investigación Príncipe Felipe
80 PUBLICATIONS **1,306** CITATIONS
SEE PROFILE

Jack C C Liao
University Health Network
13 PUBLICATIONS **648** CITATIONS
SEE PROFILE

Michael A Kennedy
Miami University
158 PUBLICATIONS **3,507** CITATIONS
SEE PROFILE

PROTEINS: Structure, Function, and Genetics 47:572–574 (2002)

STRUCTURE NOTE

NMR Structure of the Hypothetical Protein Encoded by the YjbJ Gene From *Escherichia coli*

Antonio Pineda-Lucena,[1,2] Jack Liao,[1,2] Bin Wu,[1,2] Adelinda Yee,[1,2] John R. Cort,[3] Michael A. Kennedy,[3,4] Aled. M. Edwards,[1,2,4] and Cheryl H. Arrowsmith[1,2,4]*

[1]*Division of Molecular and Structural Biology, Ontario Cancer Institute, Toronto, Ontario, Canada*
[2]*Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada*
[3]*Enviromental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington*
[4]*Northeast Structural Genomics Consortium*

Introduction. Here we describe the solution structure of YjbJ (gi|418541) as part of a structural proteomics project on the feasibility of the high-throughput generation of samples from *Escherichia coli* for structural studies. YjbJ is a hypothetical protein from *E. coli* of unknown function.[1] It is conserved, showing significant sequence identity to four predicted prokaryotic proteins, also of unkown function [Fig. 1(A)]. These include gi|16762921 from *Salmonella enterica* (*S. typhi*), gi|17938413 from *Agrobacterium tumefaciens,* gi|16265654 from *Sinorizhobium meliloti,* and gi|15599932 from *Pseudomona aeruginosa.* The structure of YjbJ reveals a new variation of a common motif (four-helix bundle) that could not be predicted from the protein sequence. Although the biochemical function is unknown, the existence of patterns of conserved residues on the protein surface suggest that the fold and function of all these proteins could be similar.

Materials and Methods. A recombinant protein consisting of the full sequence of YjbJ (69 amino acids) was expressed in *E. coli* BL21-DE3 cells containing the pET-15b expression vector (Novagen). Cells were grown at 37°C to an OD_{600} of 0.6 and induced with 1 mM IPTG for 5 h at 25°C. The protein was purified to homogeneity by using metal affinity chromatography. Subsequently, the N-terminal tag was removed by using thrombin and benzamidine-sepharose. The purified protein contained the complete sequence of YjbJ plus three additional N-terminal residues (Gly-Ser-His) remaining after proteolytic cleavage of the His_6 affinity tag. U-^{15}N and U-^{13}C,^{15}N samples were produced in standard M9 media supplemented with ^{15}N ammonium chloride (1 g/L) and ^{13}C glucose (2 g/L). ^{15}N-labeled or ^{13}C/^{15}N-labeled protein solution was prepared in 25 mM sodium phosphate (pH = 6.5), 150 mM NaCl, 1 mM DTT, 95% $H_2$0/5% D_2O. The concentration of the purified protein ranged between 1.0 and 1.5 mM.

All NMR spectra were recorded at 25°C on a Varian INOVA 600-MHz spectrometer equipped with pulsed field gradient triple-resonance probes. Linear prediction was used in the ^{13}C and ^{15}N dimensions to improve the digital resolution. Spectra were processed by using the NMRPipe software package[2] and analyzed with XEASY.[3] SPSCAN[4] was used to convert nmrPipe formatted spectra into XEASY. The assignments of the ^{1}H, ^{15}N, and ^{13}C resonances were based on the following experiments: CBCA(CO)NH, HNCACB, CC(CO)NH-TOCSY, HNHA, HC(CO)NH-TOCSY, and HCCH-TOCSY.[5,6] The backbone resonance assignment was achieved mainly by the combined analysis of the HNCACB and CBCA(CO)NH data. The side-chain resonances were identified mainly by the analysis of HCCH-TOCSY. Aromatic ring resonances were assigned on the basis of the analysis of heteronuclear NOESY. In the ^{1}H-^{15}N HSQC, 99% backbone amide resonances were assigned. Of the other backbone resonances, 99% have been assigned for C^α, and 99% for H^α. Moreover, 97% aliphatic side-chains have been assigned for YjbJ.

For structure calculation purposes, a simultaneous ^{15}N- and ^{13}C-NOESY-HSQC[7] (τ_m = 150 ms) was acquired. NOE cross-peak assignment was obtained by using a combination of manual and automatic procedures. An initial fold of the protein was calculated on the basis of unambiguously assigned NOEs, with subsequent refinement using the NOAH module in the program DYANA.[8] Peak analysis of the NOESY spectra were generated by interactive peak picking with the program XEASY. Backbone dihedral restraints were derived from $^1H^\alpha$ and $^{13}C^\alpha$ secondary chemical shifts using TALOS.[9] The program MOLMOL[10] was used to analyze the energy-minimized conformers and to prepare pictures of the structures.

Results and Discussion. YjbJ adopts a four-helix bundle structure [Fig. 1(B)] with residues in all four helices as

Grant sponsor: Ontario Research and Development Challenge Fund; Grant sponsor: National Institutes of Health: Grant number: P50 GM62413-02.

*Correspondence to: Prof. Cheryl Arrowsmith, Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Toronto, ON, Canada, M5G 2M9. E-mail: carrow@uhnres.utoronto.ca

Received 3 January 2002; Accepted 4 January 2002

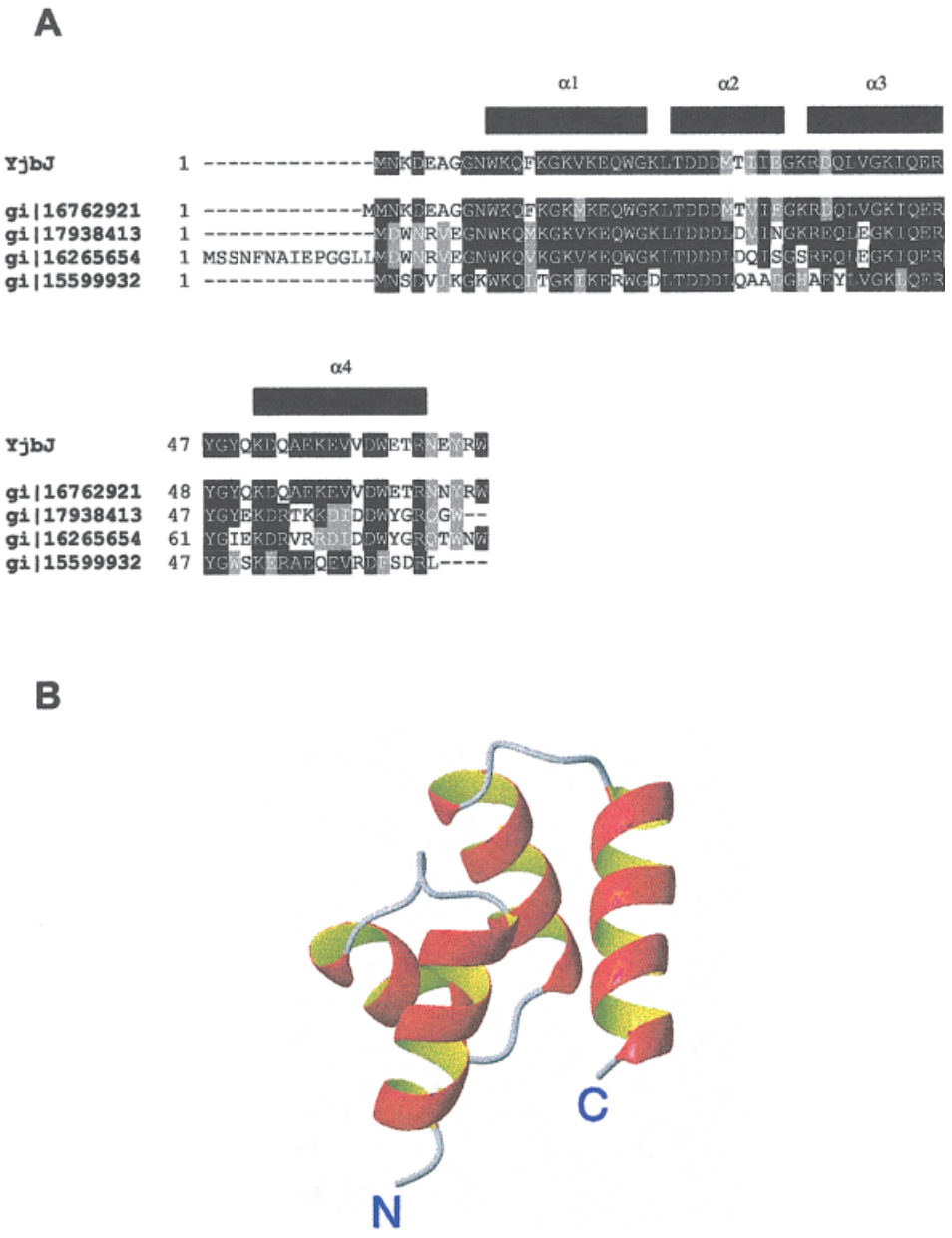

Fig. 1. **A:** Sequence alignment of YjbJ with four predicted prokaryotic proteins, gi|16762921 (*Salmonella enterica, S. typhi*), gi|17938413 (*Agrobacterium tumefaciens*), gi|16265654 (*Sinorizhobium meliloti*), and gi|15599932 (*Pseudomona aeruginosa*). Identical and similar residues are highlighted in black and gray, respectively. Black rectangles correspond to α-helical regions of YjbJ. **B:** Ribbon diagram depicting the averaged minimized NMR structure of YjbJ of *Escherichia coli* (residues 5–69).

well as in the turn regions defining a compact structural domain. Helix α1 extends from residue Trp10 to Gly22, whereas the α2, α3, and α4 helices span residues Thr25-Glu33, Arg36-Arg46, and Lys51-Arg64, respectively.

The three-dimensional structure of YjbJ was determined by using a torsion angle dynamics protocol from a total of 2036 NMR-derived constraints. A superposition of 20 low-energy structures is shown in Figure 2, and the structural statistics are given in Table I. The results obtained for the ordered regions of the protein are virtually identical. This is probably due to the small size of YjbJ and the short length of the loops connecting the ordered regions of the protein. A 3D structure search using DALI[11] showed that YjbJ shares some structural homology to the α-helical regions of the Bchi subunit of magnesium chelatase and T7 DNA polymerse (PDB accession numbers 1g8p and 1t7p, respectively). In both cases, the similarity is based on the existence of four sequential α-helical elements in these proteins, but the spatial orientation and length of these α-helices are very different compared to YjbJ.

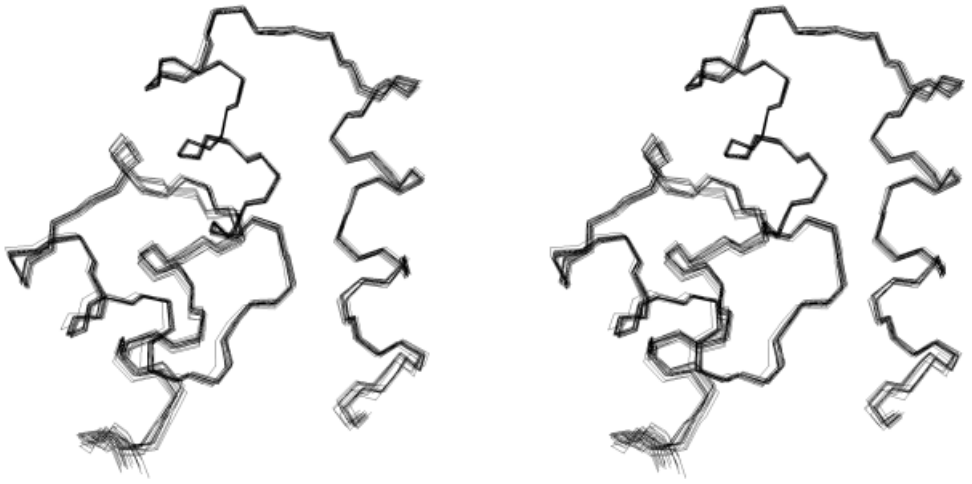

Fig. 2. Stereoview of the backbone (N, C^{α}, C′) of 20 superimposed NMR-derived structures of YjbJ of *E. coli* (residues 5–69).

TABLE I. Structural Statistics for the Ensemble Calculated for YjbJ[†]

Distance restraints	
All	2036
Intraresidue	522
Sequential ($\|i-j\| = 1$)	423
Medium range ($2 \leq \|i-j\| \leq 4$)	628
Long range ($\|i-j\| > 4$)	463
Hydrogen bonds	31×2
Dihedral angle restraints	
All	101
ϕ,ψ	51,50
Pairwise r.m.s.d.	
All residues[a]	
Backbone atoms	0.23 ± 0.10
All heavy atoms	0.99 ± 0.21
Ordered regions[b]	
Backbone atoms	0.23 ± 0.10
All heavy atoms	0.95 ± 0.18
Ramachandran plot	
Residues in most favored regions (%)	86
Residues in additional allowed regions (%)	14
Residues in generously allowed regions (%)	0
Residues in disallowed regions (%)	0

[†]Ensemble of the 20 lowest energy structures out of 100 calculated.
[a]RMSD values for residues 10–65.
[b]Only residues in α-helices are included.

The chemical shifts have been submitted to the BMRB (accession # 5105), and the structure ensemble has been submitted to the PDB (accession # 1JYG).

Acknowledgments. The authors thank A. Semesi for technical assistance. All the spectra were performed at the Environmental Molecular Sciences Laboratory (a national scientific user facility sponsored by the U.S. DOE Office of Biological and Environmental Research) located at Pacific Northwest National Laboratory, operated by Batelle for the DOE. AME and CHA are Scientists of the Canadian Institutes of Health Research.

REFERENCES

1. Blattner FR, Burland V, Plunkett G, Sofia HJ, Daniels DL. Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes. Nucleic Acids Res 1993;21:5408–5417.
2. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293.
3. Bartels C, Xia T, Billeter M, Güntert P, Wüthrich K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. J Biomol NMR 1995;6:1–10.
4. Glaser R, Wüthrich K. http://www.mol.biol.ethz.ch/wuthrich/software/spscan.
5. Bax A, Vuister GW, Grzesiek S, Delaglio F, Wang AC, Tschudin R, Zhu G. Measurement of homo- and heteronuclear J couplings from quantitative J correlation. Methods Enzymol 1994;239:79–105.
6. Kay LE. NMR methods for the study of protein structure and dynamics. Biochem Cell Biol 1997;75:1–15.
7. Pascal SM, Muhandiram DR, Yamazaki T, Forman-Kay JD, Kay LE. Simultaneous acquisition of N^{15}-edited and C^{13}-edited NOE spectra of proteins dissolved in H2O. J Magn Reson 1994;B103:197–201.
8. Güntert P, Mumenthaler C, Wüthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol 1997;273:283–298.
9. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 1999;13:289–302.
10. Koradi R, Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graphics 1996;14:51–55.
11. Holm L, Sander C. DALI: a network tool for protein structure comparison. Trends Biochem Sci 1995;20:478–480.