PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Solution NMR structure of BT_0084, a conjugative transposon lipoprotein from *Bacteroides thetaiotamicron*

Theresa A. Ramelot,[1]* Yunhuang Yang,[1] Rong Xiao,[2] Thomas B. Acton,[2] John K. Everett,[2] Gaetano T. Montelione,[2,3] and Michael A. Kennedy[1]*

[1] Department of Chemistry and Biochemistry, Miami University, Oxford, Ohio 45056 and the Northeast Structural Genomics Consortium

[2] Center for Advanced Biotechnology and Medicine, Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854 and the Northeast Structural Genomics Consortium

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

Key words: TraQ; DUF3872; PF12988; structural genomics; *tra* operon; conjugation; pili; CTnDOT; CTnERL.

INTRODUCTION

Here we report the solution NMR structure of a 121-residue fragment (residues 27–147) of BT_0084, a lipoprotein from *Bacteroides thetaiotamicron* (GenBank, NP_808997; UniProtKB ID, Q8ABM6; KEGG ID, bth:BT_0084; NESG ID, BtR376) that belongs to a family of conserved bacterial proteins found in conjugative transposons (Pfam, PF12988; DUF3872). Conjugative transposons (CTns) are mobile DNA elements that can be transferred to recipient bacteria, by conjugation or horizontal gene transfer, where they become integrated into the recipient genome. CTns have been found in gram-positive and gram-negative bacteria, but have been most studied in the *Bacteroides* group of gram-negative bacteria because of their contribution to the spread of antibiotic resistance in this clinically important pathogen. *Bacteroides* CTns typically range in size from 70 to 80 kbp and many types have been characterized.[1] BT_0084 is located in a chromosomally integrated CTn in the transfer (*tra*) operon, which is an ~20-kbp region containing genes essential for CTn transfer after it undergoes excision and formation of a circular intermediate. The genes from *traA* to *traQ* were identified in two very similar 65 and 50 kpb *Bacteroides* CTns called CTnDOT and CTnERL, and encode proteins that regulate CTn transfer as well as structural proteins needed to form the transfer apparatus.[2,3] BT_0084 is 50% identical to TraQ in CTnERL (UniProtKB ID, Q650G2; KEGG ID, bfr:BF0113), which acts as a repressor of conjugative transfer.[2] Salyers *et al.* discovered that disruption of *traQ* resulted in a 100-fold increase in CTnERL transfer as well as increased production of TraG and TraN, suggesting that TraQ is a transcriptional repressor of these two proteins.[2]

BT_0084 was selected for experimental structure determination by the Protein Structure Initiative as a conserved protein found in many bacteria. Full length BT_0084 is a 147-residue prolipoprotein with a predicted 25-residue lipoprotein signal peptide at the N-terminus (by LipoP v1.0a[4]) that targets the protein for membrane translocation. After processing, the mature protein has the signal peptide removed and lipids covalently attached to the N-terminal Cys (C26 before processing) to anchor the protein to the periplasmic surface of the inner or

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-GM074958.

*Correspondence to: Theresa A. Ramelot, Department of Chemistry and Biochemistry, Miami University, Oxford, OH 45056. or Michael A. Kennedy, Department of Chemistry and Biochemistry, Miami University, Oxford, OH 45056. E-mail: ramelota@muohio.edu

Received 23 July 2011; Revised 14 September 2011; Accepted 20 September 2011
Published online 29 October 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.23235

Table I
Summary of NMR and Structural Statistics for BT_0084[a]

Completeness of resonance assignments[b,c]		
Backbone (%)		96.3
Side chain (%)		89.9
Aromatic (%)		67.1
Stereospecific methyl (%)		100
Conformationally-restricting constraints[c]		
Distance constraints		
Total		1354
Intraresidue ($i = j$)		325
Sequential ($\|i - j\| = 1$)		375
Medium range ($1 < \|i - j\| < 5$)		104
Long range ($\|i - j\| \geq 5$)		550
Dihedral angle constraints		154
Hydrogen bond constraints		78
Number of constraints per residue		14.4
Number of long range constraints per residue		5.6
Residual constraint violations[c]		
Average number of distance violations per structure		
0.1–0.2 Å		2.8
0.2–0.5 Å		0.4
>0.5 Å		0
Average number of dihedral angle violations per structure		
1–10°		6.3
>10°		0.1
RMSD from average coordinates (Å)[c,d]		
Backbone atoms		0.7
Heavy atoms		1.2
MolProbity Ramachandran statistics[c,d]		
Most favored regions (%)		99.3
Allowed regions (%)		0.7
Disallowed regions (%)		0.0
Global quality scores (Raw/Z-score)[c]		
Verify3D	0.31	−2.4
ProsaII	0.40	−1.0
Procheck(phi-psi)[d]	−0.51	−1.7
Procheck(all)[d]	−0.32	−1.9
Molprobity clash	18.0	−1.6
RPF scores[c,e]		
Recall/Precision	0.95	0.91
F-measure/DP-score	0.93	0.74

[a]Structural statistics were computed for the ensemble of 20 deposited structures (PDB ID, 2L3B).
[b,c]Calculated for residues 2–122.
[c,]Calculated using PSVS 1.4 program.[6] Average distance violations were calculated using the sum over r^{-6}.
[d]Ordered residue ranges: 13–38, 45–54, 59–69, 72–90, 92–113, with the sum of ϕ and ψ order parameters >1.8.
[e]RPF scores (including disordered residues) to the NOESY data and resonance assignments.

outer membrane. For this reason, only the structure of residues 27–147 is reported here.

MATERIALS AND METHODS

The truncated *bt_0084* gene was cloned into a pET expression vector (NESG BtR376-27-147-21.4), which has been deposited in the PSI Materials Repository (http://psimr.asu.edu/). The truncated protein, hereafter called BT_0084, included an additional N-terminal Met and C-terminal His_6 tag (LEHHHHHH). Cloning, expression, and purification were conducted following standard protocols of the Northeast Structural Genomics Consortium (NESG) to prepare [*U*-^{13}C, ^{15}N]- and *U*-^{15}N, 5% biosynthetically directed ^{13}C (NC5) samples[5]; see Supporting Information for a detailed description of the sample preparation, NMR data acquisition, and structure determination methods. BT_0084 is a monomer under the conditions used in the NMR experiments (1.0 m*M* protein, 10% v/v D_2O, 20 m*M* MES, 100 m*M* NaCl, 10 m*M* DTT, 5 m*M* $CaCl_2$, pH 6.5, 25°C) based on analytical static light scattering in-line with gel filtration chromatography and correlation time estimates based on one-dimensional ^{15}N T_1 and T_2 relaxation data (estimated τ_c 10.2 ns, Supporting Information Fig. S1). NMR structure quality analysis was performed using the PSVS[6] software package. Structural statistics for BT_0084 are presented in Table I and a labeled ^{1}H-^{15}N HSQC spectrum is shown in Supporting Information Figure S2. The final ensemble of 20 models and NMR resonance assignments for BT_0084 were deposited to the Protein Data Bank (PDB ID, 2L3B) and BioMagResDB (BMRB accession number, 17176), respectively.

RESULTS AND DISCUSSION

The solution NMR structure of BT_0084 is a pair of β sheets with antiparallel β-strands that form a β sandwich [Fig. 1(A)]. The first sheet is composed of five β-strands (β1, S14-T17; β2, T29-H37; β7, T80-S87; β4, T59-M62; β5, V67-L68) and the second sheet of four β-strands (β6, Y74-P75; β3, F48-F52; β8, Q93-Q100; β9, L106-F113). The N-terminal tail (M1-P12), C-terminal tail (N114-E122 + L123-H130 His_6 tag), and part of the loop between β2 and β3 (D39-T45) are not well defined [Fig. 1(D)]. These disordered regions correlate well with low values in the ^{1}H-^{15}N heteronuclear NOE data [Supporting Information Fig. S1(C)]. ConSurf[7] [Fig. 1(B)] analysis revealed two patches of conserved surface residues located on each β sheet on opposite sides of the protein and electrostatic surface potential analysis[9] indicated that residues in this region are made up of both positive and negatively charged amino acids [Fig. 1(C)]. A stretch of conserved residues (V5-Q8) is also present in the disordered N-terminal tail of the protein.

The solution NMR structure of BT_0084 reported here is the first structural representative from protein domain family PF12988 (DUF3872). The structure of the 55% identical protein BVU_1572 (fragment 27-141) from *Bacteroides vulgatus* (UniProtKB ID, A6L0P5; KEGG ID, bvu:BVU_1572; NESG ID, BvR155; PDB ID 2L7Q, Supporting Information Table SI) was subsequently solved by the NESG as well. These two structures are similar (1.6 Å Cα RMSD) and both exhibit low sequence similarity (<15%) with other structurally related entries in the PDB identified by a Dali (v.3)[12] search. The structures of BT_0084 and BVU_1572 belong to the immunoglobulin

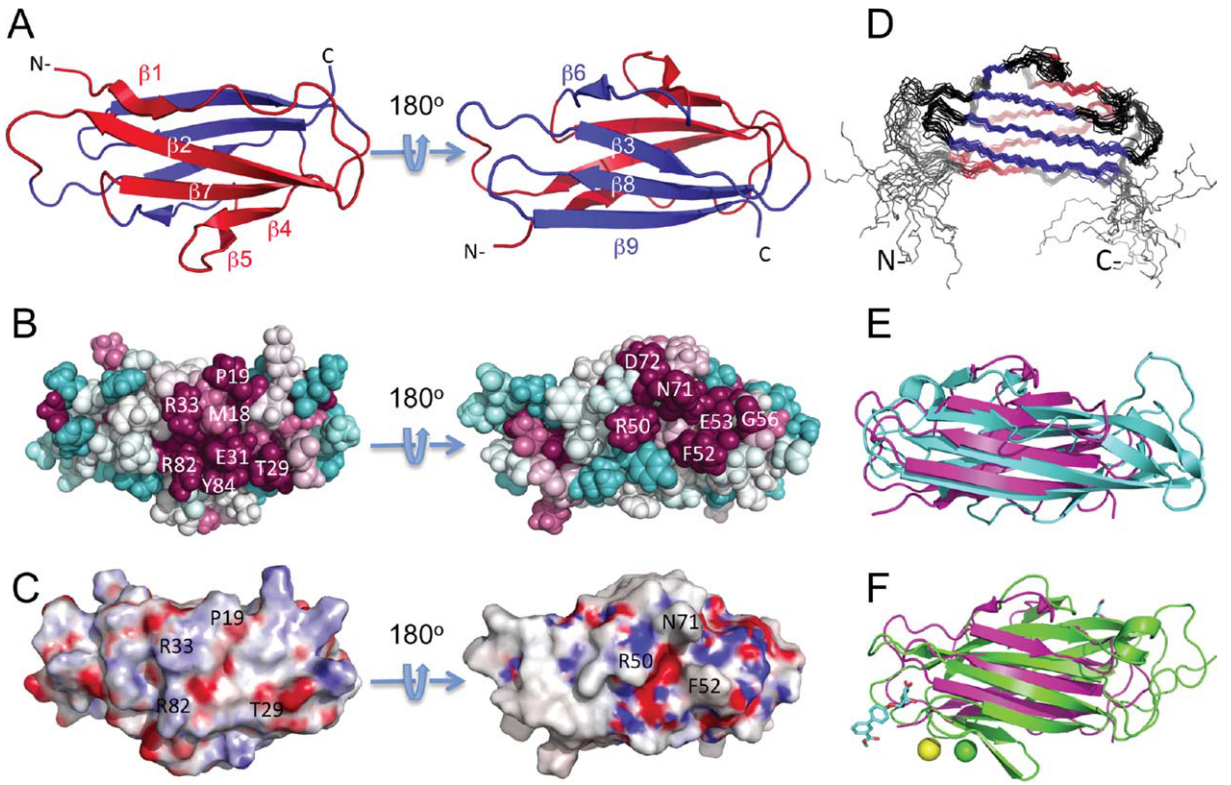

Figure 1

(**A**) Cartoon view of the lowest energy conformer from the solution NMR structure of BT_0084 (PDB ID, 2L3B, residues 11–115). β-strands are colored by sheet and labeled. Figures on right show the back view after a 180° *x*-axis rotation. (**B**) ConSurf[7] image showing the conserved surface residues (residues 12–114). Residue coloring, reflecting the degree of residue conservation over the 13 homologs found by blasting the KEGG database,[8] ranges from magenta (highly conserved) to cyan (variable). Selected highly conserved residues are labeled. (**C**) APBS[9] electrostatic surface potential diagrams (residues 12–114) colored by electrostatic potential (−20 kT/e, red to +20 kT/e, blue). Selected residues are indicated. (**D**) Superposition of the final ensemble of 20 conformers. (**E**) Overlay of the residues (11–115) from the solution NMR structure of BT_0084 (magenta) and (**F**) CfaB from *E. coli* (PDB ID, 3F84-A[10]; cyan) from the X-ray crystal structure of FimH from *E. coli* (PDB ID, 3MCY-A;[11] green). The mannose derivative, glycerol, calcium ion (yellow), and chloride (green) from 3MCY are shown. All structure figures were created with PyMOL (www.pymol.org).

superfamily and are similar to several structures of proteins involved in assembly of short attachment pili involved in bacterial cell adhesion (from 2.5 to 3.0 Å Cα RMSD). CfaB (PDB ID, 3F84-A[10]; Dali Z-score, 7.3; Cα RMSD, 2.6 Å for 96 residues) and CfaE (PDB ID, 3F83-A[10]; Dali *Z*-score, 7.3; Cα RMSD, 2.5 Å for 97 residues in the C-terminal pilin domain) are pilin subunits of colonization factor antigen I pili from enterotoxigenic *E. coli* that belong to class 5 (attach to host intestinal epithelia). CfaE is the minor adhesive subunit at the tip of the mature pilus and CfaB is the major subunit that forms a multimeric ordered structure.[10] FimC (PDB ID, 1KIU-E[13]; Dali Z-score, 6.3; Cα RMSD, 2.8 Å for 85 residues) and FimH (PDB ID, 3MCY-A[11]; Dali *Z*-score, 7.3; Cα RMSD, 2.7 Å for 95 residues in the N-terminal adhesion domain) are the periplasmic chaperone and minor adhesive subunit for type 1 pili in the well-studied *E. coli* cell surface adhesion system. BT_0084 does not have corresponding conserved surface residues with the mannose- or calcium-binding regions of FimH. Many proteins involved in different types of pili assembly have a immunoglobulin-like fold, even though sequence similarities are low.[14] The role and types of pili involved in *Bacteroides* CTn conjugation have not yet been characterized.

Genes in the *tra* operon encode structural proteins that compose the mating bridge as well as regulatory proteins.[2] Regulation of the *tra* genes, as studied in *Bacteroides* CTnDOT, occurs at the transcriptional level and involves a complex regulatory system that may be tightly regulated to ensure stable maintenance of the CTn.[15] Because TraQ is a lipoprotein with a role in transcriptional regulation of *traG* and *traN*,[2] one possible mode of function is a role in synthesis of a diffusible autorepressor that can diffuse from the periplasm to the cytoplasm to inhibit transcriptional regulation. This role may involve the other proteins implicated in repression, TraO and TraP.[2]

Interestingly, the CTn that encodes BT_0084 does not encode any antibiotic resistance genes, but instead

contains arsenic resistance genes (Supporting Information Fig. S3). Although this CTn has not been previously characterized, arsenic is a toxic environmental pollutant for which the resistance genes are often found on bacterial plasmids.[16] Additionally, heavy metal cations have antimicrobial properties and are often coselected with genes for antibiotic resistance.[17] It is well-known that *Bacteroides* CTns can mobilize DNA elements other than just themselves, such as mobilizable transposons and coresident plasmids.[18] Therefore, this can allow for the horizontal transfer of genetic material that isn't directly included in the CTn itself.

In summary, the rapid spread of antibiotic resistance, among colonic *Bacteroides* species as well as to distantly related bacteria such *E. coli*, is directly related to the spread of *Bacteroides* CTns.[19] Knowledge of the structure of the TraQ protein, with its role in transcriptional repression of CTn transfer genes, represents another step toward understanding this system and is the first structure in the *Bacteroides* TraQ family.

ACKNOWLEDGMENTS

The authors thank Shuisong Ni for valuable scientific discussion and Alex Eletski at SUNY Buffalo for providing the intav and tc macros for the Varian vnmr software. They thank D. Wang, C. Ciccosanti, H. Janjua at the Rutger's protein production facility for technical support. The majority of the data collection was conducted at the Ohio Biomedicine Center of Excellence in Structural Biology and Metabonomics at Miami University.

REFERENCES

1. Bedzyk LA, Shoemaker NB, Young KE, Salyers AA. Insertion and excision of *Bacteroides* conjugative chromosomal elements. J Bacteriol 1992;174:166–172.
2. Bonheyo GT, Hund BD, Shoemaker NB, Salyers AA. Transfer region of a *Bacteroides* conjugative transposon contains regulatory as well as structural genes. Plasmid 2001;46:202–209.
3. Bonheyo G, Graham D, Shoemaker NB, Salyers AA. Transfer region of a *Bacteroides* conjugative transposon, CTnDOT. Plasmid 2001;4:41–51.
4. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Prot Sci 2003;12:1652–1662.
5. 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.
6. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. Proteins 2007;66:778–795.
7. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 2010;38:W529–W533.
8. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res 2010;38:D355–D360.
9. 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.
10. Li YF, Poole S, Nishio K, Jang K, Rasulova F, McVeigh A, Savarino SJ, Xia D, Bullitt E. Structure of CFA/I fimbriae from enterotoxigenic *Escherichia coli*. Proc Natl Acad Sci USA 2009;106:10793–10798.
11. Han ZF, Pinkner JS, Ford B, Obermann R, Nolan W, Wildman SA, Hobbs D, Ellenberger T, Cusumano CK, Hultgren SJ, Janetka JW. Structure-based drug design and optimization of mannoside bacterial FimH antagonists. J Med Chem 2010;53:4779–4792.
12. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucleic Acids Res 2010;38:W545–W549.
13. Hung CS, Bouckaert J, Hung D, Pinkner J, Widberg C, DeFusco A, Auguste CG, Strouse R, Langermann S, Waksman G, Hultgren SJ. Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. Mol Microbiol 2002;44:903–915.
14. Kline KA, Dodson KW, Caparon MG, Hultgren SJ. A tale of two pili: assembly and function of pili in bacteria. Trends Microbiol 2010;18:224–232.
15. Jeters RT, Wang GR, Moon K, Shoemaker NB, Salyers AA. Tetracycline-associated transcriptional regulation of transfer genes of the *Bacteroides* conjugative transposon CTnDOT. J Bacteriol 2009;191:6374–6382.
16. Cervantes C, Ji GY, Ramirez JL, Silver S. Resistance to arsenic compounds in microorganisms. FEMS Microbiol Rev 1994;15:355–367.
17. Baker-Austin C, Wright MS, Stepanauskas R, McArthur JV. Co-selection of antibiotic and metal resistance. Trends Microbiol 2006;14:176–182.
18. Whittle G, Shoemaker NB, Salyers AA. Characterization of genes involved in modulation of conjugal transfer of the *Bacteroides* conjugative transposon CTnDOT. J Bacteriol 2002;184:3839–3847.
19. Shoemaker NB, Vlamakis H, Hayes K, Salyers AA. Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. App Environ Microbiol 2001;67:561–568.