Published in final edited form as:

Proteins. 2011 January ; 79(1): 335–339. doi:10.1002/prot.22869.

Solution NMR structure of the plasmid-encoded fimbriae regulatory protein PefI from *Salmonella enterica* serovar Typhimurium

James M. Aramini^{1,2,*}, Paolo Rossi^{1,2}, John R. Cort^{2,3}, Li-Chung Ma^{1,2}, Rong Xiao^{1,2}, Thomas B. Acton^{1,2}, and Gaetano T. Montelione^{1,2,4,*}

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

²Northeast Structural Genomics Consortium (NESG), Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854

³Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352

⁴Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey 08854

Keywords

PefI; plasmid-encoded fimbriae regulatory protein; solution NMR structure; winged helix-turn-helix; Northeast Structural Genomics Consortium

Introduction

The surfaces of many bacteria feature pili or fimbriae, proteinaceous filaments that play an integral role in the adhesion of bacteria to host cells and, hence, in their pathogenicity.^{1,2} These extracellular structures are key virulence factors and potential targets for antibacterial drug and vaccine development. In certain *Salmonella* enterobacteria strains, one such class of fimbriae, the virulence plasmid-encoded fimbriae encoded by the *pef* operon,^{3,4} has been shown to be important for adhesion to murine small intestine and fluid accumulation,⁵ as well as biofilm formation on eukaryotic cell surfaces.⁶ Transcription of the *pef* operon, induced under acidic conditions, is modulated by the PefI regulatory protein.⁷ PefI represses plasmid-encoded fimbrial protein production by activating the leucine-responsive regulatory protein (Lrp) mediated inhibition of DNA methylation within the *pef* promoter region.⁷ PefI expression also inhibits transcription of flagellar proteins and consequently cell motility.⁸

Here we present the solution NMR structure of the 70-residue PefI transcription regulator from *Salmonella enterica* serovar Typhimurium LT2 [UniProtKB/TrEMBL ID, Q04822_SALTY; NESG ID, StR82; hereafter referred to as stPefI], a member of the FaeA-like protein domain family (Pfam identifier, PF04703). The sequence alignment of stPefI with its functional homolog in *Escherichia coli*, PapI, is shown in Fig. 1(A). We demonstrate that the structure of stPefI adopts a winged helix-turn-helix motif,^{9,10} consistent with its role as a DNA-binding transcriptional regulator. Moreover, in spite of

*To whom correspondence should be addressed: CABM, Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854, jma@cabm.rutgers.edu; guy@cabm.rutgers.edu.

PDB ID: 2JT1; BMRB ID: 15386

their relatively low sequence identity (29%), the structure of stPefI is highly similar to that of *E. coli* PapI,¹¹ which activates the expression of pyelonephritis-associated pili.^{12,13}

Materials and Methods

Isotopically enriched samples of stPefI were cloned, expressed, and purified, and the sample buffer was optimized for NMR studies following standard protocols of the NESG consortium;^{14,15} see Supplementary Material for a complete description of the methods used in this work. Briefly, samples of [*U*-¹³C, ¹⁵N]- and [*U*-5%-¹³C, 100%-¹⁵N]-stPefI for NMR spectroscopy were concentrated by ultracentrifugation to 0.5 to 0.7 mM in 95% H₂O / 5% ²H₂O solution containing 20 mM ammonium acetate, 450 mM NaCl, 10 mM DTT, 5 mM CaCl₂ at pH 5.5. Analytical gel filtration chromatography, static light scattering (Supplementary Fig. S1) and one-dimensional ¹⁵N *T*₁ and *T*₂ relaxation data (Supplementary Fig. S2) demonstrate that the protein is monomeric under the conditions used in the NMR studies. All NMR data for resonance assignment and structure determination were collected at 20 °C on Bruker AVANCE 600 and 800 MHz spectrometers equipped with conventional 5-mm TXI probes, and a Varian INOVA 600 MHz instrument with a 5-mm HCN cold probe, and referenced to internal DSS (2,2-dimethyl-2-silapentane-5-sulfonic acid). Complete ¹H, ¹³C, and ¹⁵N resonance assignments for stPefI were determined using conventional triple resonance NMR methods, with automated backbone assignment made by AutoAssign 2.4.0,¹⁶ followed by manual side chain assignment. Resonance assignments were validated using the Assignment Validation Suite (AVS) software package¹⁷ and deposited in the BioMagResDB (BMRB accession number, 15386). The solution NMR structure of stPefI was calculated using CYANA 2.1,^{18,19} and the 20 structures with lowest target function out of 100 in the final cycle were further refined by restrained molecular dynamics in explicit water using CNS 1.2.20.²¹ Structural statistics and global structure quality factors were computed using the PSVS 1.4 suite of structure quality assessment programs.²² The global goodness-of-fit of the final structure ensemble with the NOESY peak list data was determined using the RPF analysis program.²³ The final ensemble of 20 models (excluding the C-terminal His₆) were deposited into the Protein Data Bank (PDB ID, 2JT1). All structure figures were rendered using PyMOL (www.pymol.org).

Results and Discussion

The solution NMR structure of stPefI consists of a 3-stranded winged helix-turn-helix (wHTH) motif, comprised of three α -helices (α 1, S4-D21; α 2, T27-G35; α 3, S37-V51) and three short antiparallel β -strands (β 1, V25-K26; β 2, L54-V57; β 3, L66-L69) arranged in an $\alpha\beta\alpha\beta\beta$ topology [Fig. 1(B,C)]. Structural statistics and a summary of the NMR data for stPefI are presented in Table 1 and Supplementary Fig. S3, respectively. The structure of stPefI exhibits a number of signatures of the classic wHTH motif including, i) a tri-helical core comprising the helix-turn-helix (α 2-turn- α 3) followed by a C-terminal beta hairpin wing (β 2-loop- β 3), ii) a 'shs' sequence pattern (where 's' and 'h' denote small and hydrophobic residues, respectively) in the tight turn between α 2 and α 3 (G35-L36-S37), and (iii) a conserved 'phs' sequence motif (where 'p' is a charged residue) within α 2 (D29-I30-A31).^{9,10} Electrostatic surface potential²⁴ [Fig. 1(D)] and ConSurf²⁵ [Fig. 1(E)] analyses of the stPefI structure reveal a positively charged patch and several conserved residues on the face of the protein comprising the HTH and the wing, consistent with the integral roles of the recognition helix (α 3) and the wing in DNA binding by wHTH proteins.^{9,10} The opposite face of the protein is less well conserved and features a quite negatively charged surface (not shown). In addition, the loop in the wing between strands β 2 and β 3 is not well ordered in the structure [Fig. 1(C); Supplementary Fig. S3].

The solution structure of stPefI is highly similar to the structures of a plethora of DNA binding proteins and transcriptional regulators, including the only other structure to date from the FaeA-like protein domain family, the solution structure of *E. coli* PapI (PDB ID, 2HTJ;11 Dali26 Z-score, 7.1; C α RMSD, 2.2 Å). There are similarities and differences between the structures and biological functions of stPefI and ecPapI. Both proteins adopt the wHTH motif and their structures are highly superimposable [Fig. 1(F)]. Like stPefI, ecPapI features a highly basic wHTH face opposite a highly acidic face (not shown), and a disordered loop within the wing.11 NMR chemical shift perturbation studies on ecPapI have revealed that residues within the wing and conserved residues in the major-groove binding recognition helix (α 3) are sensitive to DNA binding;11 we predict that stPefI behaves in an analogous fashion. However, ecPapI possesses an additional highly dynamic C-terminal tail, whose function is not yet known.11 Moreover, in contrast to stPefI which is a repressor of *pef* expression,7 ecPapI has been shown to be a transcriptional activator.12,13 In *E. coli*, the expression of Pap pili is regulated by DNA methylation at distal (GATC I) and proximal (GATC II) *pap* promoter sites which is modulated by the interplay between DNA adenine methylase (Dam), Lrp, and PapI.27,28 Transcription of *pap* genes is arrested by methylation of GATC I and Lrp binding to GATC II, precluding its methylation by Dam, and is activated by the PapI-mediated binding of Lrp to GATC I and concomitant methylation of the GATC II site. Similarly, in *Salmonella*, methylation of *pef* GATC II triggers Pef fimbrial expression.7 However, PefI acts in tandem with Lrp to block methylation of this site and, hence, represses transcription of *pef* genes.7 Remarkably, expression of PefI activates *pap* transcription in *E. coli*.7 Moreover, the *E. coli* ortholog FaeA represses K88 fimbrial production by an analogous mechanism.29,30 These findings suggest that the different biological roles of these functional homologs in bacterial fimbrial production are dependent upon factors in addition to their intrinsic biochemical and structural properties.7 Deciphering the exact nature of the interaction between FaeA-like proteins and their cognate Lrp and the mechanism of their regulation of fimbrial transcription awaits further biochemical and structural studies, which will be enabled by protein sample production protocols, NMR chemical shift data, and structural information such as that presented here for stPefI.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank G. V. T. Swapna, Burkhard Rost, Huang Wang, Chioma Nwosu, and Kellie Cunningham for valuable scientific discussions and technical support. This work was supported by a grant from the National Institute of General Medical Sciences Protein Structure Initiative (PSI) U54-GM074958. A portion of the NMR data was recorded at the Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory. The pET expression vector for stPefI, (NESG StR82-21.1), has been deposited in the PSI Materials Repository (<http://psimr.asu.edu/>).

References

1. Proft T, Baker EN. Pili in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease. *Cell Mol Life Sci.* 2009; 66:613–635. [PubMed: 18953686]
2. Fronzes R, Remaut H, Waksman G. Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria. *EMBO J.* 2008; 27:2271–2280. [PubMed: 18668121]
3. Friedrich MJ, Kinsey NE, Vila J, Kadner RJ. Nucleotide sequence of a 13.9 kb segment of the 90 kb virulence plasmid of *Salmonella typhimurium*: the presence of fimbrial biosynthetic genes. *Mol Microbiol.* 1993; 8:543–558. [PubMed: 8100983]

4. Rotger R, Casadesús J. The virulence plasmids of *Salmonella*. *Internatl Microbiol.* 1999; 2:177–184.
5. Bäumlér AJ, Tsolis RM, Bowe FA, Kusters JG, Hoffmann S, Heffron F. The *pef* fimbrial operon of *Salmonella typhimurium* mediates adhesion to murine small intestine and is necessary for fluid accumulation in the infant mouse. *Infect Immun.* 1996; 64:61–68. [PubMed: 8557375]
6. Ledeböer NA, Frye JG, McClelland M, Jones BD. *Salmonella enterica* serovar Typhimurium requires the Lpf, Pef, and Tafi fimbriae for biofilm formation on HEp-2 tissue culture cells and chicken intestinal epithelium. *Infect Immun.* 2006; 74:3156–3169. [PubMed: 16714543]
7. Nicholson B, Low D. DNA methylation-dependent regulation of Pef expression in *Salmonella typhimurium*. *Mol Microbiol.* 2000; 35:728–742. [PubMed: 10692151]
8. Wozniak CE, Lee C, Hughes KT. T-POP array identifies EcnR and PefI-SrgD as novel regulators of flagellar gene expression. *J Bacteriol.* 2009; 191:1498–1508. [PubMed: 19114490]
9. Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM. The many faces of the helix-turn-helix domain: Transcription regulation and beyond. *FEMS Microbiol Rev.* 2005; 29:231–262. [PubMed: 15808743]
10. Gajiwala KS, Burley SK. Winged helix proteins. *Curr Opin Struct Biol.* 2000; 10:110–116. [PubMed: 10679470]
11. Kawamura T, Le LUK, Zhou H, Dahlquist FW. Solution structure of *Escherichia coli* PapI, a key regulator of the Pap pili phase variation. *J Mol Biol.* 2007; 365:1130–1142. [PubMed: 17109885]
12. Nou X, Braaten B, Kaltenbach L, Low DA. Differential binding of Lrp to two sets of *pap* DNA binding sites mediated by PapI regulates Pap phase variation in *Escherichia coli*. *EMBO J.* 1995; 14:5785–5797. [PubMed: 8846772]
13. Kaltenbach LS, Braaten BA, Low DA. Specific binding of PapI to Lrp-*pap* DNA complexes. *J Bacteriol.* 1995; 177:6449–6455. [PubMed: 7592419]
14. Acton TB, Gunsalus KC, Xiao R, Ma L-C, Aramini J, Baran MC, Chiang Y-W, Climent T, Cooper B, Denissova NG, Douglas SM, Everett JK, Ho CK, Macapagal D, Rajan PK, Shastry R, Shih L-Y, Swapna GVT, Wilson M, Wu M, Gerstein M, Inouye M, Hunt JF, Montelione GT. Robotic cloning and protein production platform of the Northeast Structural Genomics Consortium. *Methods Enzymol.* 2005; 394:210–243. [PubMed: 15808222]
15. Rossi P, Swapna GVT, Huang YJ, Aramini JM, Anklin C, Conover K, Hamilton K, Xiao R, Acton TB, Ertekin A, Everett JK, Montelione GT. A microscale protein NMR sample screening pipeline. *J Biomol NMR.* 2010; 46:11–22. [PubMed: 19915800]
16. 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. [PubMed: 11462827]
17. 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. [PubMed: 14872126]
18. 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. [PubMed: 9367762]
19. 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. [PubMed: 12051947]
20. 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.* 1998; D54:905–921.
21. Linge JP, Williams MA, Spronk CAEM, Bonvin AMJJ, Nilges M. Refinement of protein structures in explicit solvent. *Proteins.* 2003; 50:496–506. [PubMed: 12557191]
22. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. *Proteins.* 2007; 66:778–795. [PubMed: 17186527]
23. 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. [PubMed: 15701001]

24. 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. [PubMed: 11517324]
25. 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. [PubMed: 12499312]
26. Holm L, Kääriäinen S, Rosenström P, Schenkel A. Searching protein structure databases with DaliLite v.3. *Bioinformatics*. 2008; 24:2780–2781. [PubMed: 18818215]
27. Hernday AD, Braaten BA, Low DA. The mechanism by which DNA adenine methylase and PapI activate the Pap epigenetic switch. *Mol Cell*. 2003; 12:947–957. [PubMed: 14580345]
28. Hernday A, Braaten B, Low D. The intricate workings of a bacterial epigenetic switch. *Adv Exp Med Biol*. 2004; 547:83–89. [PubMed: 15230094]
29. Huisman TT, Bakker D, Klaasen P, de Graaf FK. Leucine-responsive regulatory protein, IS1 insertions, and the negative regulator FaeA control the expression of the fae (K88) operon in *Escherichia coli*. *Mol Microbiol*. 1994; 11:525–536. [PubMed: 8152376]
30. Huisman TT, de Graaf FK. Negative control of fae (K88) expression by the ‘global’ regulator Lrp is modulated by the ‘local’ regulator FaeA and affected by DNA methylation. *Mol Microbiol*. 1995; 16:943–953. [PubMed: 7476191]

Fig. 1.

(A) Structure-based sequence alignment of stPefI and *E. coli* PapI (ecPapI). The sequence numbering for stPefI and the secondary structural elements found in its solution NMR structure (PDB ID, 2JT1) are shown above the alignment; t, turn and l, loop. Identical residues are shown in red. (B) Stereoview of the wHTH face in the lowest energy (CNS) conformer of the final solution NMR ensemble of stPefI. Secondary structural elements are labeled; α-helices and β-strands are shown in cyan and magenta, respectively, and loops are colored grey. (C) Superposition of the final ensemble of 20 conformers from the solution NMR structure of stPefI. Same color scheme as in (B). Note the disorder in the loop between strands β2 and β3, constituting the tip of the wing. (D) APBS24 electrostatic surface potential of stPefI showing negative (red), neutral (white), and positive (blue) charges. Selected basic residues are indicated. (E) ConSurf25 image showing the conserved residues in the wHTH motif of stPefI. Residue coloring, reflecting the degree of residue conservation over the entire FaeA-like protein domain family (PF04703, Pfam 24.0), ranges from magenta (highly conserved) to cyan (variable). (F) Overlay of the solution NMR structures of stPefI (red) and ecPapI (PDB ID, 2HTJ;11 green).

Table 1Summary of NMR and structural statistics for stPefI^a

Completeness of resonance assignments ^b	
Backbone (%)	97.8
Side chain (%)	95.8
Aromatic (%)	100
Stereospecific methyl (%)	100
Conformationally-restricting constraints ^c	
Distance constraints	
Total	1155
intra-residue ($i = j$)	333
sequential ($ i - j = 1$)	302
medium range ($1 < i - j < 5$)	228
long range ($ i - j \geq 5$)	292
distance constraints per residue	17.0
Dihedral angle constraints	92
Hydrogen bond constraints	
total	52
long range ($ i - j \geq 5$)	12
Number of constraints per residue	19.1
Number of long range constraints per residue	4.5
Residual constraint violations ^c	
Average number of distance violations per structure	
0.1 – 0.2 Å	0.35
0.2 – 0.5 Å	0.05
> 0.5 Å	0
average RMS distance violation / constraint (Å)	0.01
maximum distance violation (Å)	0.21
Average number of dihedral angle violations per structure	
1 – 10°	0.5
> 10°	0
average RMS dihedral angle violation / constraint (degree)	0.12
maximum dihedral angle violation (degree)	2.40
RMSD from average coordinates (Å) ^{c,d}	
backbone atoms	0.5
heavy atoms	1.0
Procheck Ramachandran statistics ^{c,d}	
most favored regions (%)	95.9
additional allowed regions (%)	4.1

generously allowed (%)	0.0
disallowed regions (%)	0.0
<hr/>	
Global quality scores ^c	
	Raw Z-score
Verify3D	0.35 -1.77
ProsaII	0.62 -0.12
Procheck(phi-psi) ^d	0.16 0.94
Procheck(all) ^d	0.11 0.65
Molprobit clash	21.27 -2.12
<hr/>	
RPF Scores ^e	
Recall	0.981
Precision	0.909
F-measure	0.944
DP-score	0.782
<hr/>	

^a Structural statistics were computed for the ensemble of 20 deposited structures (PDB ID, 2JT1).

^b Computed using AVS software¹⁷ from the expected number of peaks, excluding: highly exchangeable protons (N-terminal, Lys, and Arg amino groups, hydroxyls of Ser, Thr, Tyr), carboxyls of Asp and Glu, non-protonated aromatic carbons, and the C-terminal His₆ tag.

^c Calculated using PSVS 1.4 program.²² Average distance violations were calculated using the sum over r^{-6} .

^d Ordered residue ranges [$S(\phi) + S(\psi) > 1.8$] : 5–19, 23–57, 66–69.

^e RPF scores²³ reflecting the goodness-of-fit of the final ensemble of structures (including disordered residues) to the NOESY data and resonance assignments.