Solution NMR structure of VF0530 from Vibrio fischeri reveals a nucleic acid-binding function

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · OCTOBER 2011

Impact Factor: 2.63 · DOI: 10.1002/prot.23121 · Source: PubMed

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Solution NMR structure of VF0530 from *Vibrio fischeri* reveals a nucleic acid-binding function

James M. Aramini,[1,2]* Paolo Rossi,[1,2] Markus Fischer,[3,4] Rong Xiao,[1,2] Thomas B. Acton,[1,2] and Gaetano T. Montelione[1,2,5]*

[1] Department of Molecular Biology and Biochemistry, Center for Advanced Biotechnology and Medicine, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854

[2] Northeast Structural Genomics Consortium, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854

[3] Department of Biochemistry and Molecular Biophysics, Center for Computational Biology and Bioinformatics, Columbia University, New York, New York 10032

[4] Northeast Structural Genomics Consortium, Columbia University, New York, New York 10032

[5] Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey 08854

ABSTRACT

Protein domain family PF09905 (DUF2132) is a family of small domains of unknown function that are conserved in a wide range of bacteria. Here we describe the solution NMR structure of the 80-residue VF0530 protein from *Vibrio fischeri*, the first structural representative from this protein domain family. We demonstrate that the structure of VF0530 adopts a unique four-helix motif that shows some similarity to the C-terminal double-stranded DNA (dsDNA) binding domain of RecA, as well as other nucleic acid binding domains. Moreover, gel shift binding data indicate a potential dsDNA binding role for VF0530.

Proteins 2011; 79:2988–2991.

Key words: DUF2132; PF09905; nucleic acid-binding domain; structural genomics.

INTRODUCTION

We present the solution NMR structure of the small (80-residue), basic (pI 9.4) VF0530 protein from *Vibrio fischeri* (UniProtKB/TrEMBL ID, Q5E7H1_VIBF1; NESG ID, VfR117), a member of a conserved prokaryotic Pfam protein domain family PF09905 of unknown function (DUF2132). The structure of VF0530 was determined as part of the efforts of the Protein Structure Initiative to obtain experimental structures of representatives from broadly conserved protein domain families. Our structural and biochemical analyses of VF0530 reveal a potential nucleic acid-binding function.

MATERIALS AND METHODS

Isotopically enriched samples of VF0530 ([*U*-^{13}C,^{15}N]- and [*U*-5%-^{13}C,100%-^{15}N]-VF0530) for NMR spectroscopy were cloned, expressed, and purified following standard protocols of the Northeast Structural Genomics Consortium (NESG)[1]; see Supporting Information for a detailed description of the methods used in this work. VF0530 is a monomer under the conditions used in the NMR studies (20 m*M* ammonium acetate, 100 m*M* NaCl, 10 m*M* DTT, and 5 m*M* $CaCl_2$, pH 4.5, 20°C) based on analytical gel filtration chromatography, static light scattering (Supporting Information Fig. S1), and one-dimensional ^{15}N T_1 and T_2 relaxation data (Supporting Information Fig. S2). Backbone ^{1}H, ^{13}C, and ^{15}N resonance assignments were made using AutoAssign 2.4.0 (Sup-

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: National Institute of General Medical Sciences Protein Structure Initiative (PSI); Grant number: U54-GM094597

*Correspondence to: James M. Aramini, CABM, Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854 Email: jma@cabm.rutgers.edu or Gaetano T. Montelione, CABM, Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854 Email: guy@cabm.rutgers.edu

Received 6 March 2011; Revised 10 June 2011; Accepted 15 June 2011
Published online 15 July 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.23121

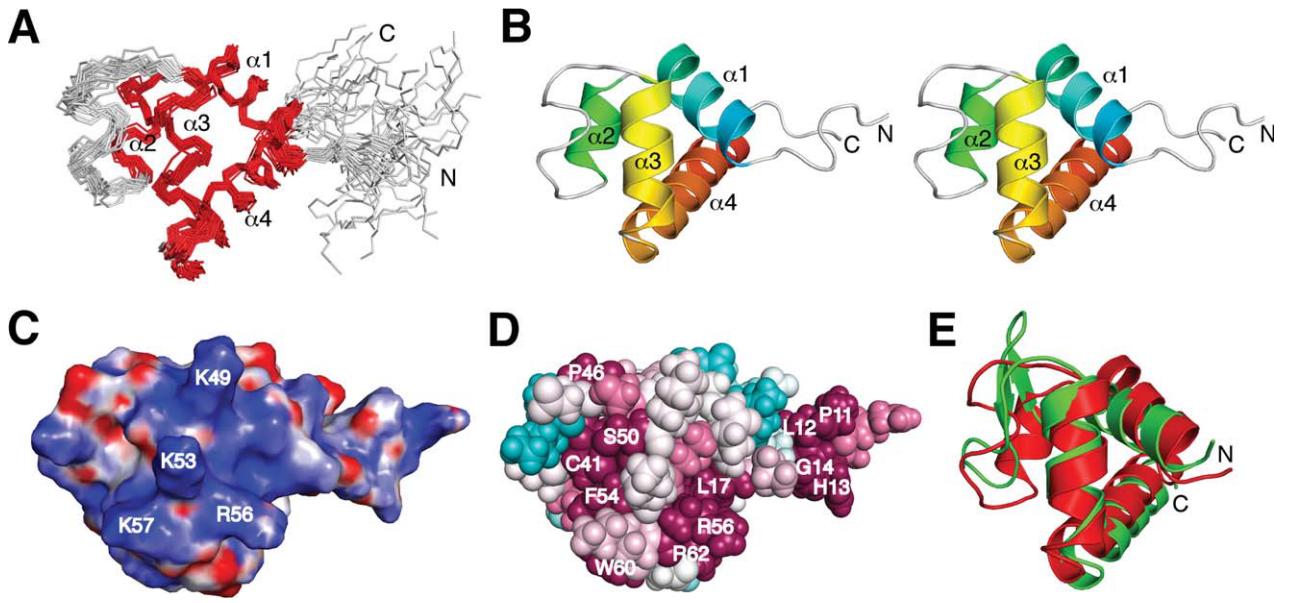

Figure 1

(**A**) Superposition of the final ensemble of 20 conformers from the solution NMR structure of VF0530 (PDB ID, 2JVW). Secondary structural elements are labeled; α-helices are shown in red and loops are colored gray. Note the disorder in the N- and C-terminal tails. For clarity, residues 1–8 have been omitted. (**B**) Stereoview of the lowest energy (CNS) conformer (residues 9–80) of the final solution NMR ensemble of VF0530. Helices are colored and loops are in gray. (**C**) APBS[11] electrostatic surface potential of VF0530 (residues 9–80) showing negative (red), neutral (white), and positive (blue) charges. Selected basic residues are indicated. (**D**) ConSurf[12] image showing the conserved residues in VF0530 (residues 9–80). Residue coloring, reflecting the degree of residue conservation over the entire PF09905 protein domain family (Pfam 24.0), ranges from magenta (highly conserved) to cyan (variable). Selected highly conserved residues are labeled. (**E**) Overlay of the ordered residues (15–75) from the solution NMR structure of VF0530 (red) and the C-terminal domain (residues 282–341) from the X-ray crystal structure of RecA from *Deinococcus radiodurans* (PDB ID, 1XP8[13]; green). All structure figures were rendered using PyMOL (www.pymol.org).

porting Information Figs. S3 and S4)[2]; side chain assignments were completed manually. The solution NMR structure of VF0530 was calculated using CYANA 2.1,[3,4] followed by refinement in explicit water using CNS 1.2.[5,6] NMR resonance assignment validation and structure quality analyses were performed using the AVS,[7] PSVS,[8] MolProbity,[9] and RPF[10] software packages. The final ensemble of 20 models (excluding the C-terminal His_6 tag) and NMR resonance assignments for VF0530 were deposited into the Protein Data Bank (PDB ID, 2JVW) and BioMagResDB (BMRB accession number, 15491), respectively. The pET expression vector for VF0530 (NESG VfR117-21.1) has been deposited in the PSI Materials Repository (http://psimr.asu.edu/).

RESULTS AND DISCUSSION

The solution NMR structure of VF0530 consists of a unique four α-helix bundle (α1, L17-Y28; α2, W30-M36; α3, I48-K57; and α4, D59-H75), comprised of two interleaved antiparallel helix-hairpin-helix motifs tethered by a 10-residue loop and flanked by disordered N- and C-terminal tails [Fig. 1(A,B)]. Structural statistics for VF0530 are presented in Table I. The well-defined core and not well-defined terminal polypeptide segments of the structure correlate well with ^{1}H-^{15}N heteronuclear NOE data for VF0530 (Supporting Information Fig. S4). Electrostatic surface potential[11] [Fig. 1(C)] and ConSurf[12] [Fig. 1(D)] analyses of the VF0530 structure reveal a largely positively charged face featuring several highly conserved residues clustered in helices α3 and α4 and in the loop between helices α2 and α3, strongly suggesting that this is a functionally important region of the molecule. A stretch of conserved residues is also present in the disordered N-terminal tail of the protein.

The solution NMR structure of VF0530 reported here corresponds to the first structural representative from the PF09905 (DUF2132) protein domain family of conserved bacterial proteins with currently unknown function and exhibits very low sequence similarity (<20%) with any other structure in the PDB. Dali[14] and Skan[15] structural alignment analyses reveal that the structure of VF0530 shows some similarity to several reported structures of the C-terminal domain of RecA [i.e., PDB ID, 1XP8[13]; Dali *Z*-score, 5.1; C^{α} RMSD, 2.6 Å (residues 15–75)]. Three of the helices within VF0530 (α1, α3, and α4) overlay well with corresponding helices in this domain, but helix α2 in VF0530 is replaced by two short antiparallel β-strands in RecA [Fig. 1(E)]. This domain of RecA

Table I
Summary of NMR and Structural Statistics for VF0530[a]

Completeness of resonance assignments[b]		
Backbone (%)		98.8
Side chain (%)		95.3
Aromatic (%)		100
Stereospecific methyl (%)		100
Conformationally restricting constraints[c]		
Distance constraints		
Total		1424
Intraresidue ($i = j$)		489
Sequential ($\|i - j\| = 1$)		380
Medium range ($1 < \|i - j\| < 5$)		285
Long range ($\|i - j\| \geq 5$)		270
Dihedral angle constraints		88
Hydrogen bond constraints		46
Number of constraints per residue		19.2
Number of long-range constraints per residue		3.3
Residual constraint violations[c]		
Average number of distance violations per structure		
0.1–0.2 Å		1.2
0.2–0.5 Å		0.05
>0.5 Å		0
Average number of dihedral angle violations per structure		
1–10°		0.1
>10°		0
RMSD from average coordinates (Å)[c,d]		
Backbone atoms		0.5
Heavy atoms		1.1
MolProbity Ramachandran statistics[c,d]		
Most favored regions (%)		96.1
Additional allowed regions (%)		3.7
Disallowed regions (%)		0.2
Global quality scores (Raw/*Z*-score)[c]		
Verify3D	0.33	−2.09
ProsaII	0.69	0.17
Procheck(phi-psi)[d]	0.21	1.14
Procheck(all)[d]	0.21	1.24
Molprobity clash	14.65	−0.99
RPF scores[e]		
Recall/Precision	0.989	0.927
F-measure/DP-score	0.957	0.808

[a]Structural statistics were computed for the ensemble of 20 deposited structures (PDB ID, 2JVW).
[b]Computed using AVS software[7] from the expected number of peaks, excluding: highly exchangeable protons (N-terminal, Lys, and Arg amino groups, hydroxyls of Ser, Thr, and Tyr), carboxyls of Asp and Glu, nonprotonated aromatic carbons, and the C-terminal His_6 tag.
[c]Calculated using PSVS 1.4 program.[8] Average distance violations were calculated using the sum over r^{-6}.
[d]Ordered residue ranges [S(phi) + S(psi) > 1.8]: 15–37, 46–75.
[e]RPF scores[10] reflecting the goodness-of-fit of the final ensemble of structures (including disordered residues) to the NOESY data and resonance assignments.

is thought to bind double-stranded DNA (dsDNA) via residues in the loop preceding its second helix.[16] In the structure of VF0530, this loop flanks the highly basic face formed by helices α3 and α4, suggesting that VF0530 may also be a nucleic acid-binding protein. Moreover, the VF0530 structure is remotely similar to structures of several protein/dsDNA complexes in the PDB, including that of the MH1 domain of the Smad3 transcription factor (PDB ID, 1OZJ).[17] Interestingly, several of the basic residues in helices 3 and 4 of VF0530 map well to residues in the positively charged dsDNA-binding surface of this Smad3 domain (Supporting Information Fig. S5). Indeed, gel shift binding experiments demonstrate an interaction between VF0530 and dsDNA (Supporting Information Fig. S6). Finally, the structure of VF0530 also resembles structures of the six-helix bundle pyrin domain (i.e., PDB ID, 2HM2),[18] which plays a role in inflammation and apoptosis.

In summary, we have determined the solution NMR structure of VF0530 from *Vibrio fischeri*, the first structure of a member of the conserved PF09905 (DUF2132) bacterial protein domain family. The all-helical structure of VF0530 shows similarity to the structures of several dsDNA-binding proteins, suggestive of distant homologous relationships and a potential dsDNA-binding function. Elucidation of the exact function of VF0530 in this symbiotic marine bacterium awaits further study.

ACKNOWLEDGMENTS

The authors thank B. Honig, H. Janjua, C. Nwosu, L. Owens, B. Rost, G. V. T. Swapna, and D. Wang for valuable scientific discussions and/or technical support.

REFERENCES

1. Acton TB, Xiao R, Anderson S, Aramini J, Buchwald WA, Ciccosanti C, Conover K, Everett J, Hamilton K, Huang YJ, Janjua H, Kornhaber G, Lau J, Lee DY, Liu G, Maglaqui M, Ma L, Mao L, Patel D, Rossi P, Sahdev S, Shastry R, Swapna GVT, Tang Y, Tong S, Wang D, Wang H, Zhao L, Montelione GT. Preparation of protein samples for NMR structure, function, and small-molecule screening studies. Methods Enzymol 2011;493:21–60.
2. Moseley HNB, Monleon D, Montelione GT. Automatic determination of protein backbone resonance assignments from triple resonance nuclear magnetic resonance data. Methods Enzymol 2001;339:91–108.
3. 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.
4. Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol 2002;319:209–227.
5. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D 1998;54:905–921.
6. Linge JP, Williams MA, Spronk CAEM, Bonvin AMJJ, Nilges M. Refinement of protein structures in explicit solvent. Proteins 2003;50:496–506.
7. Moseley HNB, Sahota G, Montelione GT. Assignment validation software suite for the evaluation and presentation of protein resonance assignment data. J Biomol NMR 2004;28:341–355.
8. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. Proteins 2007;66:778–795.
9. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, III, Snoeyink J, Richardson JS, Richard-

son DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 2007;35:W375–W383.
10. Huang YJ, Powers R, Montelione GT. Protein NMR recall, precision, and F-measure scores (RPF scores): structure quality assessment measures based on information retrieval statistics. J Am Chem Soc 2005;127:1665–1674.
11. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA 2001;98:10037–10041.
12. Glaser F, Pupko T, Paz I, Bell RE, Bechor-Shental D, Martz E, Ben-Tal N. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics 2003;19:163–164.
13. Rajan R, Bell CE. Crystal structure of RecA from *Deinococcus radiodurans*: insights into the structural basis of extreme radioresistance. J Mol Biol 2004;344:951–963.
14. Holm L, Kääriäinen S, Rosenström P, Schenkel A. Searching protein structure databases with DaliLite v.3. Bioinformatics 2008;24:2780–2781.
15. Petrey D, Fischer M, Honig B. Structural relationships among proteins with different global topologies and their implications for function annotation strategies. Proc Natl Acad Sci USA 2009;106:17377–17382.
16. Aihara H, Ito Y, Kurumizaka H, Terada T, Yokoyama S, Shibata T. An interaction between a specified surface of the C-terminal domain of RecA protein and double-stranded DNA for homologous pairing. J Mol Biol 1997;274:213–221.
17. Chai J, Wu J-W, Yan N, Massagué J, Pavletich NP, Shi Y. Features of a Smad3 MH1-DNA complex. Roles of water and zinc in DNA binding. J Biol Chem 2003;278:20327–20331.
18. Natarajan A, Ghose R, Hill JM. Structure and dynamics of ASC2, a pyrin domain-only protein that regulates inflammatory signaling. J Biol Chem 2006;281:31863–31875.