Solution NMR Structure of the TatA Component of the Twin-Arginine Protein Transport System from Gram-Positive Bacterium Bacillus subtilis

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · NOVEMBER 2010

Impact Factor: 12.11 · DOI: 10.1021/ja1053785 · Source: PubMed

5 AUTHORS, INCLUDING:

Yunfei Hu
Peking University
26 PUBLICATIONS **369** CITATIONS
SEE PROFILE

Bin Xia
Peking University
135 PUBLICATIONS **2,235** CITATIONS
SEE PROFILE

Changwen Jin
Peking University
103 PUBLICATIONS **1,213** CITATIONS
SEE PROFILE

J|A|C|S COMMUNICATIONS

Published on Web 08/20/2010

Solution NMR Structure of the TatA Component of the Twin-Arginine Protein Transport System from Gram-Positive Bacterium *Bacillus subtilis*

Yunfei Hu,[†,‡] Enwei Zhao,[†] Hongwei Li,[‡,§] Bin Xia,[†,‡,§,||] and Changwen Jin*[,†,‡,§,||]

Beijing NMR Center, College of Life Sciences, College of Chemistry and Molecular Engineering, Beijing National Laboratory for Molecular Sciences, Peking University, Beijing 100871, China

Received June 19, 2010; E-mail: changwen@pku.edu.cn

Abstract: The twin-arginine transport (Tat) system translocates folded proteins across the bacterial cytoplasmic or chloroplast thylakoid membrane of plants. The Tat system in most Gram-positive bacteria consists of two essential components, the TatA and TatC proteins. TatA is considered to be a bifunctional subunit, which can form a protein-conducting channel by self-oligomerization and can also participate in substrate recognition. However, the molecular mechanism underlying protein translocation remains elusive. Herein, we report the solution structure of the $TatA_d$ protein from *Bacillus subtilis* by NMR spectroscopy, the first structure of the Tat system at atomic resolution. $TatA_d$ shows an L-shaped structure formed by a transmembrane helix and an amphipathic helix, while the C-terminal tail is largely unstructured. Our results strongly support the postulated topology of $TatA_d$ in which the transmembrane helix is inserted into the lipid bilayer while the amphipathic helix lies at the membrane−water interface. Moreover, the structure of $TatA_d$ revealed the structural importance of several conserved residues at the hinge region, thus shedding new light on further elucidation of the protein transport mechanism of the Tat system.

Protein translocation across cellular membranes is vital to all living organisms. The Twin-Arginine Transport (Tat) system identified in bacteria and plant chloroplasts transports proteins in their fully folded states.[1,2] It is therefore challenged with the unique task of transporting target proteins of variable sizes while maintaining the membrane permeability barrier. Despite extensive research efforts, the molecular mechanism of the Tat system is still largely unknown. In particular, atomic-resolution structures of any components of this system are still lacking.

In Gram-negative bacteria (e.g., *Escherichia coli*) and plant chloroplasts, the minimal functional Tat translocase consists of three transmembrane proteins TatA, TatB, and TatC.[2] The TatA component is a predicted single-pass transmembrane protein that may form a protein conducting channel by self-oligomerization. TatB is homologous to TatA but functions together with TatC as the receptor for target proteins. Interestingly, most Gram-positive bacteria harbor multiple functional TatAC systems while lacking the TatB gene.[2,3] In *Bacillus subtilis*, a $TatA_d$−$TatC_d$ system specifically transports the phosphodiesterase PhoD.[3] Experimental evidence strongly suggests that *B. subtilis* $TatA_d$ ($BsTatA_d$) protein acts both as a receptor (similar to *E. coli* TatB) and as a pore-forming component (similar to *E. coli* TatA).[4,5] Therefore, structural information of $BsTatA_d$ protein is essential for elucidating the Tat translocation mechanism. Solution NMR studies of membrane protein structures have greatly advanced in recent years.[6] Herein, we report the solution NMR structure of $BsTatA_d$ in dodecylphosphocholine (DPC) micelles.

Complete backbone and near complete side-chain chemical shift assignments were obtained for the 70-residue full-length $BsTatA_d$ protein in deuterated DPC micelles (SI Figure 1). The chemical shift assignments have been deposited in the BioMagResBank (accession number 17075). Analysis of chemical shifts and NOE patterns demonstrated that the N-terminal region (Met1-Val48) is highly helical, while the C-terminal tail (Ser49-Gly70) is largely unstructured. The high content of helical structure in $BsTatA_d$ was further supported by the CD spectrum (SI Figure 2).

The structure calculation was carried out in two separate stages. In the first stage, the initial structures were calculated using only dihedral angle and NOE-derived distance restraints. In the second stage, backbone N−H residual dipolar couplings (RDCs) were used as a structural validation tool. Two hundred structures were calculated and 40 conformers which best satisfy both RDC and NOE data were selected to represent the $BsTatA_d$ protein (see SI Figure 3 and SI methods for more details). The atomic coordinates have been deposited in the Protein Data Bank (PDB code 2l16). As shown in Figure 2, the protein adopts an L-shaped structure formed by two helices, the transmembrane helix (TMH) and the amphipathic helix (APH). The TMH is comprised of 13 residues (Pro8-Ile20), which is in agreement with secondary structural prediction and previously reported structural characterization by solid-state NMR.[7] Residues Leu26-Val48 form the APH, which is slightly longer than predicted (residues Glu28-Lys45). RDC data contain long-range orientation information which could not be obtained from NOE restraints; therefore it is highly valuable for structure validation. The majority of the conformers calculated based on NOE information fit well with the RDC data, showing a correlation coefficient R value larger than 0.80 (SI Figures 3−4). Moreover, the backbone $\{^1H\}$-^{15}N heteronuclear NOE data (Figure 1E) showed that residues in TMH, the N-terminal half of APH, and the hinge region have generally similar flexibility on ps−ns time scales, indicating that the L-shaped structure is relatively stable.

Furthermore, RDC data are especially useful for identification of nonideal conformations in helices, such as kinks or curvatures. Fitting the experimentally measured N−H RDC values to sinusoids (Figure 1D) revealed a slight phase shift at residue Ile16 in TMH, reflecting a kink at this position. This structural feature is observable in the structure ensembles calculated without RDC data. Residues Ile16-Ile20 form the last turn of the TMH and is tilted about 14° compared to the rest of the helix (SI Figure 5). In addition, a curvature is present in the APH segment judged by the gradual change of amplitude and offset of the dipolar wave function along

[†] Beijing NMR Center.
[‡] College of Life Sciences.
[§] College of Chemistry and Molecular Engineering.
[||] Beijing National Laboratory for Molecular Sciences.

10.1021/ja1053785

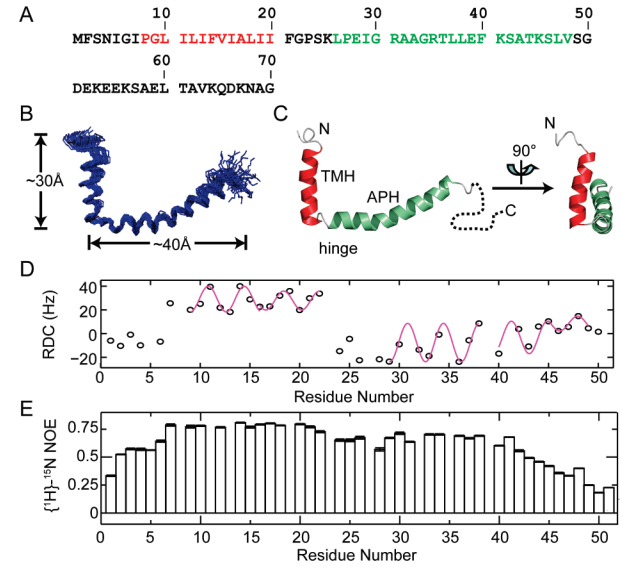

Figure 1. Solution structure of *B. subtilis* $TatA_d$. (A) Primary structure of $BsTatA_d$ with the TMH and APH segments colored in red and green, respectively. (B) Ensemble of the 40 representative conformers of $BsTatA_d$ (segment Met1-Glu52) with dimensions labeled. (C) Ribbon diagram of the solution structure of monomeric $BsTatA_d$ viewed from two angles. The C-terminal unstructured region (E52-G70) is schematically represented by a dashed line in the left panel. (D) Experimentally measured RDC values and fitting to sinusoid function using a fixed periodicity of 3.6 (shown in magenta) using a MATLAB script.[8] (E) Backbone $\{^1H\}$-^{15}N heteronuclear NOE values as a function of residue number.

the helix. In particular, the C-terminal part of APH (residues Ser42-Val48) shows a significantly reduced amplitude of the sinusoid function, indicating a change of the helix orientation or increased dynamics.

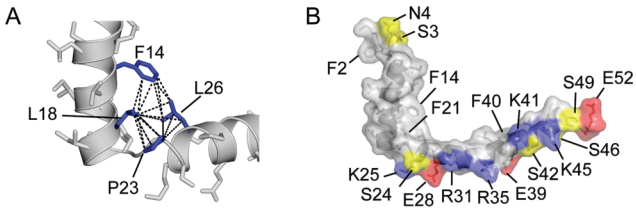

Figure 2. Structural analysis of *B. subtilis* $TatA_d$. (A) Local conformation of the hinge region. Close interactions based on H−H NOE cross-peaks are shown as dashed lines. For clarity, not all observed interactions are shown. (B) Surface representation of $BsTatA_d$. The positively charged, negatively charged, and polar residues are colored blue, red, and yellow, respectively. The positions of the four aromatic Phe residues are shown.

Careful analysis of the NOESY spectrum revealed that the L-shaped structure is largely stabilized by extensive contacts in the hinge region between TMH and APH, particularly the interactions among residues Leu18, Pro23, and Leu26 (Figure 2A). Sequence analysis indicated that the hinge region contains the most highly conserved residues for both TatA and TatB families. It has been shown that the $BsTatA_d$ protein displays sequence similarity to both *E. coli* TatA and TatB proteins (SI Figure 6). On one hand, $BsTatA_d$ possess the invariant FG dipeptide sequence of the TatA family in the hinge; on the other hand, it also contains the conserved GPxxLP motif of the TatB family. Therefore, the strictly conserved proline residues as well as the hydrophobic leucine residue in the GPxxLP motif have vital roles in maintaining the local conformation of the hinge, which may be of important biological function. In the structure of $BsTatA_d$, both residues Pro23 and Leu26 in the $G_{22}P_{23}xxL_{26}P_{27}$ motif greatly contribute to the local conformation of the hinge region. In addition, the aromatic ring of Phe14 from the TMH is also oriented toward the hinge and interacts with the hydrophobic side chains of Leu18 and Leu26, which further stabilizes the structure (Figure 2A). It is interesting to observe that, in the sequence of *E. coli* TatB, there is also a phenylalanine residue at a position equivalent to that of Phe14 in $BsTatA_d$. Therefore, it is possible that these residues are structurally important for maintaining the protein conformation, which might be required for the function of TatB family proteins. Alternatively, they may have specific roles in the processes of substrate recognition and/or interaction with the TatC component.

Mapping of the hydrophilic residues onto the surface representation of the $BsTatA_d$ structure clearly shows the amphipathic nature of the APH (Figure 2B). Analysis of NOE cross-peaks to the water signal demonstrated that the TMH is fully buried in detergent micelles without access to water, while the APH segment is partially exposed to water. Most of the hydrophobic residues in the N-terminal region of APH (residues Leu26-Phe40) showed no NOE cross-peaks to water in the NOESY spectra. Some of the hydrophilic residues can be exposed to water. In particular, the side chain hydrogen atoms of Arg31 and Arg35 showed gradually increased NOE signals to water from H^{α} to H^{δ}. In addition, the residues in the C-terminal region of APH (Lys41-Val48) are generally accessible to water, indicating that this region possibly has higher structural flexibility, which is also suggested by the reduced RDC values for this region as discussed above. Moreover, the decreased backbone $\{^1H\}$-^{15}N NOE values of this region clearly demonstrate its high flexibility (Figure 1E). These data together suggest that the C-terminal region of APH (Lys41-Val48) samples larger conformational spaces, while the calculated structures may reflect the averaged conformation. Therefore, in a lipid bilayer environment, the APH could adopt a straight helical conformation on the membrane surface with its hydrophobic side facing the hydrophobic fatty acid chains, while the polar and charged residues interact with the lipid head groups or are exposed to the aqueous environment. Previous studies on $BsTatA_d$ structure using both oriented CD and solid-state NMR methods indicated that the APH segment is oriented parallel to the lipid bilayer.[7,9] It has been suggested that TatA protein self-oligomerizes on the cell membrane and forms protein transportation pores. Electron microscopy of the membrane embedded $BsTatA_d$ protein suggested circular shaped particles of homogeneous size, with a diameter of about 10 nm.[4] In a stretched conformation, the length of the APH segment can extend to ~40 Å, which matches the radius of the complex particle (SI Figure 7).

In summary, we determined the solution NMR structure of *B. subtilis* $TatA_d$ protein, the first 3D structure of the Tat system at atomic resolution. This highlights the structural importance of several conserved residues and provides the basic building block for further modeling of the protein conducting channel complex. Moreover, the structure and resonance assignments offer a good starting point for investigations on substrate binding and TatA−TatC interactions.

Acknowledgment. All NMR experiments were carried out at the Beijing Nuclear Magnetic Center (BNMRC), Peking University. The CD spectroscopy experiments were performed at Prof. Luhua Lai's lab, Peking University. MATLAB scripts for the fitting of dipolar waves were kindly provided by Prof. Stanley Opella, University of California San Diego. This work was supported by Grant 2006AA02A323 from the National High Technology Research and Development Program and Grant 2006CB910203 from the National Basic Research Program of China to C.J.

Supporting Information Available: SI Table 1: Structural statistics; SI Figures 1−3: Backbone assignments, CD spectroscopy, and structural analyses of $BsTatA_d$; SI Figure 4: Fitting of RDC data; SI Figure 5: Representation of the kink in TMH; SI Figure 6: Multiple sequence alignment; SI Figure 7: Schematic model of membrane bound $BsTatA_d$ complex; SI Methods: Sample preparation, NMR spectroscopy, RDC measurements, heteronuclear NOE measurements, and structure calculation. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Natale, P.; Brüser, T.; Driessen, A. J. M. *Biochim. Biophys. Acta* **2008**, *1778*, 1735.
(2) Sargent, F. *Biochem. Soc. Trans.* **2007**, *35*, 835.
(3) Jongbloed, J. D. H.; Grieger, U.; Antelmann, H.; Hecker, M.; Nijland, R.; Bron, S.; van Dijl, J. M. *Mol. Microbiol.* **2004**, *54*, 1319.
(4) Westermann, M.; Pop, O. I.; Gerlach, R.; Appel, T. R.; Schlörmann, W.; Schreiber, S.; Müller, J. P. *Biochim. Biophys. Acta* **2006**, *1758*, 443.
(5) Barnett, J. P.; Eijlander, R. T.; Kuipers, O. P.; Robinson, C. *J. Biol. Chem.* **2008**, *283*, 2534.
(6) Kim, H. J.; Howell, S. C.; Van Horn, W. D.; Jeon, Y. H.; Sanders, C. *Prog. Nucl. Magn. Reson. Spectrosc.* **2009**, *55*, 335.
(7) Müller, S. D.; De Angelis, A. A.; Walther, T. H.; Grage, S. L.; Lange, C.; Opella, S. J.; Ulrich, A. S. *Biochim. Biophys. Acta* **2007**, *1768*, 3071.
(8) Mesleh, M. F.; Opella, S. J. *J. Magn. Reson.* **2003**, *163*, 288.
(9) Lange, C.; Müller, S. D.; Walther, T. H.; Bürck, J.; Ulrich, A. S. *Biochim. Biophys. Acta* **2007**, *1768*, 2627.

JA1053785