JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Solid-State NMR on a Large Multidomain Integral Membrane Protein: The Outer Membrane Protein Assembly Factor BamA

Marie Renault,[†] Martine P. Bos,[‡] Jan Tommassen,[‡] and Marc Baldus*[,†]

[†]Bijvoet Center for Biomolecular Research and [‡]Department of Molecular Microbiology, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

Supporting Information

ABSTRACT: Multidomain proteins constitute a large part of prokaryotic and eukaryotic proteomes and play fundamental roles in various physiological processes. However, their structural characterization is challenging because of their large size and intrinsic flexibility. We show here that motional-filtered high-resolution solid-state NMR (ssNMR) experiments allow for the observation and structural analysis of very large multidomain membrane proteins that are characterized by different motional time scales. This approach was used to probe the folding of the 790-residue membrane protein BamA, which is the core component of the *Escherichia coli* outer membrane protein assembly machinery. A combination of dipolar- and scalar-based two-dimensional ssNMR experiments applied to two uniformly $^{13}C,^{15}N$-labeled BamA variants revealed characteristic secondary structure elements and distinct dynamics within the BamA transmembrane protein segment and the periplasmic POTRA domains. This approach hence provides a general strategy for collecting atomic-scale structural information on multidomain (membrane) proteins in a native-like environment.

Multidomain or modular proteins constitute a large part of prokaryotic and eukaryotic proteomes and play fundamental roles in various physiological processes.[1] A detailed understanding of their function relies notably on the analysis of their supramodular architecture, dynamics, and time-resolved interactions with molecular partners at atomic resolution. However, classic structural biology methods are not generally suitable for the investigation of multidomain proteins because of their large size and intrinsic flexibility.

Such molecular properties also characterize the integral membrane protein BamA, which belongs to the family of bacterial outer membrane proteins (OMPs). BamA is the core component of the *Escherichia coli* β-barrel assembly machinery.[2,3] Its mature form consists of 790 residues organized in six distinct domains: a carboxy-terminal transmembrane (TM) domain and an amino-terminal periplasmic extension encompassing five polypeptide-transport-associated (POTRA) domains that interact with OMP substrates and accessory lipoproteins of the Bam complex[4−7] (Figure 1a). While significant structural information has been obtained for isolated POTRA domains,[6,8,9] attempts to acquire atomic-scale structural information on the full-length protein by solution-state NMR spectroscopy[10] and X-ray crystallography have to date been unsuccessful. An important aspect in such studies relates to the fold of the POTRA domains in the entire construct and their motional degrees of freedom.

Recently, progress in solid-state NMR (ssNMR) techniques has allowed the study of noncrystalline and complex biomolecules such as multidomain membrane proteins in a biologically relevant context. Such methods have been successfully applied to probe motional disordered regions of membrane-associated proteins[11] and protein amyloids[12,13] and to reduce spectral complexity using molecular constructs of different size.[14] In the present work, we investigated the use of such methods to probe structural features of distinct topological regions of BamA.

As previously demonstrated,[7] BamA lacking its signal sequence can be overexpressed in *E. coli* as intracellular inclusion bodies, subsequently refolded into its native state in the presence of detergent molecules, and incorporated into lipid vesicles at a high lipid/protein molar ratio (i.e., >800:1). For sensitivity reasons, such conditions are not compatible with conventional ssNMR structural studies on the 88 kDa full-length BamA protein. Alternatively, in vitro-refolded BamA can be precipitated under native-like conditions before ssNMR analysis [see the Supporting Information (SI)]. To reduce spectral complexity, we furthermore prepared proteoliposomes of a uniformly ($^{13}C,^{15}N$)-labeled sample of a truncated version of BamA, namely, BamA(ΔP1−P4) comprising the POTRA5 and TM domains, both of which are essential for BamA function[4] (Figure 1a). In each sample, the proteins displayed characteristic heat modifiability, with up to ∼90% of the protein in its native-like conformation (Figure S1 in the SI).

To characterize rigid and mobile protein segments, we acquired two-dimensional (2D) ($^{13}C,^{13}C$) and ($^{1}H,^{13}C$) correlation spectra employing dipolar and scalar magnetization transfer steps, respectively.[11] Figure 1b shows a comparison between 2D $^{13}C-^{13}C$ proton-driven spin diffusion (PDSD) spectra obtained using a short mixing time (40 ms) that were recorded on uniformly ($^{13}C,^{15}N$)-labeled BamA precipitate (black) and reconstituted BamA(ΔP1−P4) (blue). Overall, the spectra display good dispersion and sensitivity and comparable spectral resolution. Remarkably, BamA and its truncated variant exhibit similar intraresidue CC correlation patterns, suggesting that rigid components are largely conserved in the two preparations. According to standard amino acid-specific peak positions[15] and the overall correlation patterns, these rigid protein segments predominantly reflect protein segments in β-sheet and loop conformations in accordance with the topological profile of β-barrel TM

Received: October 21, 2010
Published: March 01, 2011

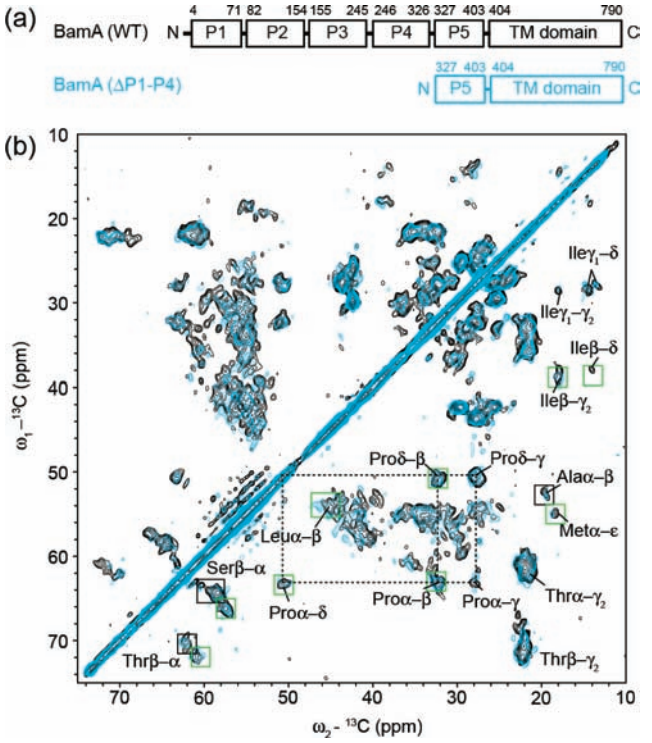

Figure 1. (a) Domain organization within the full-length BamA protein (WT) and the BamA(ΔP1−P4) variant. (b) Overlay of the aliphatic region of the 2D $^{13}C-^{13}C$ PDSD spectra (mixing time 40 ms) recorded on BamA precipitate (black) and reconstituted BamA(ΔP1−P4) variant (blue) with 10.92 kHz magic-angle spinning (MAS) and an ^{1}H Larmor frequency of 700 MHz. Amino acid-specific assignments are indicated. Characteristic β-sheet-like or putative loop CC correlations are shown in green and black squares, respectively. See the SI for close-ups of these spectra.

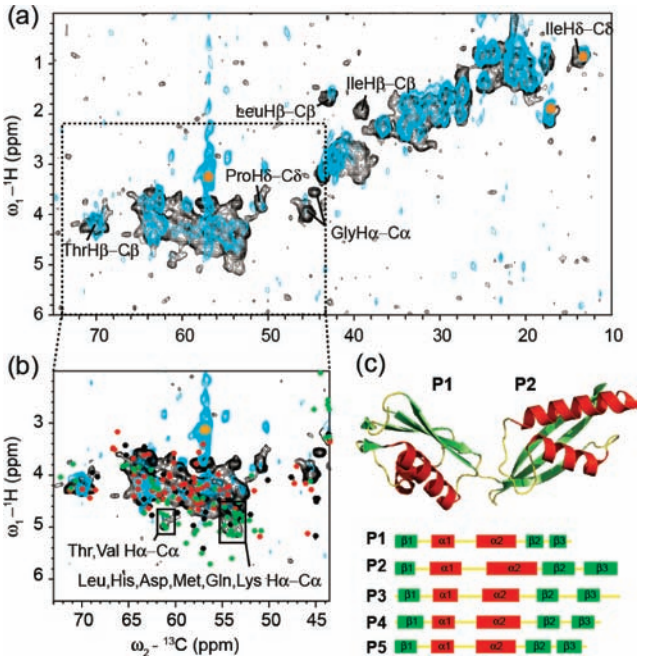

Figure 2. Flexible BamA POTRA domains revealed by scalar-based HC correlation experiments. (a) Overlay of 2D $^{1}H-^{13}C$ INEPT spectra (optimized $^{1}J_{HC}$ = 155 Hz) recorded on BamA precipitate (black) and reconstituted BamA(ΔP1−P4) variant (blue) with 10.92 kHz MAS and an ^{1}H Larmor frequency of 700 MHz. Amino acid-specific assignments are indicated. (b) Selected region of the 2D $^{1}H-^{13}C$ INEPT spectra showing backbone HC correlations of BamA proteins. Characteristic β-sheet-like Hα−Cα correlations are shown in squares. Lipid resonances are indicated with orange stars. Available solution NMR chemical shifts of the POTRA1 and POTRA2 domains (BMRB entry 15247; see ref 9) are indicated with colored circles (α-helix in red, β-sheet in green, and random coil in black). (c) Top panel: Cartoon representation of the solution NMR structure of isolated BamA POTRA domains P1−P2 (PDB entry 2V9H; see ref 9) used for ssNMR data analysis. Bottom panel: distribution of secondary structure elements present within the POTRA domains of BamA, as seen in the available high-resolution 3D structures of isolated BamA POTRA domains (PDB entries 2V9H,[9] 2QDF,[6] and 3OG5[16]). See the SI for close-ups of these spectra.

domains of OMPs. In contrast, we observed clear differences between the two protein preparations in scalar-based (^{1}H,^{13}C) correlation experiments (Figure 2a). Interestingly, the overall chemical shift dispersion in the Hα−Cα region (ω_1 = 3.5−5.3 ppm; ω_2 = 52−66 ppm) indicates that both proteins exhibit dynamic segments with β-sheet and α-helical secondary structure elements, which typify the periplasmic POTRA domains (Figure 2b). We compared our results to structural data obtained on isolated BamA POTRA domains (Figure 2c). The remarkable agreement between the P1−P2 solution NMR chemical shifts [Figure 2b; Biological Magnetic Resonance Data Bank (BMRB) entry 15247], the P3−P5 predicted chemical shifts from available crystal structures (see the SI), and the overall $^{1}H-^{13}C$ INEPT correlation pattern of full-length BamA strongly suggests that the global fold of the POTRA domains is conserved within the full-length protein. As judged by their characteristic ^{1}H and ^{13}C chemical shift dispersion, β-sheet protein segments of P1−P2 POTRA domains typically consisting of polar Thr, His, Asp, Lys, and Gln and nonpolar Val, Pro, Met, and Leu amino acids are particularly apparent in the full-length protein. In contrast, the corresponding signals are strongly attenuated in the BamA(ΔP1−P4) spectrum, which correlates with the differential amino acid distribution of the POTRA domains between the two constructs. Indeed, removal of P1−P4 should lead to a suppression of 81% of Gly and Thr, 82% of Pro, 85% of Leu, and 95% of Ile residues in POTRA segments, in line with our spectroscopic observations.

Finally, to further investigate the structural features of BamA POTRA domains, we conducted 2D $^{1}H-(^{1}H)^{13}C$ INEPT experiments on uniformly (^{13}C,^{15}N)-labeled full-length BamA using different $^{1}H-^{1}H$ mixing schemes. We observed only a few intraresidue cross-peaks using NOESY[17] mixing times between 40 and 100 ms. On the other hand, an RFDR[18] proton recoupling scheme substantially improved the intraresidue and sequential transfer efficiencies within the POTRA domain segments. Indeed, Figure 3 shows a comparison between aliphatic regions of the 2D HC-INEPT (black) and H(H)C-INEPT (red; RFDR mixing time of 40 ms) spectra of the BamA precipitate. The 2D HHC spectrum is characterized by a large number of various additional correlations corresponding to intraresidue and sequential protein contacts. Interestingly, the signals corresponding to intraresidue correlations exhibit symmetric line shapes and homogeneous ^{1}H and ^{13}C line widths of 0.1 and 0.6 ppm, respectively.

While the ^{13}C line width thus compares favorably to those typically seen in ssNMR studies on microcrystalline proteins, the observed ^{1}H line width indicates that anisotropic protein motion leads to a significant reduction of $^{1}H-^{1}H$ dipolar couplings that

dx.doi.org/10.1021/ja109469c | J. Am. Chem. Soc. 2011, 133, 4175–4177

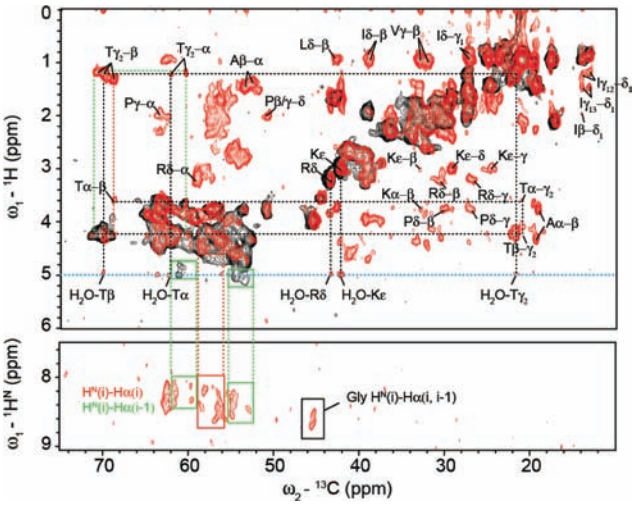

Figure 3. Overlay of the aliphatic regions of 2D $^{1}H-^{13}C$ INEPT (black) and 2D $^{1}H-(^{1}H)^{13}C$ INEPT using a RFDR mixing time of 40 ms (red) spectra recorded on BamA precipitate with 10.92 kHz MAS and an ^{1}H Larmor frequency of 700 MHz. Some typical intraresidue correlations are indicated in the spectrum. Characteristic HN−C correlations of α-helical (red) or β-sheet-like (green) protein segments are shown in squares. A dashed blue line indicates the water ^{1}H resonance. See the SI for close-ups of these spectra.

usually dominate ^{1}H spectra of solid proteins. Moreover, typical intraresidue and sequential HN−Cα correlations are observed for residues located in β-sheet and α-helical protein segments, which further demonstrates the presence of folded and flexible globular POTRA domains within the full-length protein. Along the same lines, only a subset of BamA resonances (i.e., essentially those corresponding to polar residues) are correlated to the water ^{1}H signal, indicating that residues located in secondary structure elements are confined in the hydrophobic core of each POTRA domain.

In summary, we have demonstrated that high-resolution ssNMR spectroscopy allows the investigation of the molecular structure and characterization of the dynamics of large multidomain integral membrane proteins such as the 88 kDa outer membrane protein assembly factor BamA. Embedded in a membrane bilayer or in an insoluble folded state, BamA preparations studied here exhibit rigid transmembrane protein segments and highly flexible globular POTRA domains. These results demonstrate that 2D motional-filtered ssNMR experiments can not only be used to identify molecular domains that exhibit sizable motion but also help to characterize the folding of these flexible domains. Our results suggest that individual POTRA domains can maintain their fold even when associated with the entire TM segment of BamA.

In comparison with $^{1}H-^{1}H$ polarization transfer rates seen in solid proteins under comparable conditions,[19] our results using RFDR mixing suggest a reduction in proton−proton dipolar couplings within the periplasmic POTRA domains by about 2 orders of magnitude. On the other hand, the dominant contributions in our dipolar-based spin diffusion spectra suggest the absence of such submicrosecond motions for the transmembrane regions of BamA. Additional studies will be required for further refinement of the exact details[16] of the POTRA domain motion in different molecular environments and to probe motions on significantly shorter (ns) and longer (s) time scales. Nevertheless, the data presented here provide the basis for further structural and dynamical studies involving scalar- and dipolar-based 3D ssNMR experiments on the full-length BamA protein in synthetic lipids or other cellular-like[20] environments.

■ ASSOCIATED CONTENT

Supporting Information. Sample preparation procedures and additional solid-state NMR spectra recorded on uniformly $^{13}C,^{15}N$-labeled full-length BamA. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
m.baldus@uu.nl

■ ACKNOWLEDGMENT

The research leading to these results received funding from the European Community's Seventh Framework Program [FP7/2007-2013] under Grant Agreement n°211800 and was supported by grants from the NWO (700.26.121 and 815.02.012)

■ REFERENCES

(1) Vogel, C.; Bashton, M.; Kerrison, N. D.; Chothia, C.; Teichmann, S. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 208–216.
(2) Voulhoux, R.; Bos, M. P.; Geurtsen, J.; Mols, M.; Tommassen, J. *Science* **2003**, *299*, 262–265.
(3) Wu, T.; Malinverni, J.; Ruiz, N.; Kim, S.; Silhavy, T. J.; Kahne, D. *Cell* **2005**, *121*, 235–245.
(4) Bos, M. P.; Robert, V.; Tommassen, J. *EMBO Rep.* **2007**, *8*, 1149–1154.
(5) Hagan, C. L.; Kim, S.; Kahne, D. *Science* **2010**, *328*, 890–892.
(6) Kim, S.; Malinverni, J. C.; Sliz, P.; Silhavy, T. J.; Harrison, S. C.; Kahne, D. *Science* **2007**, *317*, 961–964.
(7) Robert, V.; Volokhina, E. B.; Senf, F.; Bos, M. P.; Van Gelder, P.; Tommassen, J. *PLoS Biol.* **2006**, *4*, e377.
(8) Gatzeva-Topalova, P. Z.; Walton, T. A.; Sousa, M. C. *Structure* **2008**, *16*, 1873–1881.
(9) Knowles, T. J.; Jeeves, M.; Bobat, S.; Dancea, F.; McClelland, D.; Palmer, T.; Overduin, M.; Henderson, I. R. *Mol. Microbiol.* **2008**, *68*, 1216–1227.
(10) Tamm, L. K.; Liang, B. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 201–210.
(11) Andronesi, O. C.; Becker, S.; Seidel, K.; Heise, H.; Young, H. S.; Baldus, M. *J. Am. Chem. Soc.* **2005**, *127*, 12965–12974.
(12) Andronesi, O. C.; von Bergen, M.; Biernat, J.; Seidel, K.; Griesinger, C.; Mandelkow, E.; Baldus, M. *J. Am. Chem. Soc.* **2008**, *130*, 5922–5928.
(13) Wasmer, C.; Schutz, A.; Loquet, A.; Buhtz, C.; Greenwald, J.; Riek, R.; Bockmann, A.; Meier, B. H. *J. Mol. Biol.* **2009**, *394*, 119–127.
(14) Etzkorn, M.; Kneuper, H.; Dunnwald, P.; Vijayan, V.; Kramer, J.; Griesinger, C.; Becker, S.; Unden, G.; Baldus, M. *Nat. Struct. Mol. Biol.* **2008**, *15*, 1031–1039.
(15) Wang, Y.; Jardetzky, O. *Protein Sci.* **2002**, *11*, 852–861.
(16) Gatzeva-Topalova, P. Z.; Warner, L. R.; Pardi, A.; Sousa, M. C. *Structure* **2010**, *18*, 1492–1501.
(17) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546–4553.
(18) Bennett, A. E.; Griffin, R. G.; Ok, J. H.; Vega, S. *J. Chem. Phys.* **1992**, *96*, 8624–8627.
(19) Lange, A.; Seidel, K.; Verdier, L.; Luca, S.; Baldus, M. *J. Am. Chem. Soc.* **2003**, *125*, 12640–12648.
(20) Renault, M.; Cukkemane, A.; Baldus, M. *Angew. Chem. Int. Ed.* **2010**, *49*, 8346–8357.