Published in final edited form as:

Biomol NMR Assign. 2011 April ; 5(1): 67–70. doi:10.1007/s12104-010-9269-y.

Chemical shift assignments for F-plasmid TraI (381–569)

Nathan T. Wright,

Department of Biology, Johns Hopkins University, 3400 N. Charles St, Baltimore, MD 21218, USA

Ananya Majumdar, and

Biomolecular NMR Center, Johns Hopkins University, Baltimore, MD 21218, USA

Joel F. Schildbach

Department of Biology, Johns Hopkins University, 3400 N. Charles St, Baltimore, MD 21218, USA

Joel F. Schildbach: joel@jhu.edu

Abstract

TraI, the F plasmid-encoded nickase, is a 1,756 amino acid protein essential for conjugative transfer of F plasmid DNA from one bacterium to another. While crystal structures of N- and C-terminal domains of F TraI have been determined, central domains of the protein are structurally unexplored. These middle domains (between residues 306 and 1,500) are known to both bind single-stranded DNA (ssDNA) and unwind DNA through a highly processive helicase activity. Of this central region, the more C-terminal portion (~900–1500) appears related to helicase RecD of the *E. coli* RecBCD complex. The more N-terminal portion (306–900), however, shows limited sequence similarity to other proteins. In an attempt to define the structure of well-folded domains of this middle region and discern their function, we have isolated stable regions of TraI following limited proteolysis. One of these regions, TraI (381–569), was identified and a genetic construct encoding it was engineered. The protein was expressed, purified, and the sequence-specific chemical shifts for it were assigned.

Keywords

TraI; Bacterial conjugation; NMR chemical shift assignments; ssDNA binding protein

Biological context

Conjugative plasmids encode proteins called nickases (also called relaxases or mobilization proteins) whose activities are essential for transfer of the plasmids between bacteria (reviewed in (Garcillan-Barcia et al. 2009)). To initiate DNA transfer, these proteins bind one strand of the plasmid in a site-specific manner within the plasmid's origin of transfer (*oriT*) and cleave that DNA strand through a metal-dependant reaction (Traxler and Minkley 1988; Inamoto et al. 1991; Pansegrau et al. 1993; Inamoto et al. 1994; Sherman and Matson 1994; Llosa et al. 1995; Stern and Schildbach 2001; Datta et al. 2003; Larkin et al. 2005; Larkin et al. 2007). This reaction results in a stable, long-lived covalent intermediate that includes a phospho-tyrosyl linkage between the relaxase and the DNA (Pansegrau et al. 1990; Matson et al. 1993; Llosa et al. 1995; Larkin et al. 2005). In addition to the relaxase activity, F TraI and many other relaxases have other important activities: The MobA

relaxase of plasmid R1162 also possesses a primase activity (Henderson and Meyer 1996) and the TraI proteins from plasmids F, R1 and R100, and the TrwC protein of plasmid R388 all have a helicase activity (Traxler and Minkley 1988; Llosa et al. 1995; Llosa et al. 1996). In one model of the F TraI mechanism of action, the TraI helicase activity is required to unwind and separate plasmid strands, converting the transferred strand into a linear ssDNA molecule that can be shuttled from the donor cell through a secretory pore and into the recipient cell. Once this process is completed, the DNA in the recipient can be re-circularized and second strand synthesis can be performed in donor and recipient, resulting in two cells capable of additional rounds of conjugative transfer. Because conjugation can facilitate efficient transfer of considerable amounts of DNA, this mechanism is a major conduit for exchange of genetic information between different bacterial strains or species, including genes encoding antibiotic resistance or virulence factors (Barlow 2009).

F TraI, at 1,756 residues long, is one of the largest proteins expressed in *E. coli*. The crystal structure of the N-terminal 306 amino acids has been solved in both the apo and the ssDNA-bound state (Datta et al. 2003; Larkin et al. 2005). Additionally, the structure of a ~150 residue fragment of the C-terminus has been determined (1,476–1,629) (Guogas et al. 2009). However, there is no structural information about the middle 1,100 residues of this protein. Functionally, it is known that the middle domains of TraI function in nonspecific ssDNA binding and contain helicase activity. We report here that there are at least two proteolytically stable domains between residues 306–1,500. Of these, one (residues 381–569) displays a well-ordered HSQC profile. With the eventual goal of solving the high-resolution structure of all of the well-ordered central domains in TraI, we here present the chemical shift assignments of TraI (381–569).

Methods and experiments

We are using limited proteolysis of the large F TraI protein (192 kDa; 1,756 residues) to identify stable domains amenable to high-resolution structure determination, through either X-ray or solution methods. TraI was purified as described (Dostal and Schildbach 2010) and subjected to chymotrypsin limited proteolysis in a 1:100 chymotrypsin: TraI (w/w) ratio for 1–20 min. The resulting degradation products were run on a 10% SDS Tris-Tricine gel. From this procedure, bands corresponding to products with 20, 38, and 52 kDa masses formed within 1 min and remained stable for at least an additional 10 min in the protease mixture. A gel containing these bands was transferred to a PVDF membrane using standard western blot protocols, and each band was excised and sent for N-terminal amino acid analysis. From this experiment, coupled with mass spectroscopy, we determined that the 20-kDa fragment consisted of residues 381–569, the 52-kDa band was residues 381–858, and the 38-kDa fragment was residues 1,141–1,479. Each of these cleavage sites occurs directly after a large hydrophobic residue, as expected for chymotrypsin digestion. Because of its favorable size, we sought to assign the chemical shifts of the 20-kDa fragment using traditional NMR techniques.

A DNA sequence encoding TraI (381–569) was PCR amplified from the pET24a-TraI vector (Street et al. 2003) encoding the wild-type protein, and the PCR product was subcloned into the pET24a expression vector using *NdeI* and *EcoRI* restriction sites. The sequence of the insert was then verified by DNA sequencing. BL21(DE3) cells containing the plasmid were grown in 1 L of either ^{15}N or ^{15}N , ^{13}C MOPS minimal media (Wright et al. 2005) ($^{15}\text{NH}_4$ and ^{13}C glucose obtained from Cambridge isotopes, Andover, MA) to an OD_{600} of 0.6–0.8 and induced for 3 h with IPTG. The cells were spun down and resuspended in a buffer containing 20 mM Tris pH 7.2, 50 mM NaCl, 1 mM EDTA. After lysis via sonication and clarification by centrifugation, the resulting supernatant was applied over a SP-heparin column (GE Healthcare Life Sciences, Piscataway, NJ) in 50 mM NaCl

and the TraI fragment was eluted in ~200 mM NaCl using a salt gradient produced by an AKTA prime plus (GE Healthcare). The eluted protein, identified by OD₂₈₀ measurements, was then run on an SP-Blue column (GE Healthcare), with the desired protein eluting in a broad peak between 1.5 and 2 M NaCl. The fractions corresponding to this peak were run on an SDS gel and those fractions deemed to be >95% pure were pooled and concentrated in a 10 kDa cutoff Centricon filter to a protein concentration of between 300 and 700 μ M, as quantified by absorbance ($\epsilon_{280} = 8.3 \text{ mM}^{-1} \text{ cm}^{-1}$). The construct was then extensively dialysed into an NMR buffer consisting of 20 mM Tris, 25 mM NaCl, and 1 mM EDTA. The final NMR samples also contained 10% D₂O. Most preps produced between 10 and 20 mg of protein.

All NMR experiments were collected on either a 600 MHz Bruker AVANCE II spectrometer or an 800 MHz Varian INOVA spectrometer, both equipped with a 5 mm triple-resonance cryogenic probe and z-axis pulse field gradient coils. We collected a 2D HSQC and standard triple resonance experiments, including the HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, C(CO)NH, and H(CCO)NH. NMR data were processed with NMRPipe (Delaglio et al. 1995) and analyzed with Sparky (Goddard and Kneller 2008). While the assignments were carried out manually, we used the MARS program for verification (Jung and Zweckstetter 2004). The ¹H chemical shifts were referenced to external DSS, the ¹³C shifts were referenced indirectly to DSS using the frequency ratio ¹³C/¹H = 0.251449527 and ¹⁵N shifts were referenced indirectly to liquid ammonia using ¹⁵N/¹H = 0.101329118.

Extent of assignments and data deposition

TraI (381–569) has 184 non-proline residues, and 166 of these resonances are labeled on the HSQC (Fig. 1). All visible peaks on the HSQC were unambiguously assigned. Analysis of the triple resonance experiments provided over 90% of the ¹H and ¹⁵N assignments, over 93% of the C _{α} and C _{β} resonances (Fig. 2a), and over 90% of the C' resonances of the protein. Importantly, over 93% of side chain hydrogens expected from the H(CCO)NH (non-aromatic α , β , γ , δ , and ϵ hydrogens attached to ¹³C atoms) are assigned, and the shifts are well-dispersed. This is likely to significantly facilitate high-resolution structure determination (currently in progress). For the 5 proline residues, at least the C _{α} , C _{β} , and C _{γ} have been assigned. Of the 18 residues for which there are no ¹H or ¹⁵N data (residues T381, S382, G383, H385, D388, E389, V392, G440, G441, A442, D469, K477, I487, Q539, R540, T541, T543, and G544), 7 of them occur in the first 12 N-terminal amino acids. This observation, combined with the narrow ¹H spectral window of the remaining first ~20 amino acids, suggest that this region of the protein fragment is mostly unstructured. In contrast, the C-terminal end of the protein is fully assigned and appears to be in a more stable structural conformation (Fig. 2b). That the N-terminus is dynamic is perhaps not surprising, since chymotrypsin only cuts at exposed hydrophobic residues, which are relatively uncommon in well-ordered proteins. The HSQC shows all 18 NH₂ sidechains, which have been assigned, as well as the NH of the single tryptophan residue (W560). The sequence-specific chemical shift assignments for TraI (381–569) have been deposited in the BioMagResBank (BMRB) under the accession number 16971.

Acknowledgments

This work was supported by the National Institutes of Health GM61017 (to J.F.S.) and by the Dimitri V. d'Arbeloff fellowship (to N.T.W.). We would also like to thank the JHU Biomolecular NMR Center for providing facilities, resources and consultation.

References

- Barlow M. What antimicrobial resistance has taught us about horizontal gene transfer. *Methods Mol Biol* 2009;532:397–411. [PubMed: 19271198]
- Datta S, Larkin C, Schildbach JF. Structural insights into single-stranded DNA binding and cleavage by F factor TraI. *Structure* 2003;11:1369–1379. [PubMed: 14604527]
- 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. [PubMed: 8520220]
- Dostal L, Schildbach JF. Single-stranded DNA binding by F TraI relaxase and helicase domains are coordinately regulated. *J Bacteriol* 2010;92:3620–3628. [PubMed: 20435720]
- Garcillan-Barcia MP, Francia MV, de la Cruz F. The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* 2009;33:657–687. [PubMed: 19396961]
- Goddard, TC.; Kneller, DG. SPARKY. Vol. 3. University of California; San Francisco: 2008.
- Guogas LM, Kennedy SA, Lee JH, Redinbo MR. A novel fold in the TraI relaxase-helicase c-terminal domain is essential for conjugative DNA transfer. *J Mol Biol* 2009;386:554–568. [PubMed: 19136009]
- Henderson D, Meyer RJ. The primase of broad-host-range plasmid R1162 is active in conjugal transfer. *J Bacteriol* 1996;178:6888–6894. [PubMed: 8955311]
- Inamoto S, Yoshioka Y, Ohtsubo E. Site- and strand-specific nicking in vitro at oriT by the traY-traI endonuclease of plasmid R100. *J Biol Chem* 1991;266:10086–10092. [PubMed: 1645338]
- Inamoto S, Fukuda H, Abo T, Ohtsubo E. Site- and strand-specific nicking at oriT of plasmid R100 in a purified system: enhancement of the nicking activity of TraI (helicase I) with TraY and IHF. *J Biochem (Tokyo)* 1994;116:838–844. [PubMed: 7883759]
- Jung YS, Zweckstetter M. Mars—robust automatic backbone assignment of proteins. *J Biomol NMR* 2004;30:11–23. [PubMed: 15452431]
- Larkin C, Datta S, Harley MJ, Anderson BJ, Ebie A, Hargreaves V, Schildbach JF. Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. *Structure* 2005;13:1533–1544. [PubMed: 16216584]
- Larkin C, Haft RJ, Harley MJ, Traxler B, Schildbach JF. Roles of active site residues and the HUH motif of the F plasmid TraI relaxase. *J Biol Chem* 2007;282:33707–33713. [PubMed: 17890221]
- Llosa M, Grandoso G, de la Cruz F. Nicking activity of TrwC directed against the origin of transfer of the IncW plasmid R388. *J Mol Biol* 1995;246:54–62. [PubMed: 7853404]
- Llosa M, Grandoso G, Hernando MA, de la Cruz F. Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. *J Mol Biol* 1996;264:56–67. [PubMed: 8950267]
- Matson SW, Nelson WC, Morton BS. Characterization of the reaction product of the oriT nicking reaction catalyzed by *Escherichia coli* DNA helicase I. *J Bacteriol* 1993;175:2599–2606. [PubMed: 8386720]
- Pansegau W, Ziegelin G, Lanka E. Covalent association of the traI gene product of plasmid RP4 with the 5'-terminal nucleotide at the relaxation nick site. *J Biol Chem* 1990;265:10637–10644. [PubMed: 2191955]
- Pansegau W, Schroder W, Lanka E. Relaxase (TraI) of IncP alpha plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proc Natl Acad Sci USA* 1993;90:2925–2929. [PubMed: 8385350]
- Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 2009;44:213–223. [PubMed: 19548092]
- Sherman JA, Matson SW. *Escherichia coli* DNA helicase I catalyzes a sequence-specific cleavage/ligation reaction at the F plasmid origin of transfer. *J Biol Chem* 1994;269:26220–26226. [PubMed: 7929337]
- Stern JC, Schildbach JF. DNA recognition by F Factor TraI36: highly sequence-specific binding of single-stranded DNA. *Biochemistry* 2001;40:11586–11595. [PubMed: 11560509]
- Street LM, Harley MJ, Stern JC, Larkin C, Williams SL, Miller DL, Dohm JA, Rodgers ME, Schildbach JF. Subdomain organization and catalytic residues of the F factor TraI relaxase domain. *Biochim Biophys Acta* 2003;1646:86–99. [PubMed: 12637015]

- Traxler BA, Minkley EG Jr. Evidence that DNA helicase I and *oriT* site-specific nicking are both functions of the F TraI protein. *J Mol Biol* 1988;204:205–209. [PubMed: 2851049]
- Wright NT, Varney KM, Ellis KC, Markowitz J, Gitti RK, Zimmer DB, Weber DJ. The three-dimensional solution structure of Ca(2+)-bound S100A1 as determined by NMR spectroscopy. *J Mol Biol* 2005;353:410–426. [PubMed: 16169012]

Fig. 1. 2D ^1H - ^{15}N HSQC of Ca^{2+} -bound S100A1 at 600 MHz (^1H) with the assignments included. Backbone ^1H - ^{15}N correlations are labeled sequence-specifically, and correlations connected by *horizontal lines* correspond to glutamine and asparagine side-chain NH_2 groups

Fig. 2.

a Strips derived from ^{15}N planes of a 3D ^{15}N -edited CBCA(CO)NH (left strip of each pair) and a 3D ^{15}N -edited HNCACB spectrum (right strip of each pair) illustrating through-bond connectivities of the C_α and C_β carbon atoms from residue N529 to G538. **b** Secondary structure of TraI (381–569) predicted by TALOS+. The output from the program TALOS+ (Shen et al. 2009) is displayed as a column chart (blue, α -helix; red, β -strand) and column height reflects the probability assigned by the prediction program of TALOS+ (β -strand values are negative for clarity)