Published in final edited form as:

Proteins. 2012 February ; 80(2): 667–670. doi:10.1002/prot.23235.

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

Theresa A. Ramelot^{1,*}, Yunhuang Yang¹, Rong Xiao², Thomas B. Acton², John K. Everett², Gaetano T. Montelione^{2,3}, and Michael A. Kennedy^{1,*}

¹ Department of Chemistry and Biochemistry, Miami University, Oxford, Ohio 45056 and 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 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

Abstract

Here we describe the solution NMR structure of the 120 amino acid fragment of BT_0084, without the N-terminal lipoprotein targeting sequence, encoded in a conjugative transposon (CTn) in the genome of *Bacteroides thetaiotamicron*. BT_0084 belongs to a conserved family of TraQ lipoproteins that are encoded at the end of the *tra* operon, which contains genes essential for transfer of CTns. The structure belongs to the immunoglobulin superfamily and shares structural similarity, albeit low sequence identity (< 15%), to other proteins involved in pili production for bacterial cell attachment. Although its role in repression of CTn transfer remains to be determined, the structure of BT_0084 reported here represents the first from the *Bacteroides* TraQ family and should facilitate further understanding of the *tra* operon-regulated transfer of CTns.

Keywords

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

*Correspondence to: Theresa A. Ramelot, Department of Chemistry and Biochemistry, Miami University, Oxford, OH 45056. ramelota@muohio.edu; Michael A. Kennedy, Department of Chemistry and Biochemistry, Miami University, Oxford, OH 45056. kennedm4@muohio.edu. Phone: 1-513-593-2402 (Theresa A. Ramelot) Phone: 1-513-529-8267 (Michael A. Kennedy) Fax: (513) 529-5715. Yunhuang Yang yangy9@muohio.edu Rong Xiao rxiao@cabm.rutgers.edu Thomas B. Acton acton@cabm.rutgers.edu John K. Everett everett.jk@gmail.com Gaetano T. Montelione guy@cabm.rutgers.edu.

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 a ~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 kbp *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 v1.0a⁴) 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 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 His₆ tag (LEHHHHHH). Cloning, expression, and purification were conducted following standard protocols of the Northeast Structural Genomics Consortium (NESG) in order to prepare [*L*-¹³C, ¹⁵N]- and *L*-¹⁵N, 5% biosynthetically-directed ¹³C (NC5) samples;⁵ see Supplementary Material 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 mM protein, 10% v/v D₂O, 20 mM MES, 100 mM NaCl, 10 mM DTT, 5 mM CaCl₂, 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 ¹⁵N *T*₁ and *T*₂ relaxation data (estimated τ_c 10.2 ns, Supplementary Fig. S1). NMR structure quality analysis was performed using the PSVS⁶ software package. Structural statistics for BT_0084 are presented in Table 1 and a labeled ¹H-¹⁵N HSQC spectrum is shown in Supplementary Fig. 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 anti-parallel β -strands that form a β sandwich (Fig. 1A). The first sheet is composed of five β -strands (β 1, S14-T17; β 2, T29-H37; β 7, T80-S87; β 4, T59-M62; β 5, V67-L68) and the second sheet of four β -strands (β 6, Y74-P75; β 3, F48-F52; β 8, Q93-Q100; β 9, L106-F113). The N-terminal tail (M1-P12), C-terminal tail (N114-E122 + L123-H130 His₆ tag), and part of the loop between β 2 and β 3 (D39-T45) are not well defined (Fig. 1D). These disordered regions correlate well with low values in the ¹H-¹⁵N heteronuclear NOE data (Supplementary Fig. S1C). ConSurf⁷ (Fig. 1B) analysis revealed two patches of conserved surface residues located on each β sheet on opposite sides of the protein and electrostatic surface potential analysis⁸ indicated that residues in this region are made up of both positive and negatively charged amino acids (Fig. 1C). 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 (DUF 3872). 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, Supplementary Table S1) was subsequently solved by the NESG as well. These two structures are similar (1.6 Å C α RMSD) and both exhibit low sequence similarity (< 15%) with other structurally related entries in the PDB identified by a Dali (v.3)⁹ search. The structures of BT_0084 and BVU_1572 belong to the immunoglobulin 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 C α RMSD). CfaB (PDB ID, 3F84-A;¹⁰ Dali Z-score, 7.3; C α RMSD, 2.6 Å for 96 residues) and CfaE (PDB ID, 3F83-A¹⁰; Dali Z-score, 7.3; C α 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-E¹¹; Dali Z-score, 6.3; C α RMSD, 2.8 Å for 85 residues) and FimH (PDB ID, 3MCY-A¹²; Dali Z-score, 7.3; C α 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 an immunoglobulin-like fold, even though sequence similarities are low.¹³ 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 in order to ensure stable maintenance of the CTn.¹⁴ Since 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 in order 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 (Supplementary material 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.¹⁵ Additionally, heavy metal cations have antimicrobial properties and are often co-selected with genes for antibiotic resistance.¹⁶ It is well-known that *Bacteroides* CTns can mobilize DNA elements other than just themselves, such as mobilizable transposons and co-resident plasmids.¹⁷ 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.¹⁸ Knowledge of the structure of the TraQ protein, with its role in transcriptional repression of CTn transfer genes, represents another step towards understanding this system and is the first structure in the *Bacteroides* TraQ family.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant sponsor: National Institute of General Medical Sciences Protein Structure Initiative (PSI); Grant number: U54-GM074958. We 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. We thank D. Wang, C. Ciccocanti, H. Janjua at the Rutgers 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. [PubMed: 1309516]
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. [PubMed: 11735369]
3. Bonheyo G, Graham D, Shoemaker NB, Salyers AA. Transfer region of a *Bacteroides* conjugative transposon, CTnDOT. *Plasmid.* 2001; 4:41–51. [PubMed: 11319931]
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(8):1652–1662.
5. Acton TB, Xiao R, Anderson S, Aramini J, Buchwald WA, Ciccocanti 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. [PubMed: 21371586]
6. Bhattacharya A, Tejero R, Montelione G. Evaluating protein structures determined by structural genomics consortia. *Proteins.* 2007; 66:778–795. [PubMed: 17186527]

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. [PubMed: 20478830]
8. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A.* 2001; 98:10037–10041. [PubMed: 11517324]
9. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 2010; 38:W545–W549. [PubMed: 20457744]
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 U S A.* 2009; 106:10793–10798. [PubMed: 19515814]
11. Hung CS, Bouckaert J, Hung D, Pinkner J, Widberg C, DeFusco A, Auguste CG, Strouse R, Langemann 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. [PubMed: 12010488]
12. 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. [PubMed: 20507142]
13. Kline KA, Dodson KW, Caparon MG, Hultgren SJ. A tale of two pili: assembly and function of pili in bacteria. *Trends in Microbiol.* 2010; 18:224–232.
14. 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. [PubMed: 19700528]
15. Cervantes C, Ji GY, Ramirez JL, Silver S. Resistance to arsenic compounds in microorganisms. *FEMS Microbiol Rev.* 1994; 15:355–367. [PubMed: 7848659]
16. Baker-Austin C, Wright MS, Stepanauskas R, McArthur JV. Co-selection of antibiotic and metal resistance. *Trends in Microbiol.* 2006; 14:176–182.
17. 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. [PubMed: 12081954]
18. 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.
19. 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. [PubMed: 19880382]

Figure 1.

(A) Cartoon view of the lowest energy conformer from the solution NMR structure of BT_0084 (PDB ID, 2L3B, residues 11 to 115). β -strands are colored by sheet and labeled. Figures on right show the back view after a 180° x-axis rotation. (B) ConSurf⁷ image showing the conserved surface residues (residues 12 to 114). Residue coloring, reflecting the degree of residue conservation over the 13 homologs found by blasting the KEGG database,¹⁹ ranges from magenta (highly conserved) to cyan (variable). Selected highly conserved residues are labeled. (C) APBS⁸ electrostatic surface potential diagrams (residues 12 to 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 to 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;¹² 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).

Table ISummary 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	
intra-residue ($i = j$)	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

^aStructural statistics were computed for the ensemble of 20 deposited structures (PDB ID, 2L3B).

^bCalculated for residues 2-122.

^cCalculated using PSVS 1.4 program.⁶ Average distance violations were calculated using the sum over r^6 .

^dOrdered residue ranges: 13-38, 45-54, 59-69, 72-90, 92-113, with the sum of ϕ and ψ order parameters > 1.8 .

^eRPF scores (including disordered residues) to the NOESY data and resonance assignments.