Plasticity of the β-barrel assembly machinery investigated by NMR

Tessa Sinnige

Plasticity of the β-barrel assembly machinery investigated by NMR

Plasticiteit van de β-barrel assembly machinery onderzocht met NMR

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op donderdag 11 september 2014 des middags te 12.45 uur

door

Tessa Sinnige

geboren op 8 augustus 1987 te Leiden

Promotor: Prof. dr. M. Baldus

TABLE OF CONTENTS

CHAPTER 1

General introduction

Membrane protein folding

Proteins are the molecular machines of living cells that operate in metabolism, signal transduction, immunity, and many other vital processes. In general, proteins have to acquire their correct three-dimensional fold in order to perform their function. Already in the 1960s, Anfinsen and co-workers realized that the protein fold is encoded in the amino acid sequence and that proteins can unfold and refold reversibly (reviewed in Ref. 1), yet how a polypeptide acquires its correct structure is still not fully understood. Levinthal's paradox states that if a nascent polypeptide chain would sample all possible conformations, it would take a time longer than the existence of the universe to reach the native state.[2] Hence, pathways must exist that guide the polypeptide chain towards its lowest free energy conformation.

For integral membrane proteins, which function as e.g. channels, receptors, and enzymes and constitute a large class of drug targets, folding mechanisms are particularly puzzling, as the polypeptide chain must not only fold into its correct structure, but must also be inserted into the lipid bilayer to expose its hydrophobic side chains towards the lipidic phase and its hydrophilic parts into the aqueous phase or the protein interior. α-helical membrane proteins that reside in eukaryotic membranes and bacterial cytoplasmic membranes are generally processed by the Sec translocon and are thought to fold and insert into the membrane helix by helix (reviewed in Ref. 3,4). For β-barrel membrane proteins, which are found exclusively

in the outer membranes of bacteria, mitochondria and chloroplasts, an analogous strand-by-strand mechanism is unlikely as the neighboring β-strands have to form hydrogen bonds with one another and exposure of individual strands to the lipidic phase is energetically unfavorable. *In vitro*, these proteins rather seem to adopt a collapsed structure on the membrane-water interface followed by complete insertion (see e.g. Ref. 5-9). *In vivo*, however, the situation is quite different and β-barrels depend on dedicated folding machineries for efficient folding and insertion into the membrane, as described in the following.

Biogenesis of Gram-negative outer membrane proteins

The cell envelope of Gram-negative bacteria consists of an inner and an outer membrane, separated by the periplasmic space which contains a layer of peptidoglycan (Fig. 1). The outer membrane acts as a barrier to protect the bacteria against harmful components from the environment. It differs from the inner membrane, which is a phospholipid bilayer, by its asymmetric composition, the inner leaflet consisting of phospholipids and the outer leaflet of lipopolysaccharide (LPS). To allow for nutrient acquisition, outer membrane proteins (OMPs) exist that form open pores through which small hydrophilic solutes can pass by diffusion. Other OMPs act e.g. as receptors for active nutrient acquisition, as transporters for the secretion of virulence factors or as enzymes that can, for example, modify the membrane composition.

Like any other protein, an OMP is synthesized on the ribosome in the cytoplasm and must be transported to the outer membrane where it is folded and inserted by the β-barrel assembly machinery (BAM) (Fig. 1). From the cytoplasm, it is targeted to the SecYEG machinery in the inner membrane by virtue of its N-terminal signal sequence. Whereas hydrophobic α-helices laterally exit the SecYEG complex into the inner membrane, the amphipathic nature of β-strands designates OMPs for transport into the periplasm.[3,4] Upon exit, the signal sequence of the nascent OMP is removed by signal peptidase.[10]

Once in the periplasm, the unfolded OMP is bound by chaperones to prevent aggregation (Fig. 1). In *Escherichia coli*, several periplasmic chaperones related to OMP biogenesis are known. The seventeen-kDa protein Skp is thought to act as a "holdase" for the unfolded substrate. It comprises a trimer with α-helical tentacles[11,12] that form a cavity in which the unfolded transmembrane domain of the OMP is held, whereas globular domains can protrude to the outside of the cavity.[13,14]

SurA on the other hand has a very different architecture. It contains two peptidyl-prolyl *cis/trans* isomerase (PPIase) domains,[15] yet these appear

dispensable for chaperone function.[16] SurA has been shown to bind peptides enriched in aromatic residues as well as unfolded OMPs.[17-19] It has been cross-linked to the BAM complex,[20] indicating that it actively delivers the substrate, as has also been shown in an *in vitro* assay.[21]

DegP, which has a dual role as a chaperone and protease,[22] is also related to OMP quality control (reviewed in e.g. Ref. 23,24). In its ground state, it forms inactive hexamers, whereas interaction with substrates has been shown to induce the formation of 12- and 24-mers.[25,26] When bound to phospholipids, large oligomeric structures are formed in the absence of substrate as well, yet their shapes are different from the oligomers found in solution.[27] Whether DegP can actually assist the folding of OMPs has not been demonstrated. Its involvement in correct β-barrel assembly seems mostly related to its protease function, which is required for the clearance of toxic misfolded OMPs.[28]
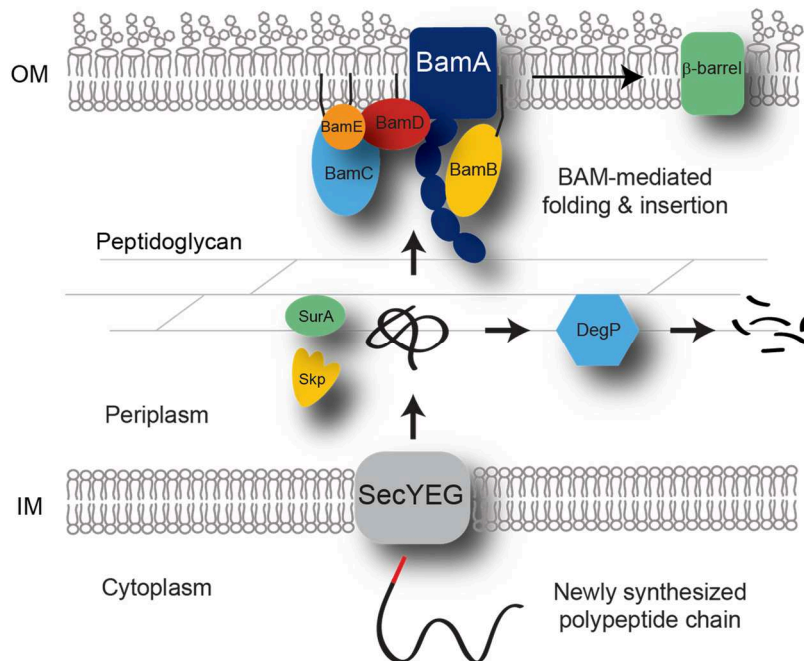
Figure 1. The cell envelope of *E. coli*, comprising the inner membrane (IM), outer membrane (OM) and periplasmic space. Nascent OMPs are transported through the SecYEG complex and assisted by chaperones Skp and SurA in the periplasm, whereas DegP degrades misfolded OMPs. The β-barrel assembly machinery, consisting of the integral membrane protein BamA and the lipoproteins BamB-E in *E. coli*, catalyzes folding and insertion of the β-barrel into the outer membrane.

Altogether, the individual roles and mutual relationships of these periplasmic chaperones are still under debate. According to one hypothesis, there are two main pathways, one involving SurA and the other Skp and DegP. SurA would be the main chaperone and Skp and DegP would serve to handle substrates that have fallen off the SurA pathway.[20] However, other groups found a profound effect on the levels of OMPs upon Skp deletion in *E. coli*,[29,30] pointing towards a more significant role for Skp. Possibly, Skp and SurA act sequentially: Skp protects the substrate from aggregation once it exits the Sec channel, and SurA delivers it to the BAM complex and assists folding.[31] The BAM complex, finally, mediates the folding and insertion of OMPs into the outer membrane (Fig. 1).[32,33]

Architecture of the β-barrel assembly machinery

The main component of the β-barrel assembly machinery, BamA, is highly conserved among Gram-negative bacteria and related proteins exist in the outer membranes of mitochondria (Sam50) and chloroplasts (Toc75-V).[34] BamA contains a β-barrel transmembrane (TM) domain and a periplasmic extension consisting of a number of POTRA (POlypeptide TRansport Associated) domains (Fig. 2).[35] The POTRA domains vary in sequence but share a conserved β–α–α–β–β fold. In *E. coli*, BamA has five POTRA domains that have been structurally well-characterized by X-ray crystallography and solution NMR.[36-40] In both crystal structures of POTRA 1-4, a short β-strand segment from POTRA 5 of a neighboring molecule binds to an exposed β-strand of POTRA 3, inspiring the hypothesis that unfolded substrates may bind to the POTRA domains by means of β-augmentation.[36,37] Solution NMR titrations of POTRA 1-2 with OMP peptides have indicated similar principles.[39]

The BamA TM domain comprises a 16-stranded β-barrel (Fig. 2), as revealed by crystal structures of full-length BamA from *Neisseria gonorrhoeae* and a truncated BamA construct from *Haemophilus ducreyi*,[41] as well as the BamA TM domain from *E. coli*.[42] Interestingly, the hydrophobic width of the BamA TM domain is highly asymmetric. Molecular Dynamics (MD) simulations indicated that on the shorter side of BamA, lipid disorder is induced and the lipid bilayer becomes thinner.[41] Furthermore, in the *N. gonorrhoeae* structure, the N- and C-terminal β-strands 1 and 16 show incomplete hydrogen bonding, which may allow the BamA β-barrel to laterally open, as was found in MD simulations.[41] On the extracellular side, however, the β-barrel is closed by the extracellular loops according to the crystal structures.

The *E. coli* BAM complex contains four lipoproteins in addition to BamA, BamB-E (Fig. 2). BamB has a β-propeller fold[43-46] and requires POTRA domains 2-4 to bind

to BamA.[36] Similarly to the POTRA domains, the BamB structure may allow for β-augmentation of the substrate.[47] Deletion of BamB only leads to mild defects in β-barrel assembly.[48] BamCDE are thought to form a sub-complex that is connected to BamA via a direct interaction of BamD with POTRA 5.[36,49,50] BamD is the only essential lipoprotein of the BAM complex in *E. coli*[49,51] and most other Gram-negative bacteria have a homolog of this lipoprotein (reviewed in Ref. 52). BamD contains five tetratricopeptide repeats (TPRs)[53,54] that are linked to protein-protein interactions.[55] Other TPR proteins have been shown to bind extended peptides[56,57] and BamD might similarly interact with unfolded OMPs. BamC contains two domains of similar fold[53,58] that have been hypothesized to localize to the outside of the outer membrane,[59] and an N-terminal extension that forms the main interaction site for BamD.[60] Its function, however, is unclear. Deletion does not seem to impair OMP biogenesis,[33] yet increases antibiotic sensitivity, suggesting a breach in the permeability barrier of the outer membrane.[51] BamE is the smallest lipoprotein of the complex being 10 kDa as a monomer, but dimeric BamE has also been identified.[53,61] Its oligomeric state within the BAM complex, however, is difficult to determine accurately due to its small size.[21] BamE is thought to provide stability to the BamCDE sub-complex[50,62] and has been shown to bind to phospholipids, which might be important for further tethering of the sub-complex to the membrane.[58]
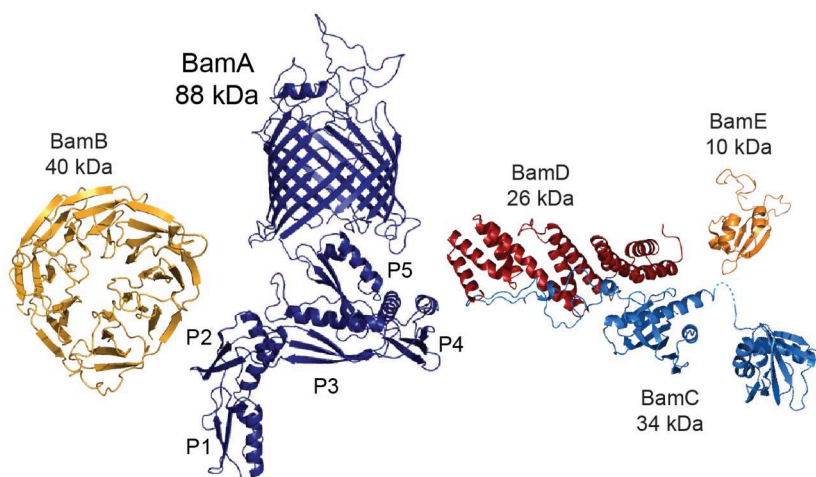
Figure 2. Structures of the *E. coli* BAM components with their molecular weights. The *E. coli* homology model of full-length BamA[41] is shown with POTRA domains P1-5 numbered from the N-terminus. BamB (PDB 3PIL) interacts with BamA independently; BamCDE form a sub-complex that binds to BamA via BamD. Shown are the BamC^{UN}D complex (PDB 3TGO) containing the N-terminal domain and unfolded extension of BamC and full-length BamD, the BamC C-terminal domain (PDB 3SNS) and BamE (PDB 2KXX).

Possible mechanisms of β-barrel assembly

Despite the availability of high-resolution structures of all BAM components from X-ray crystallography and solution NMR, the mechanism of β-barrel assembly is still unclear. Several hypotheses have been proposed based on experimental findings:

1. The substrate folds inside of the BamA β-barrel and is subsequently released into the outer membrane by lateral opening of the β-barrel (Fig. 3a). For a long time, opening was thought to be energetically unfavorable because hydrogen bonds would have to be broken, but the crystal structure of *N. gonorrhoeae* BamA revealed incomplete hydrogen bonding between β-strands 1 and 16 and MD simulations indicated that the β-barrel can open completely, supporting such a mechanism.[41] Furthermore, the BamA crystal structures of *N. gonorrhoeae* and *H. ducreyi* differ in the conformation of the POTRA domains with respect to the β-barrel, one closing the β-barrel and the other leaving it open. Exchange between these conformations could allow for opening and closure of the β-barrel interior in this model.[41] However, it is difficult to envision how the BamA β-barrel could accommodate very large substrates, e.g. the usher proteins PapC and FimD which consist of 24 β-strands.[63,64]

2. Alternatively, the substrate could be folded on the outside of BamA at the interface with the membrane. Insertion into the membrane could be promoted by the distortion of the lipid bilayer induced by BamA, as seen in MD simulations.[41] Lateral opening of BamA may also play a role in this process to form a template for the substrate β-strands to align, leading to a hybrid BamA-substrate β-barrel intermediate. Once substrate folding is complete, it is released from BamA which can then close again (Fig. 3b).

3. The BAM complex might also function as an oligomer. Higher order BamA oligomers have been observed,[65] but whether the entire BAM complex exists as oligomers remains to be determined. However, this would allow an attractive model in which BamA oligomers could form a proteinaceous cavity in which the substrate could fold, after which it is released into the membrane by dissociation of the oligomer (Fig. 3c).

4. Lastly, the possibility that the substrate folds in the periplasm prior to insertion into the outer membrane cannot be excluded. For the porin PhoE, it has been shown that the formation of tertiary structural elements can precede insertion into the membrane.[66] This mechanism may resemble the *in vitro* folding pathway of OMPs in which they form a collapsed structure on the membrane interface that is subsequently inserted.[5-9] In this model, the BAM complex would merely have a chaperoning function to shield the hydrophobic exterior of the folded substrate as long as it still resides in the aqueous phase.
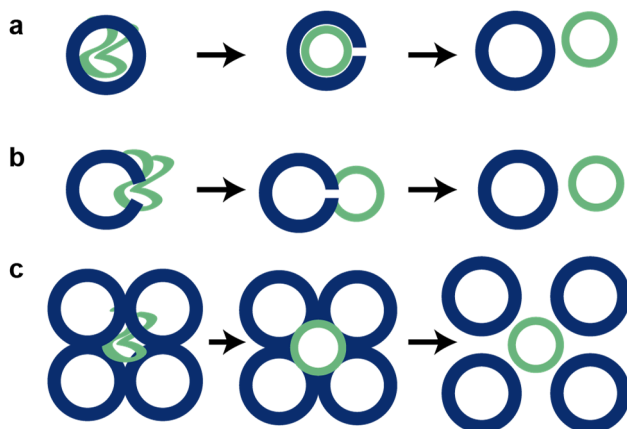
Figure 3. Possible mechanisms of β-barrel assembly, schematically depicted looking onto the membrane with BamA in blue and the substrate in green. a) The substrate folds inside the BamA β-barrel, which subsequently opens and releases the substrate into the membrane. b) BamA forms a hybrid β-barrel with the substrate. Once the substrate is folded, it is released into the membrane. c) BamA forms oligomers, creating a cavity in which the substrate can fold. Disassembly of the oligomer releases the substrate.

Some more clues are available independently of the possible models described above. The large extracellular loop 6 of BamA contains a highly conserved VRGF/Y motif that is essential for β-barrel assembly.[41,67] In the BamA crystal structures, loop 6 is well-ordered and folds slightly back into the lumen of the β-barrel,[41,42] yet in the structure of FhaC, which is also an Omp85 family member but functions in secretion, the loop extends all the way down to the periplasm.[68] *In vivo* labeling experiments have indicated that loop 6 cycles between an exposed and a protected state, meaning that conformational changes might indeed occur.[69] BamD and BamE have been proposed to play a role in this process.[69]

Furthermore, comparison between the available crystal structures of the POTRA domains suggests conformational flexibility in their inter-domain angles, especially between POTRA 2 and 3.[36-38] As mentioned above, the POTRA domains are thought to allow β-strand formation of the substrate by means of β-augmentation,[36,37,39] but conformational changes might, in addition, induce formation of β-hairpins which may represent an intermediate in β-barrel folding.[70]

In summary, further insights are needed regarding the conformational changes within BamA, the contribution of the lipoproteins and the mode of interaction with the substrate to explain the mechanism of β-barrel assembly by the BAM complex.

Solid-state NMR to study membrane proteins

Solid-state nuclear magnetic resonance (ssNMR) spectroscopy has the potential to study the structure, dynamics and interactions of membrane proteins in a native lipid bilayer environment at atomic resolution (reviewed in e.g. Ref. 71-73), and is therefore highly suitable to address the open questions regarding the mechanism of the β-barrel assembly machinery.

Membrane proteins that have been studied successfully by ssNMR include the bacterial potassium channel KcsA and a chimeric version, KcsA-Kv1.3, for which e.g. different conformations in the gating cycle have been revealed,[74,75] interactions with a toxin have been identified[76] and specific lipid binding has been studied.[77] The seven-helix receptor bacteriorhodopsin, which is a light-driven ion pump, has also been studied extensively by ssNMR, leading to better understanding of its photocycle (reviewed in Ref. 78). Other recent achievements are the ssNMR structure determination of the trimeric *Anabaena* sensory rhodopsin (ASR)[79] and the β-barrel transmembrane domain of the *Yersinia enterocolitica* adhesin A (YadA).[80]

ssNMR cannot only be used to study membrane proteins in synthetic lipid bilayers, but also in their native membranes and even in entire cells. The outer membrane enzyme PagL from *Pseudomonas aeruginosa* was studied in cell envelopes and intact bacterial cells.[81,82] In-cell ssNMR was also demonstrated for the soluble proteins thioredoxin and FKBP.[83]

Several challenges, however, are associated with high-resolution ssNMR studies on (large) membrane proteins. NMR spectroscopy intrinsically suffers from low sensitivity due to the small energy difference ΔE between the spin states α and β that is given by the product of the gyromagnetic ratio, reduced Planck constant and the magnetic field for spin ½ nuclei in the presence of an external magnetic field. Hence, the degree of polarization between the spin states remains small according to the Boltzmann equation,

$$\frac{n_\beta}{n_\alpha} = e^{-\Delta E/k_B T}$$

where n_β and n_α are the populations of the spin states, k_B is the Boltzmann constant and T is the temperature.

In solids, further complications arise from interactions that are orientation-dependent and not averaged out in the absence of molecular tumbling. In spin ½ systems these interactions relate to the dipolar couplings between nuclei close together in space, and the chemical shift anisotropy (CSA) caused by asymmetry of the electron cloud surrounding a nucleus. Dipolar couplings and the symmetric CSA tensor have a dependency of $3\cos^2\theta-1$ with respect to the external magnetic field,

B_0, which can be averaged out by Magic-angle-spinning (MAS) of the sample at an angle $\theta_m = \arccos \frac{1}{\sqrt{3}} \approx 54.73°$ (Fig. 4),[84,85] given that the spinning speed is at least equal to the size of the interaction. For example, strong ^1H-^1H dipolar couplings cannot be averaged out in solid, fully-protonated samples using currently available spinning frequencies of up to 100 kHz MAS, meaning that high levels of deuteration are required for high-resolution proton-detected ssNMR (see e.g. Ref. 86-89).
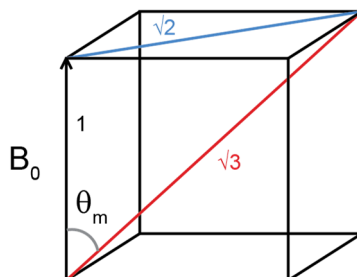

Figure 4. At the magic angle θ_m the orientation dependency of dipolar couplings and the symmetric CSA tensor with respect to the magnetic field B_0, $3\cos^2\theta-1$, cancels out.

Apart from MAS, protein dynamics can influence the effective magnitude of NMR interactions, meaning that different motional regimes can readily be separated.[90,91] Methods that rely on the presence of dipolar couplings, such as cross-polarization (CP)[92,93] or proton-driven spin diffusion (PDSD)[94,95] will only work when protein segments are relatively rigid, but will fail if motion occurs that averages out dipolar couplings. On the other hand, transfer via scalar couplings (that are much smaller than dipolar couplings) such as INEPT[96] only reveals segments that are mobile on the nanosecond timescale.

Scope of the thesis

In this thesis, NMR studies on the main components of the BAM complex and interactions among them are described. These include ssNMR investigations on the global dynamics of the POTRA domains in full-length BamA reconstituted in lipid bilayers (**Chapter 3**), instigated by the suggestion that large scale conformational rearrangements of the POTRA domains may provide a mechanism to open and close the BamA β-barrel.[41] Also the interplay between the BamA β-barrel and the lipid bilayer is described in this chapter.

For in-depth studies on the structure and dynamics of the BamA TM domain, e.g. the proposed lateral opening[41] and possible conformational changes of loop 6,[69] the ssNMR signals of this domain need to be assigned, which is not trivial for a membrane protein of this size (390 residues for the TM domain). **Chapter 4** describes the strategies that we employed towards this end, as well as the first insights into dynamics within the BamA TM domain.

Furthermore, using solution as well as ssNMR, we identified local conformational exchange in the BamA POTRA 5 domain (**Chapter 5**), which occupies a critical position at the interface of the BamA β-barrel and the membrane, where it interacts with BamD and most likely also with the substrate. We then proceeded to investigate the interaction between BamA and the BamCDE sub-complex in atomic detail (**Chapter 6**).

Finally, development of novel ssNMR methods will be critical for the study of large membrane proteins, including structure determination. In **Chapter 7** we describe a "proton-cloud" approach for high resolution proton-detected ssNMR experiments and the collection of distance information, demonstrated on ubiquitin as well as BamA.

In **Chapter 8** the insights obtained on the process of β-barrel assembly are discussed in the context of existing literature and perspectives for future research are presented.

CHAPTER 2

Materials and methods

Expression and purification of BamA

     Construction of plasmids pET11aΔssYaeT encoding *Escherichia coli bamA* without its N-terminal signal sequence and pET11aΔP1-P4yaeT encoding the BamA P5-TM construct was previously described.[65,91] To create the construct BamA P4P5-TM, the relevant DNA segment was amplified from pET11aΔssYaeT by PCR with the primers listed in Table 1 (NdeI and BamHI restriction sites underlined) and cloned into the pET11c expression vector (Novagen).

     E. coli BL21 Star(DE3) cells were transformed with the plasmids and recombinant BamA FL (790 residues; 88 kDa), BamA P4P5-TM (547 residues; 61 kDa) and BamA P5-TM (465 residues; 52 kDa) were produced as cytoplasmic inclusion bodies. For unlabeled protein, cells were grown in LB medium supplemented with 100 μg/mL ampicillin. Isotopically labeled protein was produced in M9 medium supplemented with 100 μg/mL ampicillin and 2 g/L $^{15}NH_4Cl$ and $^{13}C_6$-D-glucose as sole nitrogen and carbon sources, respectively. Specific amino acid labeling was achieved by addition of 200 mg/L of each U-^{13}C,^{15}N labeled amino acid to unlabeled M9 medium 30 minutes prior to induction, and reverse labeling with the same concentration of natural abundance amino acids to U-^{13}C,^{15}N M9 medium. To prepare the ^{1}H-cloud sample, 2 g/L ^{2}H-glucose at natural carbon abundance and 200 mg/L ^{1}H,^{13}C,^{15}N labeled valine, leucine and lysine were

added to deuterated M9 medium. Cells were grown at 37°C until reaching an OD$_{600}$ of 0.6-0.8 and BamA expression was induced by adding 1 mM IPTG to the culture after which incubation was continued for 4 h.

Inclusion bodies were purified as described before.[91] Solubilized inclusion bodies were diluted to 100 μM before refolding as determined by UV absorption at 280 nm using a molar extinction coefficient of 140,040 M^{-1} cm^{-1} for BamA FL, 117,120 M^{-1} cm^{-1} for BamA P4P5-TM or 109,790 M^{-1} cm^{-1} for BamA P5-TM. Aggregates and residual membranes were removed by ultracentrifugation at 100,000 g for 1 h at 4°C. BamA P5-TM with VLKY reverse or VLK forward labeling were solubilized in 100 mM glycine, 20 mM Tris-HCl pH 8.0 (IB buffer) containing 8 M urea and refolded by 20-fold dilution in 50 mM Tris-HCl pH 8.0 , 0.5 % N-dodecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate (sulfobetaine 12, SB12), followed by overnight incubation at room temperature. All other BamA samples were solubilized in IB buffer containing 6 M guanidinium chloride, refolded by 10-fold dilution in 50 mM sodium phosphate pH 7.0 , 1 % N,N-dimethyl-N-dodecylamine-N-oxide (LDAO) and incubated overnight at room temperature. Folding from guanidinium chloride was more reproducible than from urea, but, initially, extensive protein precipitation was observed. This problem was overcome by diluting the protein rapidly into the refolding buffer while stirring. Performing the refolding reaction in Tris-HCl at pH 8.0 or sodium phosphate pH 7.0 did not significantly affect our results. Protein aggregates were removed after refolding by centrifugation at 4,000 g for 20 min at 4°C if necessary.

Reconstitution of BamA for ssNMR

Prior to reconstitution, refolded protein was concentrated 5- to 10-fold using Amicon Ultra centrifugal filter units with a 30-kDa molecular cut-off (Merck Millipore). Lipids (obtained from Avanti Polar Lipids) were dried from chloroform solutions under a stream of nitrogen followed by vacuum drying for at least 1 h. Lipid films were resuspended in 20 mM sodium phosphate at pH 7.0 for molar LPRs of 25:1 and higher, or 20 mM sodium phosphate pH 7.0 and 5 mM MgCl$_2$ for lower LPRs (reconstitution buffer) and incubated for 5 min at 37°C. Protein was added to the lipid suspension, incubated for 30 min at 37°C and dialyzed against reconstitution buffer for one day at room temperature followed by approximately 5 days at 4°C. Biobeads were tested as a faster alternative to dialysis but no proteoliposome formation was observed. Proteoliposomes were collected by centrifugation at 100,000 g for 2 h at 4°C and packed into 3.2 mm or 1.3 mm MAS rotors.

Table 1. *Primers used to create BamA, BamC and BamD constructs.*

BamA P4P5-TM forward	5' GATACA<u>CATATG</u>GATCAGTACAAGCTTTCTGG 3'
BamA P4P5-TM reverse	5' GATAC<u>GGATCC</u>TTACCAGGTTTTACCGATGT 3'
BamA P4P5 T261 forward long	5' <u>GCCGCGCGGCAGCCTG</u>ACCGAAGGCGATCAGTACAAG 3'
BamA P4P5 T261 forward short	5' <u>TG</u>ACCGAAGGCGATCAGTACAAG 3'
BamA P4P5 G424 reverse long	5' <u>CAAGAAGAACCCCTCA</u>ACCGGTGTTGCGCTCTTTTACC 3'
BamA P4P5 G424 reverse short	5' <u>TCA</u>ACCGGTGTTGCGCTCTTTTACC 3'
BamA P5 G344 forward long	5' <u>GCCGCGCGGCAGCCTG</u>GGGTAACCGTTTCTACGTGCG 3'
BamA P5 G344 forward short	5' <u>TG</u>GGTAACCGTTTCTACGTGCG 3'
BamA P5 N422 reverse long	5' <u>CAAGAAGAACCCCTCA</u>GTTGCGCTCTTTTACCTTGTAG 3'
BamA P5 N422 reverse short	5' <u>TCA</u>GTTGCGCTCTTTTACCTTGTAG 3'
BamD full-length forward	5' GCACTC<u>CATATG</u>ACGCGCATGAAATATCTGGTGGCA 3'
BamD full-length reverse	5' ACTC<u>CTCGAG</u>TGTATTGCTGCTGTTTGCGGCGATG 3'
BamC[UN] S26 forward long	5' <u>GCCGCGCGGCAGCCTG</u>AGTTCTGACTCACGCTATAAGCGT CAGGTCAGTGGT 3'
BamC[UN] S26 forward short	5' <u>TG</u>AGTTCTGACTCACGCTATAAGCGTCAGGTCAGTGGT 3'
BamC[UN] A217 reverse long	5' <u>CAAGAAGAACCCCTCA</u>CGCGGCGTCAGTGGCAGATTTAT CCAGACC 3'
BamC[UN] A217 reverse short	5' <u>TCA</u>CGCGGCGTCAGTGGCAGATTTATCCAGACC 3'

Preparation of soluble BamA constructs

The DNA fragments for BamA P5 (G344-N422) and P4P5 (T261-G424) were amplified from pET11aΔssYaeT and cloned into pLICHIS using enzyme-free cloning[97] with the primers listed in Table 1 (LIC overhangs underlined). Proteins were produced in *E. coli* BL21 Rosetta in LB medium for unlabeled samples, in M9 minimal medium supplemented with 0.5 g/L $^{15}NH_4Cl$ for ^{15}N-labeled samples and in addition 2 g/L ^{13}C-glucose for $^{15}N,^{13}C$-labeled samples. Cultures were grown at 37°C

and induced in mid-exponential phase by addition of 1 mM IPTG, and further incubated at 25°C for 16 h. After cell lysis, the protein was purified from the soluble fraction using nickel affinity chromatography and the His-tag was removed using thrombin. The sample was applied to a Superdex 75 column (GE Healthcare) equilibrated in the appropriate buffer for solution NMR (see below). The proteins were concentrated to approximately 0.5 mM for NMR measurements using Amicon centrifugal concentrators with a cut-off of 3.5 kDa. The final concentration was determined from UV absorption at 280 nm using a molar extinction coefficient of 15,930 $M^{-1}cm^{-1}$.

Site directed mutagenesis of BamA

Site-directed mutagenesis to create expression plasmids for BamA E373A and E373K was performed using the primers indicated in Table 2 with pLICHIS P4P5 and pET11c BamA-P4P5-TM as templates. Mutant proteins were produced and purified using the same protocols as for the corresponding wild-type polypeptides.

Table 2. *Primers used for mutagenesis.*

BamA E373A forward	5' GAAATGCGTCAGATGGCTGGTGCATGGCTGGG 3'
BamA E373A reverse	5' CCCAGCCATGCACCAGCCATCTGACGCATTTC 3'
BamA E373K forward	5' GAAATGCGTCAGATGAAAGGTGCATGGCTGGG 3'
BamA E373K reverse	5' CCCAGCCATGCACCTTTCATCTGACGCATTTC 3'

Preparation of soluble BamD

BamD without signal sequence was expressed in BL21(DE3) Star from a pET16b vector carrying BamD$_{21-245}$ with an N-terminal His-tag, which was a kind gift from Frank Beckers (Utrecht University). Unlabeled protein was expressed in LB medium, whereas U-^{2}H,^{15}N,^{13}C labeled protein was expressed in D$_2$O-based M9 medium containing 0.5 g/L ^{15}NH$_4$Cl and 2 g/L ^{2}H,^{13}C-glucose. Cultures were grown at 37°C and induced at mid-exponential phase with 1 mM IPTG, after which expression was carried out at 25°C for approximately 24 h. The protein was purified by nickel affinity chromatography and eluted in 50 mM HEPES pH 7.8, 300 mM NaCl, 250 mM imidazole and 0.75 M urea to prevent protein precipitation. The His-tag was removed using factor Xa (NEB) and the final product purified by gel filtration on a Superdex 75 column (GE Healthcare) equilibrated in the appropriate NMR buffer (see below). The elution fractions of unlabeled BamD were concentrated to 250 µM

as calculated from UV absorption at 280 nm using a molar extinction coefficient of 34,840 $M^{-1} cm^{-1}$, above which the protein precipitated extensively. U-^{2}H,^{13}C,^{15}N BamD was concentrated to 93 µM.

Preparation of the soluble BamC^{UN}D complex

The construct BamC^{UN} comprising the N-terminal extension and the first helix-grip domain, but not the N-terminal signal sequence (residues S26-A217) was amplified from *E. coli* genomic DNA and cloned into the pLICHIS vector using enzyme-free cloning[97] with the primers listed in Table 1 (LIC overhangs underlined). His-tagged BamC^{UN} was expressed in *E. coli* BL21(DE3) Star cells that were grown at 37°C until OD$_{600}$ 0.8, induced with 1 mM IPTG and further incubated at room temperature without agitation overnight. Cells were lysed by sonication and the protein was purified by nickel affinity chromatography in 25 mM HEPES pH 7.5, 250 mM NaCl with 300 mM imidazole in the elution buffer.

Purified BamC^{UN}-His was mixed with BamD$_{21-245}$ from which the His-tag had been cleaved. The complex was purified by gel filtration on a Superdex 75 column (GE Healthcare) equilibrated in 50 mM sodium phosphate pH 6.25, 100 mM NaCl. Peak fractions containing equimolar amounts of both proteins were pooled and concentrated to 95 µM as determined by UV absorption at 280 nm with a combined molar extinction coefficient of 63,260 $M^{-1} cm^{-1}$.

Expression, purification and reconstitution of native BamD

BamD with its N-terminal signal sequence was amplified from *E. coli* genomic DNA with the primers listed in Table 1 and cloned into the pET24b(+) expression vector (Novagen) using NdeI and XhoI restriction sites (underlined in Table 1). Full-length BamD was expressed in BL21(DE3) at 30°C overnight without IPTG induction. Cell envelopes were isolated by sonication of the cells in 50 mM Tris-HCl pH 8.0, 2 mM EDTA after which lysates were cleared from unbroken cells by a centrifugation step at 10,000 g for 10 min. The membrane fraction was pelleted by ultracentrifugation at 100,000 g for 1h and solubilized in 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM imidazole, 40 mM decyl β-D-maltopyranoside (DM). Another ultracentrifugation step at 100,000 g for 1h was performed to remove non-solubilized components. Lipid-anchored BamD was isolated using nickel affinity chromatography in 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 4 mM DM. Purified BamD detergent micelles were added to refolded BamA in LDAO micelles in the appropriate molar ratios calculated from the protein concentrations as determined

by the Pierce BCA assay (Thermo Scientific), after which reconstitution in DLPC at a molar lipid-to-protein ratio of 25:1 (calculated on BamA) was performed according to the same procedure used for isolated BamA.

Expression, purification and reconstitution of native BamCDE

BamCDE with N-terminal signal sequences were produced by co-expression from the plasmids pSK46 and pBamE-His[21] that were a kind gift from Prof. Daniel Kahne (Harvard Medical School). BL21(DE3) Star cells transformed with both plasmids were grown in LB medium at 37°C until an OD_{600} of 0.6, after which expression was induced by addition of 0.1 mM IPTG. Cultures were further incubated at 37°C for 3 h. Cells were lysed by sonication in 50 mM Tris-HCl pH 8.0, 2 mM EDTA and the BamCDE complex was purified from the membrane fraction by nickel affinity purification using the His-tag on BamE according to the same procedure used for lipid-anchored BamD described above. The eluted fraction was further purified by gel filtration on a Highload 16/60 Superdex 200 column (GE Healthcare) equilibrated in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.03% dodecyl β-D-maltopyranoside (DDM). The purified BamCDE complex was mixed with refolded BamA in a 1:1 molar ratio as judged from band intensities on SDS PAGE, after which reconstitution in DLPC at a molar lipid-to-protein ratio of 50:1 (per complex) was performed according to the standard procedure.

Preparation of ubiquitin for ssNMR

A glycerol stock of *E. coli* BL21 Rosetta with a plasmid harboring ubiquitin was a kind gift from Dr. Huib Ovaa (NKI Amsterdam). Ubiquitin was expressed in D_2O based M9 minimal medium containing 2 g/L ^2H-glucose at natural carbon abundance and 200 mg/L of each $^1H,^{13}C,^{15}N$ labeled amino acid to obtain proton clouds. The fully protonated sample was produced in H_2O based M9 medium supplemented with 2 g/L ^{13}C-glucose and 0.5 g/L $^{15}NH_4Cl$. Cells were lysed by sonication and the protein was purified by cation exchange on a MonoS column (GE Healthcare) in 20 mM ammonium acetate pH 5.1 with a salt gradient of 0 to 1 M NaCl, followed by gel filtration on a Superdex 75 column (GE Healthcare) in 20 mM ammonium acetate pH 5.1, 150 mM NaCl. After buffer exchange to H_2O and lyophilization, the protein was precipitated with MPD as described previously[98] to obtain microcrystals. These were pelleted by ultracentrifugation at 100,000 g for 2h at 4°C and packed into 1.3 mm rotors.

ssNMR spectroscopy

ssNMR experiments at moderate MAS frequencies were carried out on a spectrometer operating at 700 MHz [1]H Larmor frequency (16.4 T) equipped with a 3.2 mm ^{1}H,[13]C,[15]N MAS probe (Bruker BioSpin). For dipolar-based experiments, the MAS frequency was set to 11 kHz, 13 kHz or 15 kHz and experiments were done at an effective sample temperature of -2°C unless stated otherwise. Hartman-Hahn cross-polarization (CP) was performed with a contact time of typically 500-800 μs (which was kept constant in case spectra were used for comparison) and a linear ramp of 70-100 %. For ^{13}C,[13]C mixing, the PARIS scheme[99] was used with a mixing time of 30 ms, 40 ms or 150 ms and 8 kHz irradiation on [1]H. Decoupling was performed using SPINAL64[100] with 78 kHz irradiation on [1]H. The NCA experiment was performed using a 600-900 μs [1]H-^{15}N CP step followed by a 4 ms SPECIFIC-CP,[101] typically employing 30 kHz and 17 kHz irradiation on ^{15}N and ^{13}C, respectively. 1D and 2D ^{1}H,[13]C INEPT-HETCOR experiments were performed at an effective sample temperature of 8°C with 10 kHz GARP decoupling.[102]

Proton-detected experiments at fast MAS rates were performed on Bruker spectrometers operating at 700 MHz and 800 MHz [1]H Larmor frequency, equipped with 1.3 mm ^{1}H,[13]C,[15]N MAS probes. MAS frequencies between 40 and 60 kHz were employed. CP was established using standard zero-quantum Hartmann-Hahn conditions (n=1 or 2). In [1]H-cloud samples (ubiquitin and BamA), low-power (LP) decoupling on ^{1}H and ^{13}C channels was carried out with PISSARRO.[103,104] The decoupling amplitude $\nu_{LP\text{-}decoupling}$ was set to $\nu_{LP-decoupling}^{1H/13C} \approx 0.25\nu_{rot}$; pulse length τ_p was shortly optimized around 360° pulses: $\tau_p \approx \frac{15}{16\nu_{LP-decoupling}}$. In fully protonated samples (ubiquitin and BamA), high-power PISSARRO decoupling (160 kHz decoupling amplitude on ^{1}H) was applied during evolution times and LP-PISSARRO during direct acquisition times. Solvent suppression was carried out using the MISSISSIPI[105] scheme $[\tau_{p(x)}\ \tau_{p(-x)}\ \tau_{p(y)}\ \tau_{p(-y)}]_L$ (with τ_p = 10 ms and L = 2) over a total of 80 ms.

Data were processed with Bruker TopSpin 3.0 and analyzed using Sparky.[106]

Solution NMR spectroscopy

Solution NMR experiments on BamA P4P5 were performed using AVANCE III Bruker spectrometers operating at 600 MHz [1]H Larmor frequency equipped with regular or cryogenic TXI probes. Resonance assignments were obtained using standard 3D triple- (HNCA, HNcoCA, HNCACB, CBCAcoNH, HNCO, HAHBcoNH) and double-resonance (hCCH-DIPSY, HcCH-DIPSY) experiments. Arginine side chain Hε-

Nε resonances were assigned using 2D HccneHE and hCcneHE (with mixing times of 12 ms and 24 ms, respectively)[107] in combination with a 3D ^{15}N-NOESY-HSQC (mixing time 100 ms). ^{1}H-^{15}N HSQC spectra of P4P5 mutants (E373A, E373K) were assigned by comparison with WT spectra and where possible confirmed using 3D HNCA and CBCAcoNH spectra for E373K. Secondary chemical shifts (H^N, N, HA, CA, CB, C') were computed using TALOS+[108] and used for secondary structure determination. ^{15}N relaxation dispersion experiments were performed as described by Tollinger et al.,[109] using a constant time delay T_{CP} of 30 ms and 4.5 kHz ^{15}N CPMG pulses. The spectrum was recorded as a pseudo 3D with the CPMG frequency ν_{CPMG} as third dimension, using 19 increments between 0 and 960 Hz, including the reference and twice the 400 Hz experiment. R_2^{eff} was calculated from the ratio of intensities, where I_0 represents the intensity in the reference experiment:[110]

$$R_2^{eff}(\nu_{CPMG}) = \frac{-1}{T_{CP}} ln \frac{I(\nu_{CPMG})}{I_0}$$

The titration of U-^{13}C,^{15}N BamA P4P5 with unlabeled BamD was performed at 750 MHz ^{1}H Larmor frequency with stock concentrations of 50 µM P4P5 and 250 µM BamD, both in 50 mM sodium phosphate pH 7.0, 100 mM NaCl. The reverse titration was performed at 600 MHz ^{1}H Larmor frequency using 93 µM U-^{2}H,^{13}C,^{15}N BamD and 720 µM unlabeled P4P5 in 50 mM sodium phosphate pH 7.0, 50 mM NaCl. The titration of U-^{13}C,^{15}N P4P5 with the unlabeled BamC^{UN}D complex was performed at 600 MHz ^{1}H Larmor frequency with 50 µM P4P5 and 95 µM BamC^{UN}D complex both in 50 mM sodium phosphate pH 6.25, 100 mM NaCl.

Solution NMR data were processed using NMRPipe[111] and analyzed with Sparky.[106]

Electron microscopy

Samples at approximate protein concentrations of 1 mg/mL were adsorbed to an amorphous carbon film atop of 200 mesh copper grids (JEOL, Netherlands) for 2 minutes before rinsing three times with water and staining with 2 % w/v uranyl acetate (Merck, Netherlands). Grids were imaged on a Tecnai 10 or Tecnai 12 microscope (FEI Company, Netherlands) operating at 100 and 120 kV, respectively, and images recorded at nominal magnifications of between 38,000 - 97,000 x on a SIS Megaview II CCD detector (Olympus, Netherlands).

Electron microscopy was performed at Electron Microscopy Utrecht.

Molecular Dynamics simulations

Molecular Dynamics simulations were carried out using the Groningen Machine for Chemical Simulations (GROMACS) simulations package, version 4.5.3,[112] with the Groningen Molecular Simulation computer program package (GROMOS53a6) force field[113] and the Berger lipid parameters.[114] The simulation system was represented by the homology model of *E. coli* BamA[41] (residues Asp264 to Trp810), embedded in a DMPC bilayer in an aqueous solution of NaCl. Potassium and chloride ions were added to electrically neutralize the system and to mimic a 150 mM NaCl solution. The final system consisted of 212,222 atoms, comprising BamA, 491 DMPC lipids, 174 sodium ions, 151 chloride ions and 61,217 water molecules. Ions and water molecules, as well as lipids and the protein were each jointly coupled to the thermostat at 303 K. After initial equilibration in a NVT ensemble, the system was simulated in a NPT ensemble using semi-isotropic pressure coupling for 50 ns with gradually reduced force constants and then further evolved for 300 ns without restraints.

Conservation analysis

Full-length BamA from *E. coli* (Uniprot P0A940) was subjected to a BLAST search using the Uniref 90 database to prevent incorporation of redundant sequences. The results were filtered for proteobacteria, which generally have five POTRA domains. Sequences were aligned using Clustal Ω and Consurf[115] was used to analyze the residue conservation and plot it on the BamA structure.

Sequence conservation of BamD was performed with Consurf[115] using the crystal structure (PDB 3TGO) as input, the Uniref 90 database to collect homologues and Clustal W for alignment.

Docking

Docking was performed using HADDOCK 2.1[116] with active residues on BamA derived from the NMR data and passive residues on BamD defined from sequence conservation (see Table 3). From the solution NMR data, P5 residues were selected that experienced a CSP of more than µ+2σ in the titration with BamD. From the ssNMR data, residues were selected that experienced line broadening in the experiment with co-reconstituted BamD and/or CSPs in the BamCDE complex. For docking with the BamA P4P5-TM construct, residues from pL2 and pL3 were included (Table 3). Surface accessibility of residues identified in solution and ssNMR data was analyzed with NACCESS[117] and residues with less than 15% main chain or

side chain accessibility were discarded. Residues E373 that is known to be involved in the BamD interaction from mutagenesis studies[69,118] was included in all docking runs as active residue. Passive residues for BamA constructs were automatically defined in HADDOCK.

For BamD, passive residues were chosen based on conservation. Residues A175, E177, V181, A182, Y185, R188, A190, A193, V194, N196 and R197 formed a highly conserved patch on the C-terminal side of the molecule that was previously implied in the stable interaction with BamA. Therefore, the region encompassing these residues, A175-R197, was defined as passive (Table 3).

Input structures were taken from PDB 3Q6B for P4P5, 3TGO for BamD and the *E. coli* homology model of BamA[41] truncated to P4P5-TM. Random removal of restraints was set to 50% when a sufficient number of active residues was available, and otherwise reduced to 33% or switched off (Table 3). Other HADDOCK parameters were kept at the default values.

Table 3. HADDOCK parameters.

	BamA (active)	BamD (passive)	Removal of restraints
P4P5 solution NMR	Y348, I352, A363, R366, E368, E373, A375, F394	A175-R197	50%
P4P5 ssNMR	I352, V364, R366, E373, A375		33%
			off
P4P5-TM ssNMR	I352, R366, E373, A375, F478-G482, P518		50%

CHAPTER 3

Solid-state NMR studies of full-length BamA in lipid bilayers suggest limited overall POTRA mobility

Tessa Sinnige, Markus Weingarth, Marie Renault, Lindsay Baker,
Jan Tommassen, Marc Baldus

Abstract

The outer membrane protein BamA is the key player in β-barrel assembly in Gram-negative bacteria. Despite the availability of high-resolution crystal structures, the dynamic behavior of the transmembrane domain and the large periplasmic extension consisting of five POTRA domains remains unclear. We demonstrate reconstitution of full-length BamA in proteoliposomes at low lipid-to-protein ratio, leading to high sensitivity and resolution in solid-state NMR (ssNMR) experiments. We detect POTRA domains in ssNMR experiments probing rigid protein segments in our preparations. These results suggest that the periplasmic region of BamA is firmly attached to the β-barrel and does not experience fast global motion around the angle between POTRA 2 and 3. We show that this behavior holds at lower protein concentrations and elevated temperatures. Chemical-shift variations observed after reconstitution in lipids with different chain lengths and saturation levels are compatible with conformational plasticity of BamA's transmembrane domain. Electron microscopy of the ssNMR samples shows that BamA can cause local disruptions of the lipid bilayer in proteoliposomes. The observed interplay between protein-protein and protein-lipid interactions may be critical for BamA-mediated insertion of substrates into the outer membrane.

Introduction

The outer membrane protein (OMP) BamA is the central component of the β-barrel assembly machinery that folds and inserts OMPs in Gram-negative bacteria.[32,33] Not only is BamA highly conserved in these organisms, but homologues in chloroplasts and mitochondria have also been identified.[31,119,120] BamA consists of a transmembrane β-barrel domain with 16 strands[41] and a periplasmic extension of five POlypeptide-TRansport Associated (POTRA) domains.[35]

Experimental evidence suggests that the POTRA domains serve as docking sites for the BamB/C/D/E lipoproteins and have a chaperone-like function for substrates.[36] In *Escherichia coli*, a deletion analysis of the POTRA domains has shown that POTRA 2-5 are required for BamB association and POTRA 5 is required for the BamC/D/E interaction.[36] POTRA 1, on the other hand, has been shown to control assembly of BamA itself and to interact with SurA, a major periplasmic chaperone.[121] Moreover, NMR titration experiments suggested the ability of POTRA domains to bind unfolded OMPs,[39] possibly by means of β-augmentation.[36] Apart from mediating protein-protein interactions, the POTRA domains are likely to experience larger domain reorientations during the process of β-barrel assembly to drive folding and insertion of the substrate. Bending of the periplasmic region of BamA could allow for the formation of β-hairpins[37,38] which are thought to be an intermediate in β-barrel folding.[70]

Several crystallographic structures of the POTRA domains have been solved, but their overall conformation and flexibility are still under debate.[36-39] Recently, crystal structures of full-length BamA from *Neisseria gonorrhoeae* and a truncated construct from *Haemophilus ducreyi* have become available, showing two different orientations of the POTRA domains with respect to the β-barrel.[41] In the first structure, POTRA 5 closes the lumen of the β-barrel, whereas in the latter, the POTRA domains are turned away from the transmembrane domain. Taken together, these findings were interpreted as a conformational switch that might allow the substrate to enter the β-barrel.[41]

In the following we investigated the dynamics of *E. coli* BamA and its POTRA domains in membranes by Magic Angle Spinning (MAS) solid-state NMR (ssNMR). In addition, we studied the interplay between the protein and its lipid environment by ssNMR and electron microscopy (EM). MAS ssNMR is a powerful technique to study membrane proteins in lipid bilayers (see e.g. Ref. 74-76,79,122-124) and hence provides insight on functionality in the native(-like) environment (see e.g. Ref. 71). Unless dedicated signal enhancements such as Dynamic Nuclear Polarization (DNP) are employed (see e.g. Ref. 125,126) signal-to-noise considerations in multidimensional ssNMR studies require high protein concentrations.[127,128] Hence,

care must be taken to reconstitute the purified protein in lipid bilayers at low lipid-to-protein ratio (LPR) such that the correct fold and stability are maintained.

Previously, we showed that the POTRA domains and the β-barrel of BamA could be separated in motion-filtered ssNMR studies on precipitated protein and a truncated construct in liposomes.[91] In both cases, our NMR results suggested that POTRA domains exhibit motional degrees of freedom. In the current work, we extended our studies by investigating the full-length, 88-kDa protein after reconstitution in liposomes by high-resolution ssNMR. Conditions for refolding and reconstitution were found that yield well-resolved 2D ssNMR spectra, allowing us to probe the fold and mobility of the complete periplasmic part of BamA, i.e. POTRA domains 1-5, as well as the transmembrane domain embedded in lipid bilayers. In our analysis we utilized published NMR resonance assignments[39] for POTRA domains 1 and 2 as well as chemical-shift prediction routines that have become a general structure elucidation tool in NMR (see e.g. Ref. 129-131). In addition, we studied the influence of the lipid type and the lipid-to-protein ratio by ssNMR as well as EM.

Results and discussion

Full-length BamA stably incorporates into liposomes

Full-length (FL) BamA (88 kDa) was expressed in *E. coli* without its N-terminal signal sequence (Fig. 1a), targeting it to intracellular inclusion bodies. With this procedure, high yields of recombinant protein can be obtained. After solubilization of inclusion bodies and protein purification in the presence of a chaotropic agent, however, the protein needs to be refolded in detergent micelles, which requires careful optimization. Various detergents and dilution methods were tested, yielding an optimum of ~85% FL BamA refolding in 1% N-dodecyl-N,N-dimethylamine-N-oxide (LDAO) as judged from heat-modifiability on semi-native PAGE, on which folded BamA migrates faster than heat-denatured protein.[65] Heat-modifiability (i.e. the different electrophoretic mobility of the properly folded and heat-denatured forms of OMPs) is a well-known property of OMPs[132,133] and established as a method to monitor the proper refolding of an OMP in vitro after its production in inclusion bodies (see e.g. Ref 134).

From detergent micelles, BamA was reconstituted into liposomes by dialysis. Zwitterionic phosphatidylcholine (PC) lipids with different chain lengths and saturation levels were tested, namely dilauroyl-PC (C12:0, DLPC), dimyristoyl-PC (C14:0, DMPC) and dioleyl-PC (C18:1, DOPC), as well as an *E. coli* polar lipid extract

composed of phosphatidylethanolamine, phosphatidylglycerol, cardiolipins and other lipids. Different LPR's were screened to search for the lowest one supporting correct protein folding and membrane insertion and yielding the highest sensitivity in ssNMR experiments.

E. coli polar lipids are often used for functional reconstitution of membrane proteins because they mimic the lipid environment of the bacterial inner membrane.[135] Notably, the outer membrane however has a very different lipid composition with the outer leaflet consisting of lipopolysaccharides (LPS) (*vide infra*). In the case of BamA, the use of *E. coli* polar lipids led to the detection of protein in the supernatant after harvesting the proteoliposomes by centrifugation for all LPR's tested. We concluded that BamA was not fully incorporated into the lipid bilayer under the conditions we used for reconstitution. By contrast, we did succeed in fully reconstituting BamA in proteoliposomes of any of the pure phosphatidylcholine lipids (DLPC, DMPC and DOPC) at a molar LPR as low as 25:1, corresponding to around 1:5 in weight. Note that such conditions are comparable to the bacterial outer membrane itself that is also densely packed with protein with a ratio of phospholipids:LPS:protein estimated to be around 1:1:5.[136] Yields of folded protein in the lipid vesicles could be estimated from the relative band intensities of folded and unfolded protein on semi-native SDS-PAGE and were found to be similar for all types of PC used, with an average of 76 ± 1.6 % folded (Fig. 1b).

Additional experiments were performed to confirm that the BamA β-barrel had correctly inserted into the lipid bilayer. No protein was observed in the supernatant on SDS-PAGE when washing the proteoliposomes with 400 mM NaCl or 1 M urea (data not shown), indicating that the protein was not unspecifically bound to the liposome surface. In addition, limited proteolysis with trypsin yielded a protein fragment of about 50 kDa on SDS-PAGE corresponding to the isolated transmembrane domain of BamA (Fig. 1c). These results show that BamA FL is homogeneously and properly embedded in the lipid bilayer. Importantly, all BamA molecules showed accessible for trypsin digestion meaning that their POTRA domains are solvent-exposed and not involved in aggregation. In summary, correct folding and reconstitution of BamA into lipid bilayers were confirmed by heat-modifiability on SDS PAGE, urea- and salt-extractability and protease resistance.

Next, we characterized the proteoliposomes by EM (Fig. 1d). At a molar LPR of 25:1, all PC proteoliposomes formed mostly round vesicles with a diameter of a few hundred nanometers. However, they show notches at the surface. This effect is particularly apparent for DLPC - which is the thinnest bilayer - at LPR 25:1, for which the liposomes appear rather disrupted. At a higher molar LPR of 150:1, the DLPC proteoliposomes look more intact, although small notches are still present.
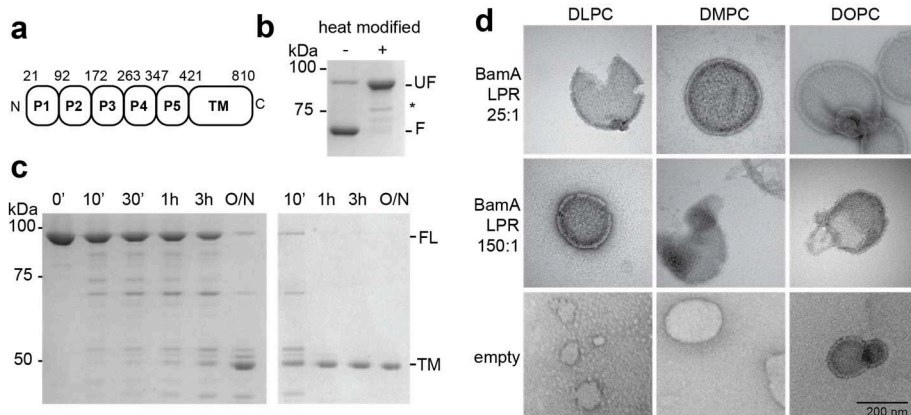
Figure 1. BamA full-length (FL) is well folded and stably inserted in liposomes at low LPR. a) Domain organization of mature BamA after processing of the N-terminal signal sequence. P1-5: POTRA domains 1-5; TM: transmembrane domain. b) Heat-modifiability of BamA FL analyzed on semi-native SDS-PAGE to assess the fraction of folded protein. As an example, the analysis of BamA FL reconstituted in DOPC vesicles at a molar LPR of 25:1 is shown. Comparison between native (left) and heat-denatured (right) proteoliposomes reveals a characteristic shift of the electrophoretic mobility of folded BamA FL (F) and unfolded BamA FL (UF) from 70 to 88 kDa, respectively. The amount of well-folded BamA FL is estimated by measuring the relative intensity of the F band relative to the UF band and represents about 76% of the overall protein content. The asterisk indicates slight degradation of the sample after one week of dialysis. c) Limited proteolysis with trypsin shows that the BamA β-barrel is properly inserted into the lipid bilayer and all protein molecules are accessible. Time-points after addition of 1:100 w/w trypsin:BamA to BamA/DOPC proteoliposomes of LPR 25:1 were analyzed on SDS-PAGE. Incubation at 4°C (left panel) causes the full-length protein (FL) to be gradually degraded; incubation at room temperature (right panel) leaves a band migrating at around 50 kDa (TM) which corresponds to the β-barrel that has a MW of 44 kDa. The samples were heat-denatured prior to loading on the gel. d) Representative electron micrographs of BamA proteoliposomes in DLPC, DMPC and DOPC at a molar LPR of 25:1 (top row) and 150:1 (middle row), as well as empty liposomes (bottom row).

Electron micrographs of DMPC and DOPC proteoliposomes at a molar LPR of 150:1 revealed mainly aggregated structures and distorted vesicles. To the best of our knowledge, the notches that we discovered in the BamA samples have not been observed in other β-barrel containing proteoliposomes. Control liposomes prepared by solubilization of the lipids with the BamA refolding buffer followed by dialysis yielded vesicles similar to the BamA proteoliposomes for DMPC and DOPC, although on average somewhat smaller. DLPC only formed very small liposomes that did not yield a pellet after centrifugation. However, no sharp notches were observed in the empty liposomes. The disturbance of the lipid bilayer by BamA might be functionally relevant for substrate insertion. Indeed, the crystal structure of BamA showed that

the hydrophobic thickness of the transmembrane domain is asymmetric, which may induce large distortions of the lipid bilayer in proximity to BamA, as demonstrated by molecular dynamics simulations.[41] Extending this phenomenon to a larger scale, clusters of BamA molecules might be responsible for the deformation of the lipid bilayer in our proteoliposomes.

POTRA domains 1 and 2 exist as rigid components of BamA in lipid bilayers

To characterize the overall fold and dynamics of BamA reconstituted in proteoliposomes, ssNMR measurements were carried out on uniformly (U-) $^{13}C,^{15}N$ labeled BamA in DLPC lipid bilayers at a molar LPR of 25:1 (Fig. 2). The 2D $^{13}C,^{13}C$ dipolar-based correlation spectrum (Fig. 2a) is indicative of a well-folded protein with ^{13}C linewidths of 0.7-1.0 ppm. As we have shown in earlier work, ssNMR schemes that invoke dipolar polarization transfer steps primarily report on protein domains that exhibit limited dynamics on the timescale of the ssNMR experiment, while through-bond experiments report on protein segments that exhibit fast motion (see e.g. Ref. 90). In the following, we concentrated on studying overall motion of specific POTRA domains which should reduce the ssNMR signal intensities of all POTRA residues seen in dipolar-based experiments and enhance signals detected in through-bond experiments. For such an analysis, we concentrated on well-resolved resonances in our ssNMR data that must stem from POTRAs 1-3. To identify such correlations we compared ssNMR data obtained on FL BamA to spectra obtained on a truncated construct of BamA lacking the first three POTRA domains at its N-terminus (BamA P4P5-TM, 547 residues). This comparison revealed a set of well-resolved correlations, for example in the spectral region typical for intra-residue correlations of serine, threonine and isoleucine (Fig. 2b) that matched with previous NMR assignments obtained on isolated POTRA domains 1 and 2 (BMRB 15247).[39] These correlations hence most likely stem from POTRA domains 1 and 2. For 24 out of the 153 residues that these domains comprise, resolved signals were found corresponding to the assignments (Fig. 2c, indicated in black) and these were subsequently used as spectral probes for *overall* domain motion.

The presence of well-folded POTRA domains in our dipolar-based experiments speaks in favor of restricted molecular motions in DLPC proteoliposomes used in Fig. 2. In previous work on folded precipitates of FL BamA, we have shown that POTRA domains can undergo faster molecular motions on nano- up to microsecond timescales,[91] pointing to a potential stabilizing effect of the lipid bilayer.
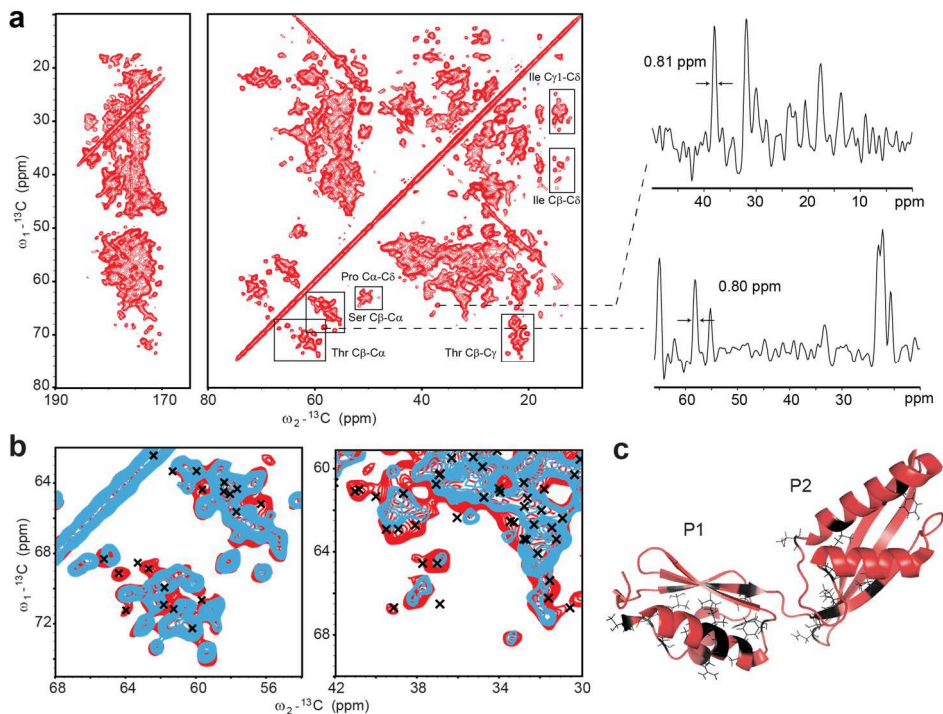
Figure 2. ssNMR signals from periplasmic POTRA domains 1 and 2 are observed in experiments probing rigid protein segments. a) Left, 2D ^{13}C,^{13}C correlation spectrum obtained on U-^{13}C,^{15}N labeled BamA FL in DLPC proteoliposomes at a molar LPR of 25:1 using PARIS recoupling[99] with a mixing time of 30 ms. Right, lines extracted from the 2D ^{13}C,^{13}C correlation spectrum indicate ^{13}C linewidths of about 0.8 ppm. Spectra were acquired at an effective sample temperature of -2°C. b) Overlay between selected regions from the 2D ^{13}C-^{13}C correlation spectrum on U-^{13}C,^{15}N BamA FL as shown in a) (red) and U-^{13}C,^{15}N BamA P4P5-TM (blue) in DLPC bilayers at a molar LPR of 25:1. The CP contact times were set to 600 μs and 500 μs, respectively. Solution NMR assignments of POTRA 1 and 2 (BMRB 15247) are superimposed (black crosses). c) NMR structure of BamA POTRA 1-2 domains (PDB 2V9H). The 24 residues that appear as resolved cross-peaks in the 2D ^{13}C,^{13}C spectrum of BamA FL are shown in black.

Conformational flexibility of the POTRA domains has been investigated extensively before.[36-39] POTRA 1 and 2 seem to have a rather defined juxtaposition in crystal structures,[36,37] whereas the solution NMR structure consists of a variable ensemble,[39] the lowest energy structure of which has a markedly different domain-domain angle compared to the crystal structures. The latter outcome, however, seemed to result from the lack of sufficient NOE restraints since incorporation of PELDOR restraints in the structure calculation led to a more defined conformation.[137] Our results suggest that both domains are rigid with respect to each other.

The angle between POTRA 2 and 3 has been under more debate as the two crystal structures of POTRA 1-4 strikingly differ in this aspect, resulting in either a "fish hook"[36] or extended conformation.[37] In the structure of full-length *N. gonorrhoeae* BamA, the POTRA domains however adopt a conformation that is in between these two folds.[41] The fact that we observe POTRA domains 1 and 2 as rigid protein segments implies that either the POTRA domains largely populate one preferred conformation or the exchange between the different conformations is much slower than the timescales probed in our ssNMR experiments. At this point, it would be difficult to discern differences in the angle between POTRA domains 2 and 3, as this would only cause residues at their interface to exhibit chemical-shift variations. To resolve such changes, three- or higher-dimensional ssNMR, possibly in combination with advanced isotope labeling (see below), would be required to reduce spectral crowding and to obtain residue-specific ssNMR chemical shift assignments for POTRA 3.

BamA transmembrane domain and POTRA 4 and 5 can be studied by a specific labeling scheme

To study the BamA transmembrane domain and the adjacent POTRA domains with optimal sensitivity, we revised our reconstitution protocol to yield a sample of BamA P4P5-TM in DLPC at a molar LPR of 10:1. At this low LPR, we expect the BamA β-barrel domains to be in close contact with one another, which is unfavorable given its negative charge at neutral pH (pI = 4.5). Indeed, 5 mM $MgCl_2$ was necessary in the reconstitution buffer to provide counter ions. Despite the slight modification of the sample preparation protocol, the 2D $^{13}C,^{13}C$ correlation spectrum of this sample was virtually identical to the one obtained on proteoliposomes at a molar LPR of 25:1 described above, yet the sensitivity of the 10:1 sample was significantly better (data not shown). Notably, we confirmed that the protein was properly incorporated in the lipid bilayer by a two-dimensional T_2 filtered H(H)C ssNMR experiment in which we monitored lipid-protein correlations (see e.g. Ref. 138). The results were consistent with protein residues being in close proximity to lipid chains (Fig. 3).
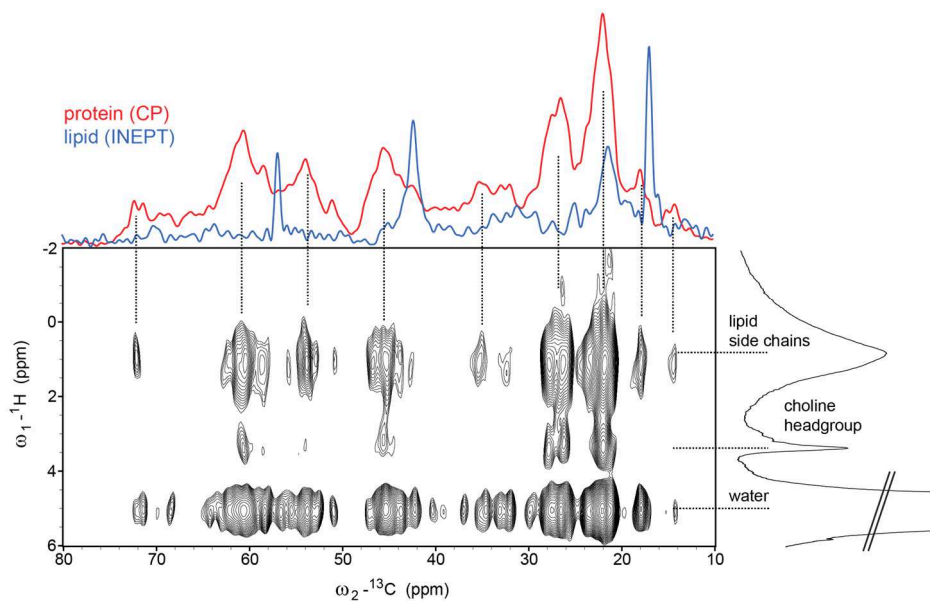
Figure 3. BamA contacts the lipid bilayer in our ssNMR preparations. Shown is a T_2-edited H(H)C experiment with 5 ms ^{1}H,^{1}H mixing time recorded on BamA P4P5-TM with amino acids GSAVLTI ^{15}N,^{13}C labeled, reconstituted in DLPC at a molar LPR of 10:1, revealing cross-peaks between lipid and protein. The ^{1}H frequencies of the lipid acyl chains and choline head groups (1D ^{1}H spectrum, right) correlate with the ^{13}C frequency of protein moieties that appear in a 1D CP (top, red), rather than lipid natural abundance ^{13}C signals that are detected in a 1D INEPT experiment (top, blue).

In spite of the truncation, the BamA P4P5-TM construct still consists of 547 residues. A specific labeling scheme was designed to alleviate spectral crowding. Depending on the protein sequence, different labeling strategies can be pursued. Reverse labeling approaches have been successfully used to assign relatively large α-helical membrane proteins in ssNMR,[139-141] since they are characterized by a substantial fraction of hydrophobic residues that cause overlap in the spectra. Outer membrane β-barrels, however, contain a much more homogeneous distribution of amino acids, including a large fraction of polar and charged residues.[142] Removal of hydrophobic residues by reverse labeling is therefore not highly effective, and removal of charged and polar residues is often not straight-forward due to the metabolic pathways of amino acid synthesis. Glutamate and glutamine for example cannot be reverse labeled without scrambling because they are subject to transamination. Apart from reverse labeling, labeling with precursors such as 1,3- and 2-^{13}C-glycerol has been used to reduce spectral complexity and aid resonance assignment.[143] However, the resulting labeling pattern is not straightforward for all

amino acid types and can be difficult to interpret in the absence of a well-characterized uniformly labeled sample. 1- and 2-^{13}C-glucose labeling has been successfully applied to fibrillar systems[144] but leads to labeling of only 50% of the molecules at a certain position. We thus decided to specifically label amino acids which are easily recognized by their chemical shift pattern and make up a large part of the BamA transmembrane domain. Based on these requirements, a FANDAS analysis[145] led us to label Gly, Ser, Ala, Val, Leu, Thr and Ile with the resulting ^{13}C,[13]C spectrum shown in Figure 4a confirming our expectations.

FANDAS predictions based on the crystal structure of POTRA 4-5 (Fig. 4a,b; blue dots) and the *E. coli* homology model of the transmembrane domain[41] (Fig. 4a,b; red dots) match well with the observed signals within the confidence limits of ShiftX2[146] chemical shift predictions investigated earlier.[130] The main deviations are the serine signals that appear low in intensity in the 2D ^{13}C,[13]C spectrum due to scrambling (Fig. 4a, dashed arrow). Interestingly, a set of well-resolved peaks in the 2D ^{13}C,[13]C experiment is not supported by predictions (Fig. 4a, solid arrows). Presumably these discrepancies arise from differences between the homology model of the transmembrane domain and the structure of *E. coli* BamA reconstituted in proteoliposomes. These differences can be especially pronounced for loops that are not highly conserved in sequence and length, yet modeled according to the structure of the loops in the template structure.[41]

In addition to the most N-terminal POTRA domains (Fig. 2), the results of Figure 4 suggest that the transmembrane domain and adjacent POTRA domains are rigid in our preparations. Indeed, when using 2D INEPT-HETCOR experiments on U-^{13}C,[15]N BamA P4P5-TM invoking only scalar-based magnetization transfer steps to reveal mobile components,[27] we observed lipid signals and flexible protein side chains, but no backbone signals (Fig. 4c), which provides additional evidence that there is no flexible hinge between POTRA 5 and the β-barrel. Notably, this protein region is likely involved in targeting the substrate to enter either the membrane or the lumen of the β-barrel.[41] Possibly, POTRA 5 and the β-barrel share a stabilizing interface in our liposomal preparations, rendering the periplasmic BamA extension rigid. Interestingly, one of the recently published crystallographic conformations showed closure of the β-barrel by the POTRA 5 domain, although the interface residues were shown to be non-essential for function by mutagenesis.[41] An alternative possibility was given by the crystal structure of *E. coli* TamA, which is also a member of the Omp85 protein family, in which the POTRA domain adjacent to the transmembrane domain is stabilized by periplasmic loops without full occlusion of the β-barrel.[147]
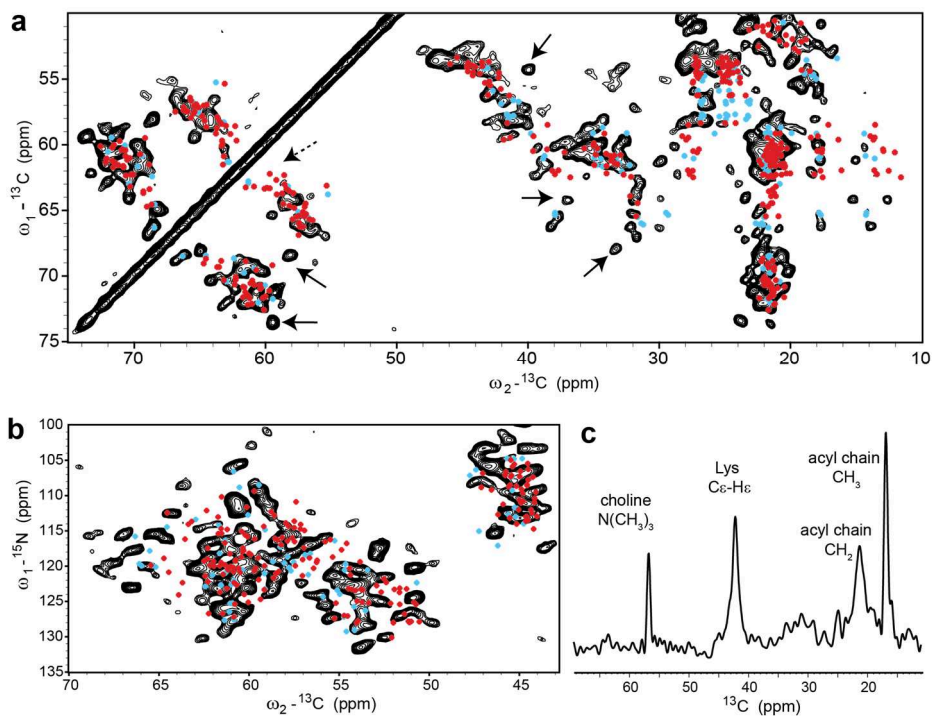
Figure 4. BamA transmembrane domain and POTRA 4-5 can be probed in a specifically labeled ssNMR sample. a) 2D ^{13}C,^{13}C spectrum with 30 ms PARIS mixing recorded on BamA P4P5-TM in DLPC at LPR 10:1 with amino acids GSAVLTI ^{15}N,^{13}C labeled. The CP contact time was 750 µs. Dots show FANDAS[145] predictions for this BamA construct and labeling scheme. Chemical shift predictions were generated with ShiftX2[146] based on the crystal structure of POTRA 4-5 (PDB 3Q6B) (blue) and the *E. coli* homology model[41] of the transmembrane domain (red). Dashed arrow points at lacking serine cross-peaks; solid arrows point at outstanding peaks that are not predicted. b) NCA spectrum of GSAVLTI ^{15}N,^{13}C BamA P4P5-TM in DLPC at LPR 10:1 with FANDAS predictions in the same color coding as in a). Spectra in a) and b) were acquired at an effective sample temperature of -2°C. c) 1D ^{1}H,^{13}C INEPT-HETCOR on U-^{15}N,^{13}C BamA P4P5-TM in DLPC at LPR 10:1, measured at an effective sample temperature of 8°C. Characteristic ^{13}C signals from natural abundance DLPC (based on published assignments for phosphatidylcholine lipids[148]) and from the lysine side chain (Cε-Hε) of labeled BamA are indicated.

BamA structure and dynamics are largely preserved under differential experimental conditions

The observed overall rigidity of the POTRA domains in our proteoliposome preparations might be influenced by protein concentration and temperature. At low LPR, stabilizing protein-protein interactions could be responsible for a decrease in protein motion. When comparing dipolar-based 2D ^{13}C,^{13}C spectra of BamA FL in DLPC at a molar LPR of 25:1 and of 150:1, we observed that the overall intra-residue

correlation pattern was similar for both preparations, even though some signals are lacking at higher LPR due to poor signal-to-noise (Fig. 5a). An overlay with the truncated construct BamA P4P5-TM and solution NMR assignments for POTRA 1 and 2 (BMRB 15247)[39] confirmed the presence of the N-terminal POTRA domains as rigid components also at a molar LPR of 150:1 (Fig. 5a, arrows). Moreover, in the 2D ^{1}H,^{13}C INEPT-HETCOR spectrum on this sample, only lipid signals and some protein side chains were observed (Fig. 5b) as shown above for a far more concentrated sample (Fig. 4c), again suggesting that the POTRA domains do not experience fast overall motion even in more diluted samples. We made similar observations when comparing a 2D ^{13}C,^{13}C spectrum recorded at 20°C to the data recorded at -2°C. Also in this case correlation patterns largely overlap and match with FANDAS POTRA predictions including solution NMR assignments of POTRA 1 and 2 (BMRB 15247)[39] (Fig. 6).
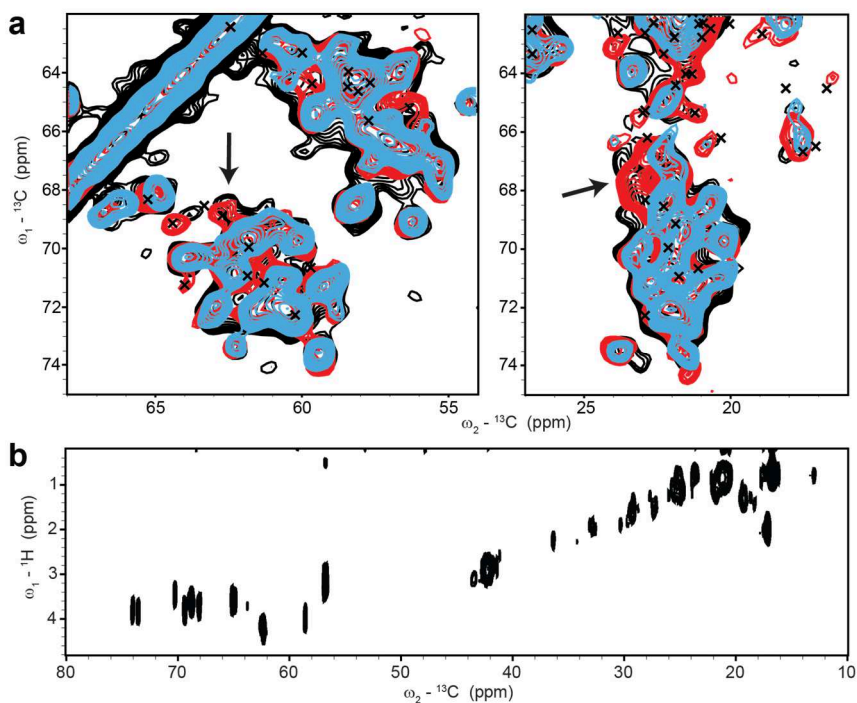
Figure 5. BamA POTRA domains remain rigid at lower protein concentration in the lipid bilayer. a) Overlay of 2D ^{13}C,^{13}C spectra of BamA FL in DLPC bilayers at a molar LPR of 150:1 (black) and 25:1 (red), as well as BamA P4P5-TM in DLPC at a molar LPR of 25:1 (blue). Solution NMR chemical shifts of POTRA 1-2 (BMRB 15247)[39] are shown as crosses. Arrows point at signal intensity from POTRA domains conserved in both samples of BamA FL. All experiments were performed at an effective sample temperature of -2°C, employing 30 ms PARIS mixing. b) 2D ^{1}H,^{13}C INEPT-HETCOR of BamA FL in DLPC at LPR 150:1, recorded at an effective sample temperature of 8°C.
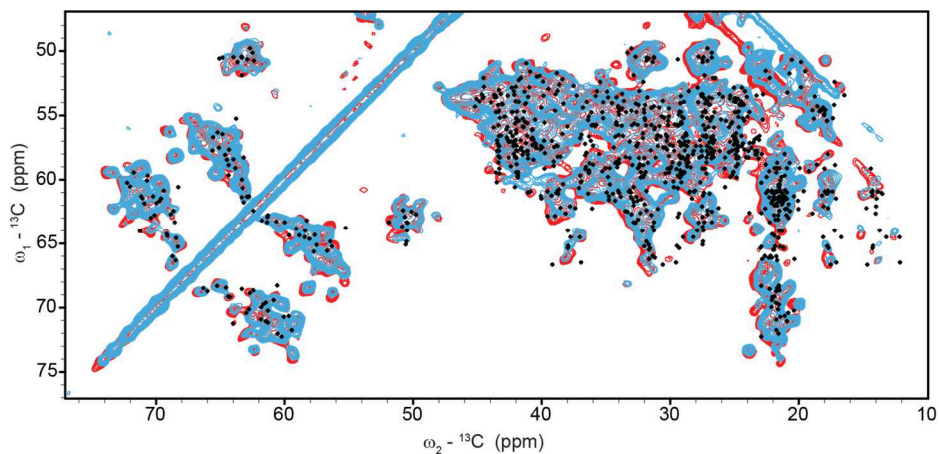
Figure 6. POTRA domains remain rigid at higher temperature. Shown is an overlay of 2D $^{13}C,^{13}C$ spectra with 30 ms PARIS mixing of BamA FL in DLPC at LPR 25:1 recorded at -2°C (red) and at 20°C (blue). Crosses are FANDAS predictions for POTRA 1-5 based on solution NMR chemical shift assignments of POTRA 1-2 (BMRB 15247)[39] and ShiftX2[146] predictions on the crystal structures (PDB 2QCZ) and (PDB 3Q6B) of the other POTRA domains.

Lastly, we analyzed FL BamA reconstituted in lipid bilayers of different chain lengths and saturation levels to investigate whether our findings extend to different membrane environments. Interestingly, the 2D $^{13}C,^{13}C$ spectra in DLPC, DMPC and DOPC lipid bilayers are highly similar (Fig. 7a). Only minor differences are apparent for signals that most likely stem from the transmembrane domain, since these correlations cannot be explained by FANDAS predictions of the POTRA domains (Fig. 7b).

Threonine and serine correlations that show differences (Fig. 7b, indicated by arrows) appear to be in β-strand conformation based on average chemical shift analysis.[149] These observations point towards a certain degree of conformational plasticity within the BamA β-barrel that is induced by the hydrophobic thickness of the membrane in our preparations. In liquid crystalline phase, the hydrophobic thicknesses of DLPC, DMPC and DOPC are 21, 23 and 27 Å, respectively (Fig. 8a).[150,151] At an effective measurement temperature of -2°C, which is favorable for sensitivity in dipolar-based experiments, DMPC is in the gel phase, whereas DOPC is in the liquid crystalline phase and DLPC is around its phase transition (Fig. 8b). Notably, at this temperature the sample is not frozen as evidenced by a sharp water peak (data not shown).
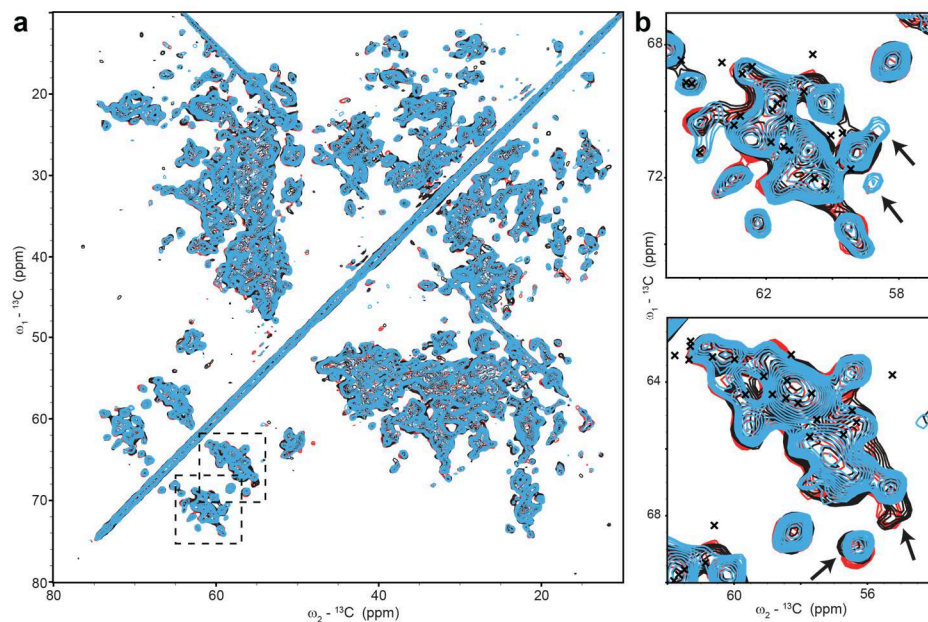
Figure 7. Structure and dynamics of BamA are largely preserved in different lipid bilayers. a) 2D ^{13}C,^{13}C correlation experiment with 30 ms PARIS mixing on BamA FL in DLPC (red), DMPC (black) and DOPC (blue) at a molar LPR of 25:1. b) Close-ups of the spectra shown by boxes in a). Crosses are FANDAS predictions for POTRA 1-5 as in Figure 6. Arrows point at slight variations in the spectra caused by a different lipid environment for some signals, which appear to come from the transmembrane domain rather than the POTRA domains.
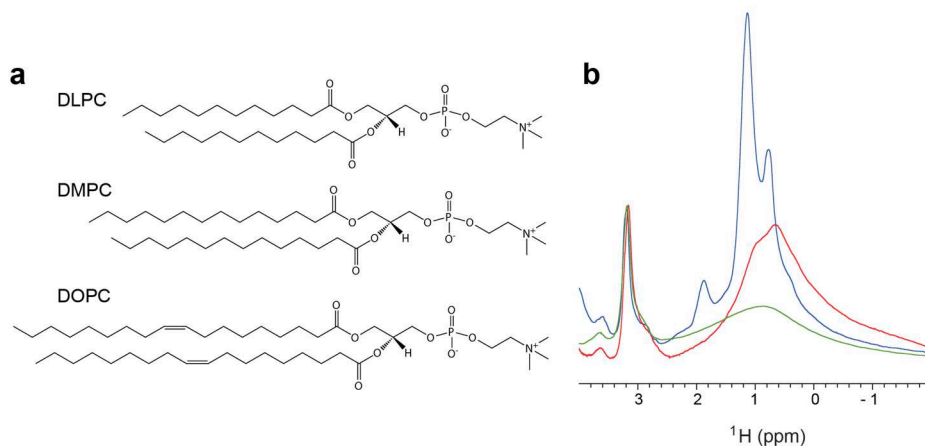
Figure 8. Chain lengths and phase behavior of PC lipids. a) Structures of DLPC (C12), DMPC (C14) and DOPC (C18:1). b) ^1H spectrum of BamA FL in DLPC (red), DMPC (green) and DOPC (blue) at LPR 25:1 recorded at 13 kHz MAS and an effective temperature of -2°C, scaled on the choline signal at 3.2 ppm. Liquid crystalline bilayers are characterized by two sharp acyl chain signals around 1 ppm, whereas in the gel phase the signals are broadened.

Which of these environments is closest to the Gram-negative outer membrane? The hydrophobic thickness of outer membrane proteins with known structure was found to be 23.9 ± 1.7 Å, which is much thinner than inner membrane proteins that span 30.2 ± 1.9 Å.[152] DLPC or DMPC would thus be appropriate in terms of chain length; however DMPC is extended to ~30 Å in the gel phase.[153] An additional complication arises from the fact that the outer leaflet of the outer membrane consists of LPS, of which the lipidic core is thought to be in the gel phase in intact *E. coli* cells at 37°C.[154] The commonly shared view that the liquid crystalline state is more native-like thus does not hold for outer membrane proteins. Our observations do however indicate that the BamA transmembrane domain can accommodate to different lipid bilayer environments and *vice versa*, as evidenced by EM (Fig. 1d).

Conclusions

Our results indicate that the entire periplasmic part of BamA, POTRA 1-5, is tightly connected to the β-barrel when reconstituted in lipid bilayers. Correct folding and reconstitution of BamA were probed by heat-modifiability on SDS PAGE as well as salt- and urea-extraction and limited proteolysis and these results were supported by ssNMR experiments probing lipid-protein contacts. Well-resolved correlations matching solution NMR assignments for POTRA 1 and 2 were observed in dipolar-based ssNMR experiments. Further experiments showed that overall motion of the POTRA domains is restricted even if protein concentrations are reduced, as well as at elevated temperatures. Our previous studies on a folded precipitate of full-length BamA demonstrated fast motion of the POTRA domains,[91] that may be slowed down by the stabilizing effect of the lipid bilayer in our current preparations. However, from the current data we cannot exclude local motion in POTRA domains whose ssNMR signals remain unresolved in our ssNMR spectra presented here. Likewise, our data may still be compatible with global domain motion on the millisecond timescale or slower that may occur to drive BamA-mediated folding and insertion of the substrate.

The ssNMR spectra of BamA in DLPC, DMPC and DOPC bilayers were found to be very similar, suggesting that the overall structure and dynamics are unaffected despite differences in lipid chain lengths and phase behavior. Differences in cross-peaks were confined to the transmembrane domain, implying a certain degree of conformational plasticity of the BamA β-barrel. These protein regions may also be responsible for the perturbance of the lipid bilayer as seen in our EM studies, suggesting that local high concentrations of BamA can disrupt the lipid bilayer. Deformation of the lipid bilayer by the BamA β-barrel was recently observed in

molecular dynamics simulations and linked to its remarkable asymmetric hydrophobic thickness.[41] Together with the conformational plasticity that we observe, this membrane deformation may have functional relevance for inserting substrates into the outer membrane.

With the intrinsic plasticity of POTRA domains being limited in membranes, we expect that the presence of the lipoproteins BamB-E as well as the composition of the cell envelope (peptidoglycan, LPS) will be critical to fully explain the conformational changes associated with the BAM-mediated process of protein insertion into outer membranes (Fig. 9). Again, ssNMR may be well suited to study such complexes in lipid bilayers. In extended conformation, the POTRA domains might span the entire periplasmic region and even contact the SecYEG machinery (Fig. 9). Hence, the present study also paves the way to investigate these phenomena in cellular preparations such as previously described by our group for the outer membrane enzyme PagL from *Pseudomonas aeruginosa*.[81,82]
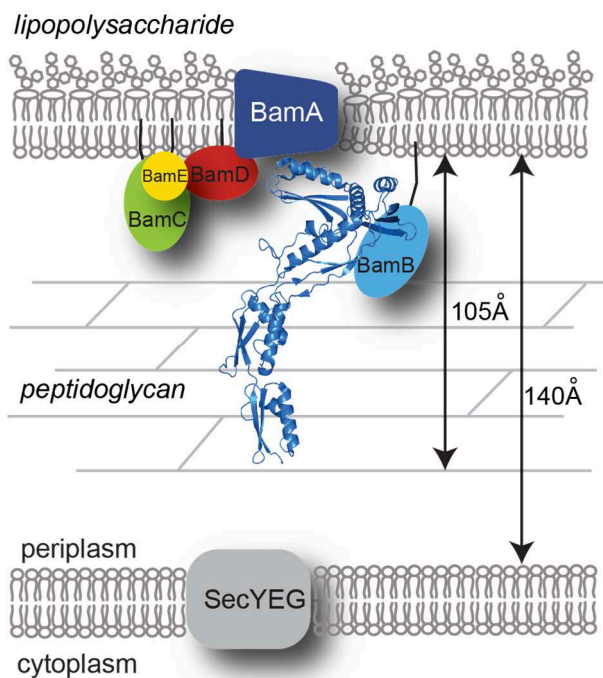
Figure 9. Schematic view of the BAM complex inside the cell envelope of *E. coli*. The BamA β-barrel has an asymmetric hydrophobic thickness,[41] which perturbs the lipid bilayer. In extended conformation, the POTRA domains measure 105 Å meaning they can almost span the periplasm, which has been shown to be at most ~140 Å.[38,155,156] The POTRA structural model is composed of crystal structures of POTRA 1-4 (PDB 3EFC) and POTRA 4-5 (PDB 3Q6B) aligned on POTRA 4.

CHAPTER 4

Solid-state NMR assignments of membrane-embedded BamA: implications for dynamics within the transmembrane domain

Tessa Sinnige & Marc Baldus

Abstract

Solid-state NMR (ssNMR) is a powerful technique to study the structure and dynamics of membrane proteins in a lipid bilayer environment. We investigated the main component of the β-barrel assembly machinery (BAM), BamA, a 790 residue protein that is involved in folding and insertion of outer membrane proteins in Gram-negative bacteria. Using two truncated constructs including the BamA transmembrane domain and one or two soluble POTRA domains (465 and 547 residues respectively) in combination with specific labeling schemes tailored to the sequence of BamA, we succeeded in assigning a number of residues based on 2D and 3D ssNMR experiments. Secondary chemical shifts from the ssNMR assignments were in close agreement to previous crystal structures. Interestingly, we obtained resonance assignments for residues in the extracellular loop 6 that was shown to be crucial for BamA-mediated folding and insertion of substrates, implying that this loop is relatively rigid in isolated BamA in lipid bilayers. However, the BamA transmembrane domain may exhibit local structural fluctuations. Such processes would explain the absence of ssNMR assignments around β-strands 1 and 16, which are hypothesized to form a lateral gate for substrate release into the lipid bilayer.

Introduction

The outer membrane protein (OMP) BamA is the main component of the β-barrel assembly machinery (BAM)[32,33] that folds and inserts OMPs into the outer membrane of Gram-negative bacteria, where they serve e.g. as porins, receptors and enzymes. BamA consists of a transmembrane (TM) domain, which is a 16-stranded β-barrel, and a periplasmic extension consisting of five POTRA domains in *Escherichia coli*.[35,41] Despite recently solved crystal structures of BamA from *Neisseria gonorrhoeae* and *Haemophilus ducreyi* in bicelles[41] and the TM domain of *E. coli* BamA in detergent micelles,[42] the mechanism of BamA-mediated OMP folding and insertion is still poorly understood.

In recent years, magic angle spinning (MAS) solid-state NMR (ssNMR) has proven a valuable method for structural studies of membrane proteins (see e.g. Ref. 75,76,79,80,123,140) that can be probed in a native-like environment consisting of lipid bilayers, a feature that is not easily accessible by other structural biology techniques. The power of NMR spectroscopy is that atomic-resolution information on the structure, dynamics and interactions of the biomolecule of interest can be obtained. Previously, ssNMR studies from our lab showed that the periplasmic POTRA domains of BamA can undergo fast motion in a precipitated state or when truncated to the C-terminal POTRA 5 domain in membranes.[91] In contrast, our recent ssNMR experiments of membrane-embedded BamA suggested the absence of fast overall POTRA motion in constructs containing at least POTRA 4 and POTRA 5.[157] In the latter studies, we made use of the fact that soluble POTRA domains encompassing around 80 residues each can easily be studied using solution NMR methods[39] and the resonance assignments can be transferred to the ssNMR spectra (see Chapter 5). However, to obtain atomic-resolution information on the structure and dynamics of the BamA TM domain in a native-like environment consisting of lipid bilayers, ssNMR resonance assignments are required.

Compared to solution NMR methods for sequential assignments which rely on ^{1}H detection,[158,159] ssNMR has the drawback of strong ^{1}H,^{1}H dipolar couplings that are not averaged out in the absence of molecular tumbling. Therefore, ssNMR studies using MAS and randomly oriented biomolecules have relied on the detection of heteronuclei (^{13}C,^{15}N). In such studies, ^{13}C,^{13}C correlation spectra yield intra- or inter-residue correlations depending on the mixing time and NCACX and NCOCX experiments are typically used to establish backbone correlations (see Ref. 160). Applications of such methods resulted in the first ssNMR assignments for microcrystalline protein samples of e.g. the α-spectrin SH3 domain (62 residues),[161] Crh (dimer of 2x 85 residues)[162] or ubiquitin (76 residues).[163-165] More recently, progress has been made on membrane proteins such as the ion channel KcsA (4 x

160 residues)[139], the bacterial inner membrane protein DsbB (176 residues)[141,166] and the seven-helix receptor *Anabaena* sensory rhodopsin (ASR, 235 residues)[167] that were assigned to near completion. A large part of the 249-residue sensory rhodopsin II from *Natronomonas pharaonis* (NpSRII) was also assigned.[140] Crystals of the C-terminal domain of the Ure2p prion (285 residues) represent the largest protein that has been assigned by ssNMR so far.[168]

BamA encompassing 790 amino acids or its TM domain (390 residues) hence represent a challenge for current ssNMR methods. Because a possible effect of the POTRA domains on the conformation of the β-barrel cannot be excluded, we in the following describe an ssNMR approach that applies tailored isotope labeling strategies on constructs containing one or two POTRA domains in addition, leading to membrane proteins of 465 and 547 residues, respectively (Fig. 1).

With this strategy, we succeeded in acquiring well-resolved 2D and 3D ssNMR spectra, leading to *de novo* NMR assignments of the BamA β-barrel. We obtained secondary structure information in functionally important regions of the transmembrane part of BamA that, to a large extent, agreed with available crystal structures. Moreover, these studies shed light onto protein dynamics of the BamA TM domain in lipid bilayers, suggesting that the β-barrel region that closely interacts with the extracellular loop 6 in the crystal structures is relatively rigid.

Results

In the following, we describe two and three-dimensional ssNMR studies using optimized labeling schemes to obtain resonance assignments of the BamA TM domain. These efforts range from protein labeling patterns such as the VLKY reverse labeling (Fig. 1) aiming at an overall characterization of all structural elements of BamA, to more advanced schemes that minimize spectral crowding and allow to unambiguously study secondary structure and dynamics of specific protein regions critical for BamA function.

```
                eL1          pL1         eL2        pL2           eL3
        β1  ┤├ β2        ┤├  β3    ┤├ β4       ┤├  β5  [α1]     β6

VLKY reverse  RNTGSFNFGI GYGTESGVSF QAGVQQDNWL GTGYAVGING TKNDYQTYAE LSVTNPYFTV DGVSLGGRLF YNDFQADDAD LSDYTNKSYG TDVTLGFPIN
VLK forward   RNTGSFNFGI GYGTESGVSF QAGVQQDNWL GTGYAVGING TKNDYQTYAE LSVTNPYFTV DGVSLGGRLF YNDFQADDAD LSDYTNKSYG TDVTLGFPIN
GS(C)AVLTI    RNTGSFNFGI GYGTESGVSF QAGVQQDNWL GTGYAVGING TKNDYQTYAE LSVTNPYFTV DGVSLGGRLF YNDFQADDAD LSDYTNKSYG TDVTLGFPIN
VRGS ¹⁵N/¹³C   RNTGSFNFGI GYGTESGVSF QAGVQQDNWL GTGYAVGING TKNDYQTYAE LSVTNPYFTV DGVSLGGRLF YNDFQADDAD LSDYTNKSYG TDVTLGFPIN
              430        440        450        460        470        480        490        500        510        520


        pL3                α2          eL4              pL4          eL5
          β7        ┤├          ┤├           β8    ┤├      β9   ┤├        β10

VLKY reverse  EYNSLRAGLG YVHNSLSNMQ PQVAMWRYLY SMGEHPSTSD QDNSFKTDDF TFNYGWTYNK LDRGYFPTDG SRVNLTGKVT IPGSDNEYYK VTLDTATYVP
VLK forward   EYNSLRAGLG YVHNSLSNMQ PQVAMWRYLY SMGEHPSTSD QDNSFKTDDF TFNYGWTYNK LDRGYFPTDG SRVNLTGKVT IPGSDNEYYK VTLDTATYVP
GS(C)AVLTI    EYNSLRAGLG YVHNSLSNMQ PQVAMWRYLY SMGEHPSTSD QDNSFKTDDF TFNYGWTYNK LDRGYFPTDG SRVNLTGKVT IPGSDNEYYK VTLDTATYVP
VRGS ¹⁵N/¹³C   EYNSLRAGLG YVHNSLSNMQ PQVAMWRYLY SMGEHPSTSD QDNSFKTDDF TFNYGWTYNK LDRGYFPTDG SRVNLTGKVT IPGSDNEYYK VTLDTATYVP
              530        540        550        560        570        580        590        600        610        620


        pL5                α3                    eL6
          β11       ┤├          ┤├                                      β12

VLKY reverse  IDDDHKWVVL GRTRWGYGDG LGGKEMPFYE NFYAGGSSTV RGFQSNTIGP KAVYFPHQAS NYDPDYDYEC ATQDGAKDLC KSDDAVGGNA MAVASLEFIT
VLK forward   IDDDHKWVVL GRTRWGYGDG LGGKEMPFYE NFYAGGSSTV RGFQSNTIGP KAVYFPHQAS NYDPDYDYEC ATQDGAKDLC KSDDAVGGNA MAVASLEFIT
GS(C)AVLTI    IDDDHKWVVL GRTRWGYGDG LGGKEMPFYE NFYAGGSSTV RGFQSNTIGP KAVYFPHQAS NYDPDYDYEC ATQDGAKDLC KSDDAVGGNA MAVASLEFIT
VRGS ¹⁵N/¹³C   IDDDHKWVVL GRTRWGYGDG LGGKEMPFYE NFYAGGSSTV RGFQSNTIGP KAVYFPHQAS NYDPDYDYEC ATQDGAKDLC KSDDAVGGNA MAVASLEFIT
              630        640        650        660        670        680        690        700        710        720


        pL6               β13              eL7              pL7      β15        eL8      β16

VLKY reverse  PTPFISDKYA NSVRTSFFWD MGTVWDTNWD SSQYSGYPDY SDPSNIRMSA GIALQWMSPL GPLVFSYAQP FKKYDGDKAE QFQFNIGKTW
VLK forward   PTPFISDKYA NSVRTSFFWD MGTVWDTNWD SSQYSGYPDY SDPSNIRMSA GIALQWMSPL GPLVFSYAQP FKKYDGDKAE QFQFNIGKTW
GS(C)AVLTI    PTPFISDKYA NSVRTSFFWD MGTVWDTNWD SSQYSGYPDY SDPSNIRMSA GIALQWMSPL GPLVFSYAQP FKKYDGDKAE QFQFNIGKTW
VRGS ¹⁵N/¹³C   PTPFISDKYA NSVRTSFFWD MGTVWDTNWD SSQYSGYPDY SDPSNIRMSA GIALQWMSPL GPLVFSYAQP FKKYDGDKAE QFQFNIGKTW
              730        740        750        760        770        780        790        800        810
```

Figure 1. Labeling schemes used in this study shown on the sequence of the BamA TM domain. The topology and secondary structural elements (β-strands, α-helices, periplasmic and extracellular loops, taken from crystal structure of the *E. coli* BamA TM domain, PDB 4N75) are indicated at the top. The residues are color coded according to the labeling: red ^{13}C,^{15}N labeled; blue ^{13}C labeled; green ^{15}N labeled. Sequentials of labeled residues are shown underlined. Labeling schemes GS(C)AVLTI are combined into one row with cysteine residues in grey.

Reverse labeling of VLKY yields well-resolved ssNMR spectra of BamA P5-TM

Firstly, we utilized the construct BamA P5-TM, containing the transmembrane domain and POTRA 5 which is connected N-terminally to the β-barrel. To alleviate spectral crowding for this 465 residue construct, a sample with reverse labeling of the residues VLKY was prepared (denoted BamA P5-TM (-VLKY)), since these amino acids are highly abundant (31, 25 and 21 residues respectively, Fig. 1). Moreover, we made use in our analysis of our recent observation that the P5 domain remains largely unfolded in the absence of P4 (see Chapter 5).

The protein was reconstituted in dilauroyl phosphatidylcholine (DLPC) at a low molar lipid-to-protein ratio (LPR) of 21:1 (0.25 w/w) to obtain sufficient sensitivity in ssNMR experiments. The resulting sample yielded a high resolution 2D ^{13}C,^{13}C correlation spectrum, in which the spin systems of several residues could be recognized (Fig. 2).

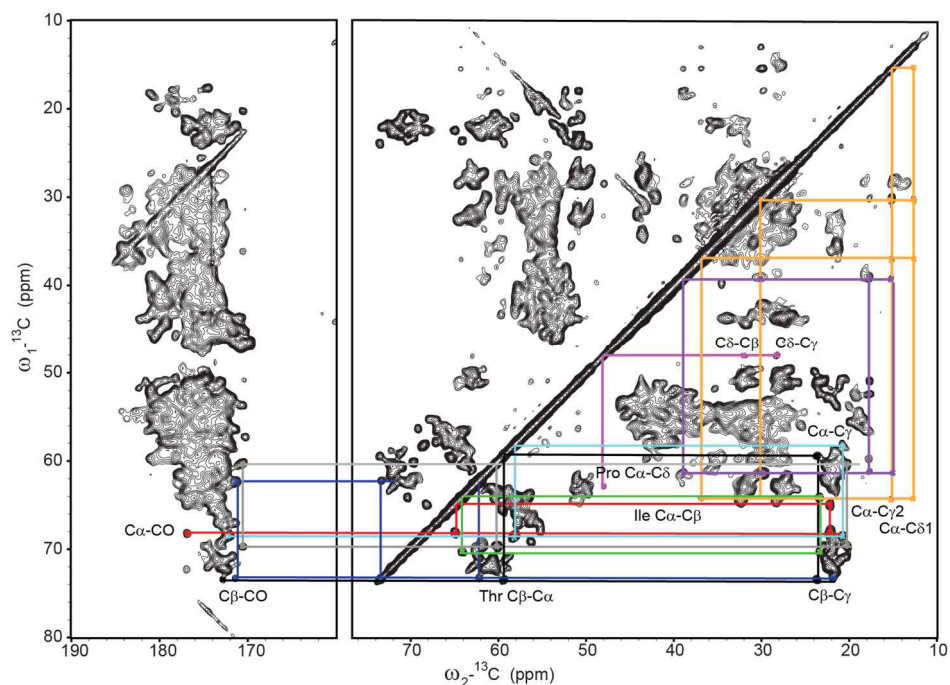**Figure 2.** 2D ^{13}C,^{13}C experiment with 40 ms PARIS mixing time recorded on BamA P5-TM (-VLKY) reconstituted in DLPC liposomes at a LPR of 21:1. Colored lines show examples of spin systems connecting cross-peaks within a given residue. The spectrum was recorded at 13 kHz MAS and an effective sample temperature of -2°C.

To obtain sequential information, a 3D NCOCX was recorded (Fig. 3), in which correlations only occur if two neighboring residues are both labeled (Fig. 1). Although many peaks were resolved in this experiment and some amino acid types could be recognized, most planes of the 3D spectrum still showed considerable overlap. For this reason, an NCACX experiment, in which more correlations are expected, was not performed on this sample.
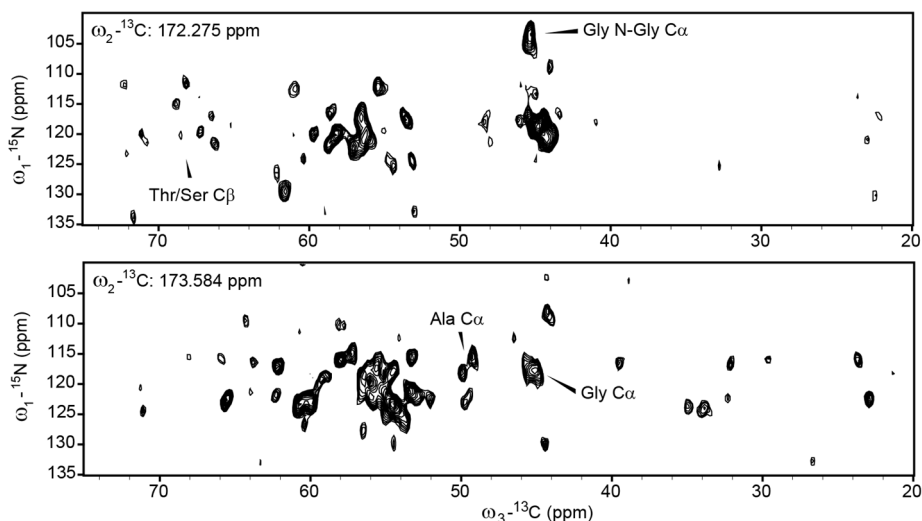
Figure 3. Examples of 2D planes from the 3D NCOCX spectrum with 30 ms PARIS mixing time on BamA P5-TM (-VLKY) reconstituted in DLPC liposomes at a LPR of 21:1. Labels indicate characteristic chemical shifts for certain residue types. The spectrum was recorded at 11 kHz MAS and an effective sample temperature of -2°C.

Finding unique sequentials using a minimal forward labeling scheme

Instead, a sample was prepared with only three residue types labeled to obtain a minimum of sequential correlations (Fig. 1, second lane). For this forward labeling strategy, U-^{13}C,^{15}N Val, Leu and Lys were added to the culture medium. The resulting VLK sample should reveal the unique sequentials VVL, KK, LV and KL within the TM domain of BamA. Based on a set of 2D experiments, we could identify and partially assign the sequentials K580-L581, V628-V629-L630 and K792-K793 (Fig. 4). These residues are located in β-strand 8, β-strand 11, and the border of extracellular loop 8 and β-strand 15, respectively (vide infra).
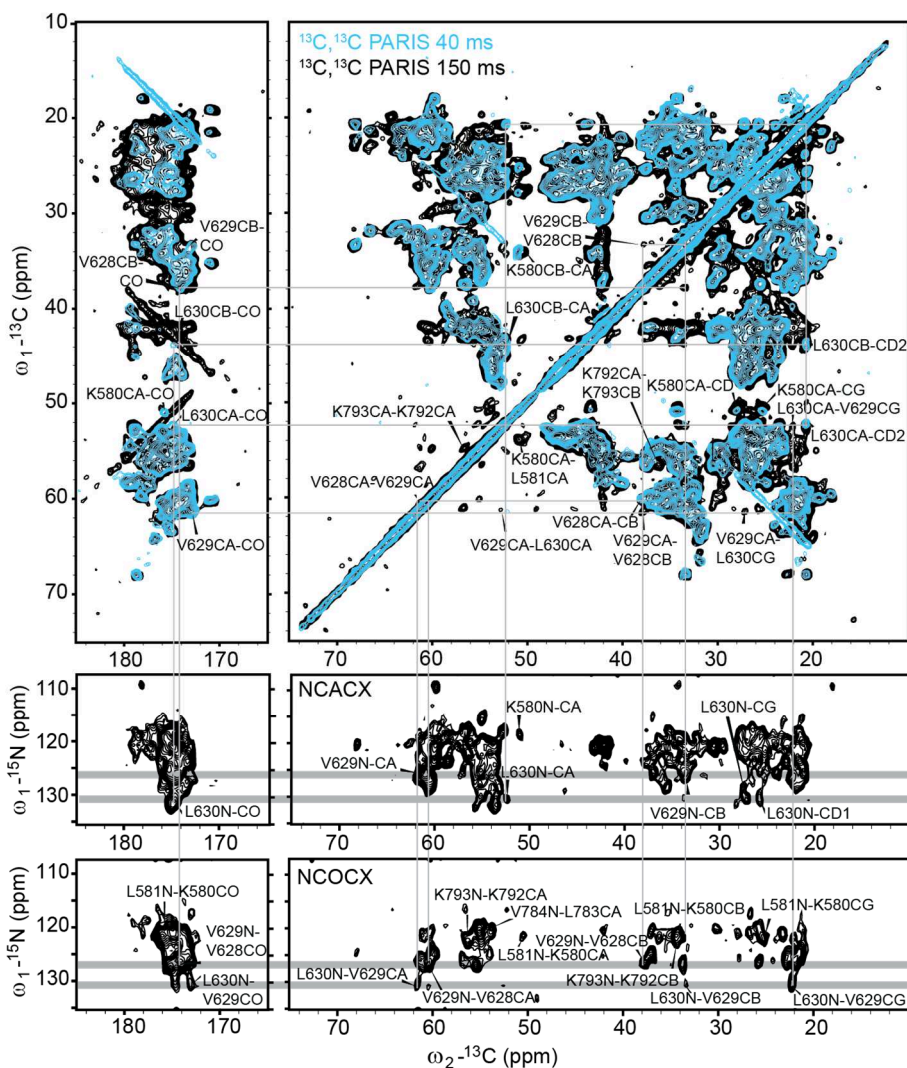
Figure 4. Assignment of unique sequentials in BamA P5-TM (VLK) reconstituted in DLPC at LPR 21:1. Shown are 2D ¹³C,¹³C spectra with a PARIS mixing time of 40 ms (blue, top) and 150 ms (black, top); and 2D NCACX (middle) and NCOCX (bottom) experiments with 40 ms PARIS mixing. Grey lines and bars indicate intra- and inter-residue correlations that led to the assignments that are labeled in the spectra. Experiments for intra-residue correlations (¹³C,¹³C with short mixing time and NCACX) were recorded at 15 kHz MAS, whereas the experiments for sequential contacts were recorded at 11 kHz MAS. In all experiments the effective sample temperature was kept at -2°C.

More extended labeling scheme at high spectral resolution

Next, we proceeded with an approach that increased residue-specific labeling at minimum spectral overlap. Residues GSCAVLTI were selected (Fig. 1, third row) because they can easily be identified by virtue of their ^{13}C chemical shift correlations. They were labeled using a combination of forward labeling of Thr and Ile and reverse labeling of the amino acids DEFHKMNPQRYW. Unfortunately, this strategy resulted in scrambling. Since glutamine and glutamate are subject to transamination in *E. coli* metabolism, scrambling of the ^{14}N atoms of these reverse labeled amino acids caused the overall ^{15}N intensity of the sample to drop dramatically compared to the previous samples and relative to the ^{13}C intensity. A 3D NCOCX experiment was therefore not feasible, but instead a 3D NCACX was recorded that was helpful to assign spin systems. This spectrum (Fig. 5a) was well-resolved and the individual spin systems could be traced to their respective amino acid types. Interestingly, BamA contains only two cysteine residues that reside in extracellular loop 6 (eL6) and presumably form a disulfide bond.[69] However, the chemical shift values we detected experimentally did not correspond to oxidized cysteine (Fig. 5b), suggesting flexibility of this part of eL6 (see below).
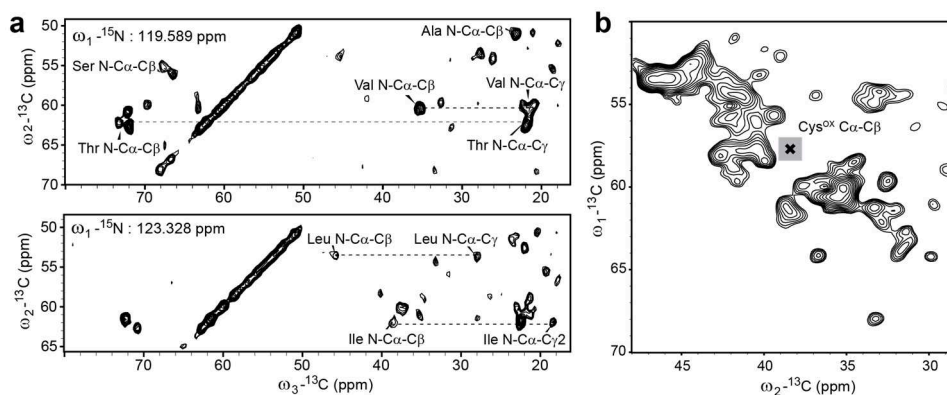

Figure 5. a) 2D planes from a 3D NCACX spectrum with 30 ms PARIS mixing recorded on BamA P5-TM (GSCAVLTI) in DLPC bilayers at a LPR of 8:1. b) Close-up of a 2D $^{13}C,^{13}C$ experiment with 30 ms mixing time, showing the expected position of the C_α-C_β correlation for oxidized cysteine in random coil conformation (black cross; grey box indicates the standard deviation).[149] The spectra were acquired at 13 kHz MAS and an effective sample temperature of -2°C.

Identification of residues in BamA's extracellular loop 6

Since the combination of reverse and forward labeling resulted in nitrogen scrambling, we subsequently opted for forward labeling of all residues GSAVLTI. Because the two cysteine residues were not observed previously, this amino acid was not further included. This labeling strategy was applied to the BamA P4P5-TM construct. We note that including an additional POTRA domain did not affect the resonances from the TM domain (data not shown).

2D ^{13}C,^{13}C experiments with different mixing times (Fig. 6) matched well with FANDAS predictions[145] for sequential correlations based on ShiftX2 predictions[146] for the TM domain and solution NMR assignments for P4P5 (see Chapter 5). By analyzing these experiments, a 2D NCOCX data set (not shown) and the 3D experiments on the samples discussed before, the unique sequential stretches T667-I668-G669 and T600-I601 could be assigned. Remarkably, the stretch T667-I668-G669 is located in the extracellular loop 6 (eL6), which we speculated might be dynamic (see also Ref. 69). However, in the crystal structures of BamA from *N. gonorrhoeae*, *H. ducreyi* and *E. coli*, the loop partially folds back into the β-barrel.[41,42] Notably, the assigned TIG stretch is very close to the conserved motif VRGF which in the crystal structures is locked to the wall of the β-barrel by several interactions.

To examine eL6 and in particular the VRGF motif in more detail, we designed a tailored labeling strategy to be able to unambiguously assign these residues. We labeled BamA P4P5-TM with ^{13}C-Val, ^{13}C,^{15}N-Arg and ^{15}N-Gly and Ser, leading to V-R/S/G and R-R/G sequentials in 2D NCO and NCOCX experiments (Fig. 7, see also Fig. 1). Comparison with solution NMR assignments of P4P5 (see Chapter 5) corroborated the labeling strategy (Fig. 7, orange symbols). In addition, several cross-peaks were present that could be attributed to the TM domain, yet not all predictions matched well enough to serve as an initial assignment guess (Fig. 7, purple symbols). In the 2D NCO experiment, more correlations were observed than expected, which may point to the presence of different conformations or to unspecific magnetization transfer (Fig. 7, left). The 2D NCOCX experiment on the other hand lacked many cross-peaks (Fig. 7, right), caused by limited sensitivity and potentially by protein dynamics. However, based on their different secondary structure, V660-R661 from the VRGF motif (random coil) and V733-R734 (β-strand) could be tentatively assigned.[149]
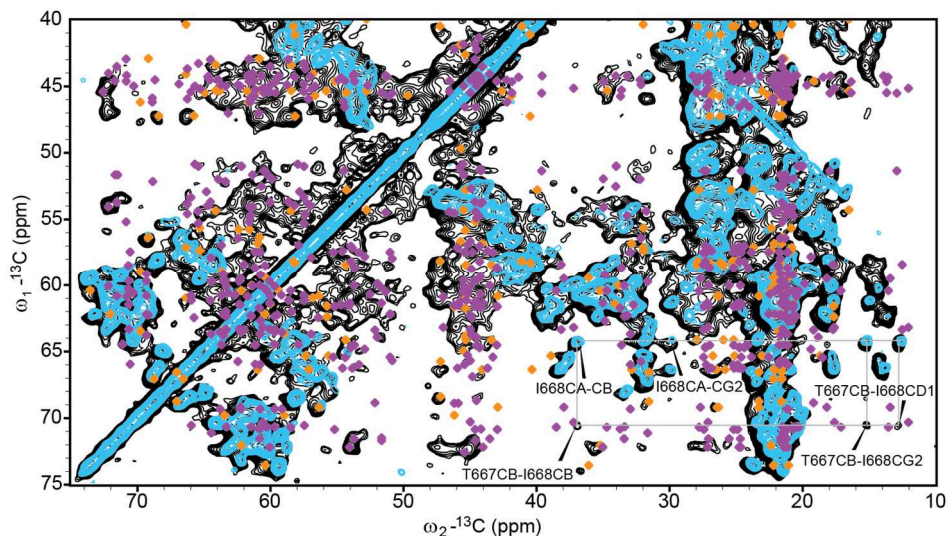
Figure 6. 2D ^{13}C,^{13}C spectra with 30 ms (cyan) and 150 ms (black) PARIS mixing time on the sample BamA P4P5-TM (GSAVLTI) reconstituted in DLPC at LPR 10:1. The sequential correlations between T667 and I668 are labeled and connected with grey lines. Dots indicate FANDAS predictions[145] for sequential correlations based on ShiftX2 predictions[146] for the TM domain (purple) and solution NMR assignments of P4P5 (orange). The experiment with 30 ms mixing was recorded at 13 kHz MAS frequency, whereas the experiment with 150 ms mixing was recorded at 11 kHz. In both cases the effective sample temperature was set to -2°C.
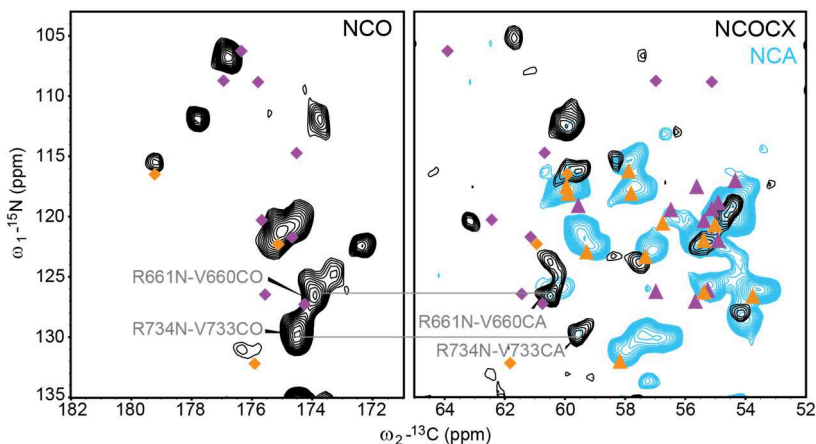

Figure 7. Sequential correlations in BamA P4P5-TM labeled with ^{13}C-Val, ^{13}C,^{15}N-Arg and ^{15}N-Gly and Ser, reconstituted in DLPC at LPR 10:1. Shown are a 2D NCO spectrum (left panel), and an overlay of a 2D NCOCX (right panel, black) and a 2D NCA (right panel, cyan). Diamonds represent predicted inter-residue correlations and triangles intra-residue correlations (orange: P4P5; purple: TM domain). Labels indicate tentative sequential assignments based on secondary structure. Spectra were recorded at 11 kHz MAS (NCO, NCOCX) or 13 kHz MAS (NCA) and an effective sample temperature of 4°C.

All tentative assignments obtained for the BamA TM domain from the different samples are listed in Table 1. The resulting conformation-dependent secondary chemical shifts based on C_α and C_β assignments (Fig. 8a) were in good agreement with the crystal structure of the *E. coli* BamA TM domain (see Fig. 1).[42] The only exception is K793 which showed a β-strand rather than random coil value, but this residue is located at the tip of β-strand 15 (Fig. 8b). Possibly, this β-strand is slightly elongated in our proteoliposomal BamA preparations. In addition, I668 showed a rather α-helical secondary chemical shift, although it has a random coil conformation in the crystal structures, being located in eL6 (Fig. 8b).

Table 1. *Tentative ssNMR chemical-shift (CS) assignments of the BamA β-barrel in DLPC bilayers. Black: tentative assignments with high confidence based on sequential correlations in at least two experiments; italics: tentative assignments based on secondary structure.*

Residue	Atom	CS (ppm)	Residue	Atom	CS (ppm)
K580	N	118.2	L630	CA	52.3
K580	CA	50.9	L630	CB	44.0
K580	CB	34.2	L630	CG	27.2
K580	CD	28.0	L630	CD1	25.8
K580	CG	25.2	L630	CD2	20.6
K580	CE	43.2	*V660*	*CO*	*173.9*
K580	CO	175.8	*V660*	*CA*	*60.5*
L581	N	121.3	*R661*	*N*	*126.6*
L581	CA	54.0	T667	CB	70.5
T600	CO	172.2	I668	N	115.9
T600	CA	61.7	I668	CA	64.2
T600	CB	72.3	I668	CB	37.0
T600	CG	22.4	I668	CG1	30.0
I601	N	130.1	I668	CG2	15.2
I601	CA	61.3	I668	CD1	12.6
I601	CB	38.8	I668	CO	176.6
I601	CG1	29.9	G669	N	111.8
I601	CG2	17.7	G669	CA	45.8
I601	CD1	15.0	*V733*	*CO*	*174.5*
V628	CA	60.2	*V733*	*CA*	*59.6*
V628	CB	37.9	*R734*	*N*	*129.8*
V629	N	126.7	K792	CA	56.6
V629	CA	61.5	K793	N	122.2
V629	CB	33.5	K793	CA	54.8
V629	CG	22.2	K793	CB	36.2
L630	N	130.5			

Figure 8. a) Secondary chemical shifts of BamA TM residues for which C_α and C_β were assigned (or only C_α for glycine). $\Delta C_{\alpha/\beta} = C_{\alpha/\beta} - C_{\alpha/\beta,rc}$.[149] Positive values are indicative of α-helical conformation; negative of β-strand. b) Assigned residues shown as sticks on the *E. coli* BamA TM domain (PDB 4N75).[42] Purple residues are located in extracellular loops (eL); blue residues in β-strands.

Discussion

Specific labeling schemes applied to BamA

In this work, specific isotope labeling schemes were used to obtain ssNMR resonance assignments for the β-barrel domain of the outer membrane protein BamA. This domain with its 390 residues is larger than any protein that has so far been assigned using MAS ssNMR on unoriented samples. Moreover, we included one or two POTRA domains not to cause artifacts on the interface of the TM domain, leading to constructs of 465 and 547 residues, respectively. Given these sizes, several problems were encountered when trying to obtain resonance assignments using standard ssNMR routines. Reverse labeling of the amino acids VLKY, which eliminated 77 residues and shortened sequential stretches (Fig. 1), did not provide sufficient opportunities for sequential assignment. Besides spectral crowding, finding a starting point for assignment proved difficult because sequential pairs that are unique in the sequence rarely occur. In addition, the ^{15}N frequencies were too degenerate to connect the spin systems based on NCACX and NCOCX experiments. A CONCA experiment could help to solve this problem, but since this pulse sequence contains an additional SPECIFIC-CP element,[101] the signal-to-noise ratio is significantly less and not sufficient to record a 3D experiment on a sample such as BamA in a reasonable amount of time. Sequential correlations may also be derived

from 2D ^{13}C,^{13}C spin diffusion experiments,[169] but these suffered heavily from spectral crowding in the case of BamA.

For these reasons, we used sparse labeling schemes with selectively labeled amino acids, VLK and GSAVLTI, that sufficiently limited spectral crowding and sequence ambiguity (see Fig. 1) to enable the tentative assignment of several short sequentials in the BamA TM domain based on 2D and 3D ssNMR experiments (Table 1). These residues are located in β-strands and at the beginning of extracellular loops, but also in extracellular loop 6 (eL6) (Fig. 8b). Secondary chemical shifts largely corresponded to the secondary structural elements in the crystal structure of the *E. coli* BamA TM domain (PDB 4N75), although β-strand 15 appeared elongated, as K793 was found to be in a β-strand conformation in our experiments, whereas it is the first residue of eL8 in the crystal structure (Fig. 8b).

Implications for BamA dynamics

The long extracellular loop 6 is crucial for functioning of BamA[41,67] and has been proposed to undergo a conformational change related to substrate insertion.[69] A short stretch from eL6, T667-I668-G669, was tentatively assigned using the GSAVLTI labeling strategy. We further examined this loop using an even sparser labeling scheme including amino acids that are only ^{15}N or ^{13}C labeled, targeted at the conserved VRGF motif in eL6. Using this strategy, we obtained tentative assignments for V660-R661 from the conserved motif. These results suggest that, in steady-state, at least part of this loop is relatively rigid. In all BamA crystal structures currently available,[41,42] eL6 similarly folds back into the β-barrel, where the conserved VRGF motif is stabilized by interactions with the β-barrel wall (Fig. 9a). Perhaps, binding of the substrate is required for a conformational change in eL6 to occur. Also the lipoproteins BamD and BamE have been implied in conformational cycling of BamA between a protected and exposed eL6.[69] However, the part of eL6 that protrudes further away from the β-barrel may be slightly more mobile in our sample preparations. The two cysteine residues that are present in the insertion in eL6 of *E. coli* BamA and supposedly form a disulfide bond[69] were not observed in our experiments, in agreement with lacking electron density in the crystal structures (Fig. 9a).

Remarkably, most of the assigned residues map to the side of the BamA TM domain where eL6 is coordinated (Fig. 8b). Clearly, these results may be biased by the choice of labeling schemes and the fact that spectrally resolved chemical shifts can be more readily assigned, but they do imply that this part of the domain is altogether rigid. On the other hand, no residues around the proposed lateral gate[41]

between β-strands 1 and 16 could be assigned in this study, although with the labeling schemes VLK and GSAVLTI, a significant portion of the terminal β-strands was labeled in our experiments (Fig. 9b). Whether these residues could not be assigned because of spectral crowding and sequence ambiguity, or as a consequence of dynamics, remains to be determined. A tailored labeling scheme to probe specific sequential residues in β-strands 1 and 16 would hence be highly desirable.
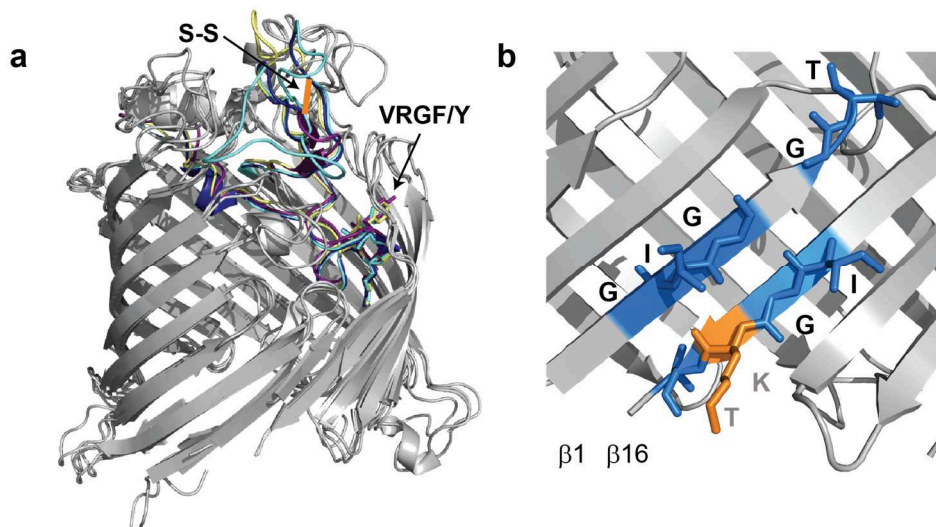

Figure 9. a) Overlay of available BamA crystal structures with eL6 colored (*E. coli* PDB 4C4V blue; *E. coli* PDB 4N75 cyan; *N. gonorrhoeae* PDB 4K3B purple; *H. ducreyi* PDB 4K3C yellow). The conserved VRGF/Y motif is shown in sticks and the expected disulfide bond in *E. coli* BamA that is not resolved in the crystal structures is indicated with an orange line. b) Labeling of β-strands 1 and 16 resulting from the amino acid schemes GSAVLTI (blue) and VLK (orange) shown in sticks on the *E. coli* BamA TM structure (PDB 4N75). Residues indicated in black form sequentials in one of the labeling schemes, whereas those in grey are isolated in the individual labeling schemes.

Other assignment strategies for large proteins

We have shown that a highly specific labeling scheme using amino acids with only ^{13}C or ^{15}N labeling is well suited for the spectral elucidation of specific protein sites. To obtain a larger number of assignments, however, this strategy may become impractical because many samples have to be prepared. Apart from specific amino acid labeling, labeling with precursors such as 1,3-^{13}C and 2-^{13}C glycerol[143] or 1-^{13}C and 2-^{13}C glucose[144,170] has been applied in ssNMR assignments of membrane proteins and fibrils. These methods have the advantage that spectral resolution is further improved by elimination of ^{13}C scalar couplings. Such labeling schemes can

be expected to yield high-resolution spectra for BamA, but they rely on a well-characterized uniformly labeled sample and seem to be more useful in later stages of finalizing the assignments and collecting distance restraints for structure determination.

A different approach to limit spectral crowding is segmental isotope labeling, in which only the domain of interest is labeled and subsequently fused to the unlabeled remainder of the protein, using the intein or sortase ligation systems (see e.g. Ref. 171-174). For BamA, segmental isotope labeling of the TM domain is not very beneficial compared to the constructs used in our current study, because the POTRA domains make up only a small number of the total amount of residues and P4P5 has been assigned in solution (see Chapter 5). However, BamA can be refolded from a denatured state, so if conditions could be established to perform the ligation under denaturing conditions, one could imagine expressing part of the β-barrel labeled and the remainder unlabeled, after which the ligation product could be refolded. Developing such a protocol is however not trivial. For large soluble proteins, optimization of the ligation conditions to obtain sufficient material for NMR studies is already tedious[175,176] and to the best of our knowledge segmental isotope labeling has not yet been reported for ssNMR applications on membrane proteins.

Lastly, proton-detected methods are rapidly evolving in ssNMR (see e.g. Ref. 87-89,177-179) and may greatly improve signal-to-noise as well as resolution provided by the additional dimension. Such methods, possibly combined with specific labeling schemes, bear the potential to assign a larger number of residues of the BamA β-barrel in the future (see Chapter 7).

In conclusion, we presented the first ssNMR resonance assignments for the BamA TM domain based on specific labeling schemes. Although at present the assignments are sparse, we could conclude that the region of eL6 carrying the conserved VRGF motif is relatively rigid in the context of isolated BamA in lipid bilayers. Dynamics in the TM domain may however occur in other parts of the domain, possibly around β-strands 1 and 16. Further assignments as well as interaction studies with lipoproteins and substrate will be required to resolve the dynamic behavior of BamA in relation to OMP folding and insertion.

CHAPTER 5

Conformational plasticity of the POTRA 5 domain in the outer membrane protein assembly factor BamA

Tessa Sinnige, Klaartje Houben, Iva Pritisanac, Markus Weingarth, Marie Renault, Mark Daniëls, Rolf Boelens, Alexandre M.J.J. Bonvin, Jan Tommassen, Marc Baldus

Abstract

BamA is the main component of the β-barrel assembly machinery (BAM) that folds and inserts outer membrane proteins in Gram-negative bacteria. Crystal structures have suggested that this process involves conformational changes in the transmembrane β-barrel of BamA that allow for lateral opening, as well as large overall rearrangements of its periplasmic POTRA domains. Here, we identify local dynamics of the BamA POTRA 5 domain by solution and solid-state NMR. The protein region undergoing conformational exchange is highly conserved and contains residues critical for interaction with BamD and for correct β-barrel assembly *in vivo*. We show that mutations known to affect the latter processes influence the conformational equilibrium, suggesting that the plasticity of POTRA 5 is related to its interaction with BamD and possibly to substrate binding. Taken together, a view emerges in which local protein plasticity may be critically involved in the different stages of outer membrane protein folding and insertion.

Introduction

The process of protein folding and insertion into bacterial membranes is essential for physiological, pathogenic and drug resistance functions.[23] The mechanism for membrane integration of α-helical membrane proteins is relatively well understood.[4,180] On the other hand, the insertion and folding of β-barrel membrane proteins that are only found in the outer membranes of Gram-negative bacteria and eukaryotic organelles (mitochondria and chloroplasts) have remained elusive. Transport of proteins into or through these outer membranes usually requires complex molecular machines. In *Escherichia coli*, the precursors of outer-membrane β-barrel proteins are synthesized in the cytoplasm. The unfolded precursor proteins are recognized and translocated across the inner membrane by the Sec translocase. Insertion into the outer membrane is subsequently coordinated by the β-barrel assembly machinery (BAM).[24,32,33,181] Its main component, BamA, is a highly conserved integral membrane protein,[52] consisting of a transmembrane β-barrel and five periplasmic POTRA domains in *E. coli*. The *E. coli* BAM complex contains four lipoproteins in addition, BamB and the BamCDE sub-complex.[33,50] Only BamA and BamD are essential, but deletion of the other lipoproteins leads to defects in OMP assembly and increased sensitivity to antibiotics and stress conditions.[33,49,50]

The BamA POTRA domains seem to be the key player in substrate binding. Evidence exists that they form a template for the formation of β-strands in the substrate polypeptide chain.[36,37,39] Even though the sequences of POTRA 1-5 are not very similar, they share a common fold that may allow for β-augmentation.[36] Furthermore, some POTRA domains play a more specific role. POTRA 1 has been cross-linked to the periplasmic chaperone SurA and is expected to serve as an initial delivery site for the substrate.[121] Deletion of this domain affects the assembly of a subset of OMPs.[121] POTRA 2-4 are required for binding of BamB,[36,43-45] whereas POTRA 5 (P5) mediates the interaction with the essential lipoprotein BamD and thereby the BamCDE sub-complex.[36,53] P5 moreover occupies a central position closest to the β-barrel, where the substrate should be guided into either the lumen of the barrel, the interior of an oligomeric assembly or directly into the membrane.[41] It is noteworthy that P5 is the only essential POTRA domain in *Neisseria meningitidis* BamA.[182]

Recently, the crystal structures of two BamA proteins, those from *Neisseria gonorrhoeae* and *Haemophilus ducreyi*, revealed novel insights into the mechanism of β-barrel assembly.[41] The structure of the *N. gonorrhoeae* BamA β-barrel showed incomplete hydrogen bonding between the first and last β-strand, which folds back into the lumen of the barrel. Molecular Dynamics (MD) simulations suggested that the β-barrel can open entirely to form a lateral gate. This was proposed as a possible

mechanism in which a substrate OMP could be folded inside the BamA β-barrel and subsequently released. Alternatively, it might provide a way for BamA to form a template for the newly formed β-strands of the substrate to align.[41]

Furthermore, comparison between the *N. gonorrhoeae* and *H. ducreyi* BamA structures,[41] as well as the crystal structure of the paralogue TamA from *E. coli*,[147] revealed different orientations of the POTRA domains with respect to the β-barrel, suggesting that the POTRA domains could undergo large conformational changes and thereby open or close the interior of the β-barrel. Interestingly, solid-state NMR (ssNMR) studies from our lab have shown that global motion of the POTRA domains does not occur on a fast (ns) timescale in BamA proteoliposomes, but might take place on slower timescales (see Chapter 3).[157]

In the following, we show how the combination of NMR studies of soluble POTRA and membrane-embedded BamA constructs with mutagenesis and MD simulations led us to identify P5 segments exhibiting local conformational plasticity on the μs-ms timescale. This region encompasses residue E373, which was shown crucial for binding of BamD and functioning of the BAM complex *in vivo.*[118] We observed that removing the negative charge by mutation of this residue strongly perturbs the conformational equilibrium in P5, suggesting that electrostatic interactions play an important role in P5 plasticity. Interestingly, the dynamic region of P5 is highly conserved yet absent in the other POTRA domains, implying that the conformational plasticity is unique to the function of P5 and critical for β-barrel assembly.

Results

P5 domain depends on P4 to acquire its native fold

NMR is a sensitive method to probe fold and dynamics of soluble and membrane-embedded proteins (reviewed in e.g. Ref. 71,72,183,184). For example, chemical-shift correlations and NMR signal intensities can provide direct insight into protein fold as well as the presence of global motion. Previously, we have shown that the periplasmic POTRA domains of BamA do not experience fast global motion in proteoliposomes (see Chapter 3).[157] In the following we examined the influence of protein-protein and protein-lipid interactions on the fold of P5, which is located closest to the BamA β-barrel and has been shown to mediate association of the BamCDE sub-complex.[36] Firstly, we compared NMR spectra of the isolated P5 domain to results obtained on the tandem POTRA 4-5 (P4P5) construct in solution (Fig. 1a). Unexpectedly, the ^{1}H,^{15}N HSQC spectra, which serve as a fingerprint of the

protein fold, showed large differences, indicating that the fold of P5 is not preserved between both constructs. For further analysis we obtained resonance assignments of P4P5 using standard solution NMR triple-resonance experiments (see Chapter 2 for methods; BMRB deposition 19928). The resulting $C\alpha$ and $C\beta$ secondary chemical shifts indicated a secondary structure in line with previous crystallization studies using the P4P5 construct (data not shown).[38,40] In contrast, the correlation pattern of the isolated P5 domain was reminiscent of a protein that is largely unfolded. Only the central β-sheet seemed to remain partially intact, judged by P5 resonances that remain close to the P4P5 spectrum (Fig. 1a). These findings suggest that the correct fold of P5 requires stabilization by the P4P5 interface.
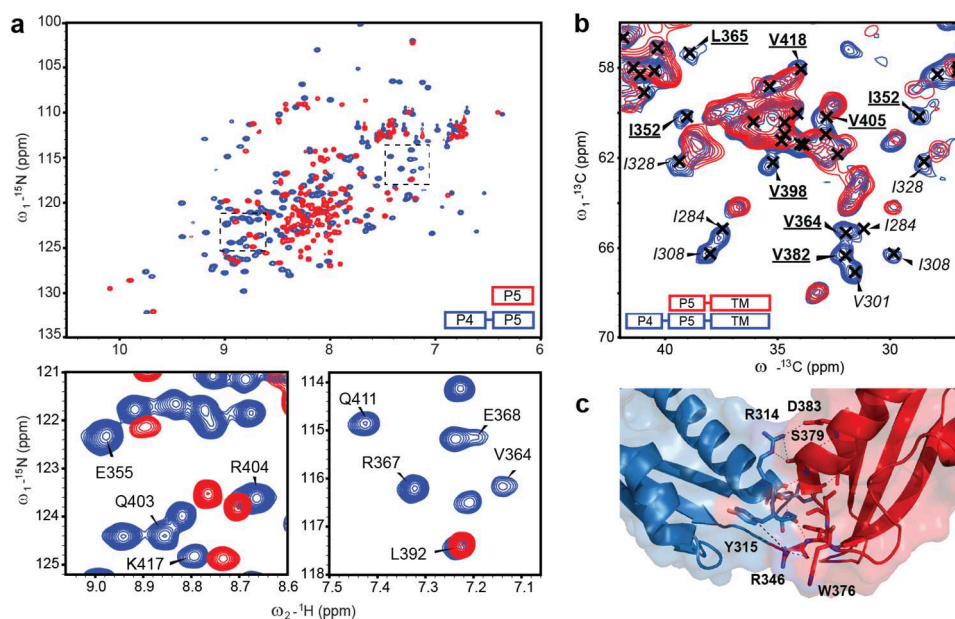

Figure 1. P5 folding critically depends on the presence of P4. a) 2D ^{1}H,^{15}N HSQC spectra of BamA P5 (red) and P4P5 (blue) in solution (upper panel). Close-ups of the boxed regions (lower panels) show assignments for P5 residues in the P4P5 construct. b) Close-up of 2D ^{13}C,^{13}C PARIS ssNMR spectra of BamA P5-TM (red) and P4P5-TM (blue) (both labeled with ^{13}C,^{15}N amino acids GSAVLTI) in dilauroyl phosphatidylcholine (DLPC) bilayers. Black crosses are assignments for P4P5 in solution. P4 peaks are labeled in italic, P5 bold and underlined. c) P4P5 crystal structure (PDB 3OG5)[38] with P4 in blue and P5 in red. Labeled residues contribute to the stabilizing interface by polar and cation-π stacking interactions (dashed lines).

Crystal structures of BamA[41] and TamA[147], which is also a member of the Omp85 protein family, have provided evidence for contacts between P5 and periplasmic loops of the transmembrane (TM) domain, which could provide another way of stabilizing the fold of P5. Hence, we investigated constructs consisting of the BamA TM domain preceded by either P5 alone or P4P5 in lipid bilayers using solid-state NMR (ssNMR). The protein variants were ^{13}C,^{15}N-labeled at the amino acids GSAVLTI, reducing spectral overlap (see Chapters 3 and 4). Signals from P5 were not detected in dipolar-based ssNMR experiments on the P5-TM construct using our solution-state NMR assignments as a reference (Fig. 1b, red). In scalar-based experiments, which select for protein segments that are mobile on the ns timescale,[90,91] protein signal was observed but the peak positions did not correspond to the chemical shift assignments that we obtained in solution (data not shown). Likely, this signal stems from unfolded or misfolded P5. Indeed, when we measured dipolar-based spectra of P4P5-TM, we could identify a series of correlations that were absent in the P5-TM construct and matched the solution NMR assignments of the isolated P4P5 construct (Fig. 1b, blue). Taken together, these findings confirmed that P5 folding critically depends on the interface with P4, consisting mainly of polar interactions and cation-π interactions (Fig. 1c), even when associated with the membrane-embedded BamA part and that the domains as a whole are relatively rigid in this sample preparation, as we observed in previous work.[157]

A set of P5 residues experiences conformational exchange

Previous NMR and SAXS data suggested a well-defined orientation between the two domains and thus a common correlation time for the P4P5 construct in solution.[38] Yet, our solution NMR studies on the P4P5 tandem construct revealed that, on average, P5 signals were broader than those of P4, with some regions even broadened beyond detection (Fig. 2a). To examine whether line broadening is induced by local dynamics within the P5 domain, we performed CPMG experiments that probe conformational exchange on the µs to ms timescale (see e.g. Ref. 109,185). Indeed, several residues on P5 exhibited a contribution of exchange to the relaxation rate (Fig. 2b), suggesting that conformational exchange was the also the cause of the extensive line broadening of residues for which assignments could not be obtained. When plotted on the 3D structure, these residues mostly cluster to the α-helices and connecting loop (Fig. 2c,d) while on P4, only some residues that are involved in the inter-domain interface, e.g. A318, showed broadening in the ^{1}H,^{15}N HSQC spectrum and conformational exchange in CMPG experiments (Fig. 2d).
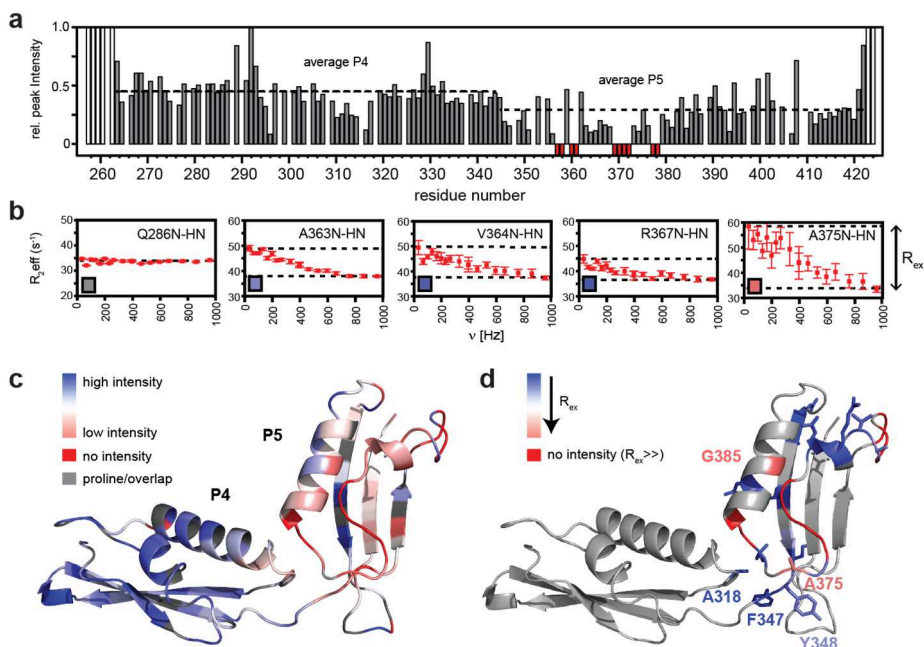
Figure 2. BamA P5 experiences conformational exchange in solution. a) Peak intensities from the 2D $^1H,^{15}N$ HSQC of BamA P4P5 plotted along the amino acid sequence. Negative red bars represent residues that could not be assigned due to lacking intensities. b) Relaxation dispersion profiles for several residues of P4P5. R_{ex} is indicative of the contribution of conformational exchange to the relaxation rate. Boxes in the left bottom corners indicate the extent of R_{ex} as in d) and no R_{ex} in grey. c) HSQC intensities as in a) plotted on the crystal structure of P4P5 (PDB 3Q6B). d) R_{ex} plotted on the crystal structure of P4P5. Residues experiencing detectable conformational exchange are shown in sticks.

Next, we investigated to what extent the observed local dynamics in P5 might be modulated by the presence of the BamA β-barrel and the lipid bilayer. To this end, we analyzed 2D ssNMR spectra of the BamA P4P5-TM construct carrying $^{13}C,^{15}N$-labeling at the amino acids GSAVLTI, allowing us to track several resonances of the POTRA domains and of the β-barrel. For our spectral analysis, we used *de novo* ssNMR resonance assignments of the BamA TM domain (see Chapter 4) in addition to our solution NMR assignments for the P4P5 domains. To distinguish between static and motional disorder, we compared data obtained at -2 °C and 20 °C (Fig. 3a).

First, we examined signal intensities of P4 and P5 domains in a 2D NCA spectrum at the lower temperature. Here, we observed that residues I284, V301 and T302 from P4 had similar intensities to L630 of the β-barrel, whereas isolated P5 residues, including A375 and S408, exhibited lower intensity (Fig. 3b). On the other hand, peak integration showed that residue S408 from P5 had the same integral

value as L630 from the β-barrel (Fig. 3c), indicating static conformational heterogeneity that leads to peak broadening. In contrast, A375 exhibited both lower intensity and a smaller peak integral, which would be consistent with residual motion that partially averages out the ^{15}N-^{13}C dipolar coupling. Note that at -2 °C, the sample was not yet frozen as judged by one-dimensional ^{1}H NMR (data not shown). We next recorded a 2D NCA spectrum at 20 °C to probe dynamic effects that increase as a function of temperature. From P5, A375 signal intensity decreased further, but S408 was not affected by the increase in temperature. These findings relate very well to our results from solution NMR: A375 is located in the region of conformational exchange, which is expected to become faster at higher temperature. Furthermore, relative to L630 from the β-barrel, signal intensities relating to residues from P4 (namely I284, V301 and T302) now also diminished (Fig. 3b). Again, this matches our solution NMR data where the P4P5 interface was found to be dynamic (e.g. A318, F347, Y348; Fig. 2d), causing the P4 domain to become more mobile relative to the membrane-embedded β-barrel in ssNMR (Fig. 3d). Note that S408 is not affected by the increase in temperature, indicating that the domain orientation of P5 with respect to the β-barrel remains stable.
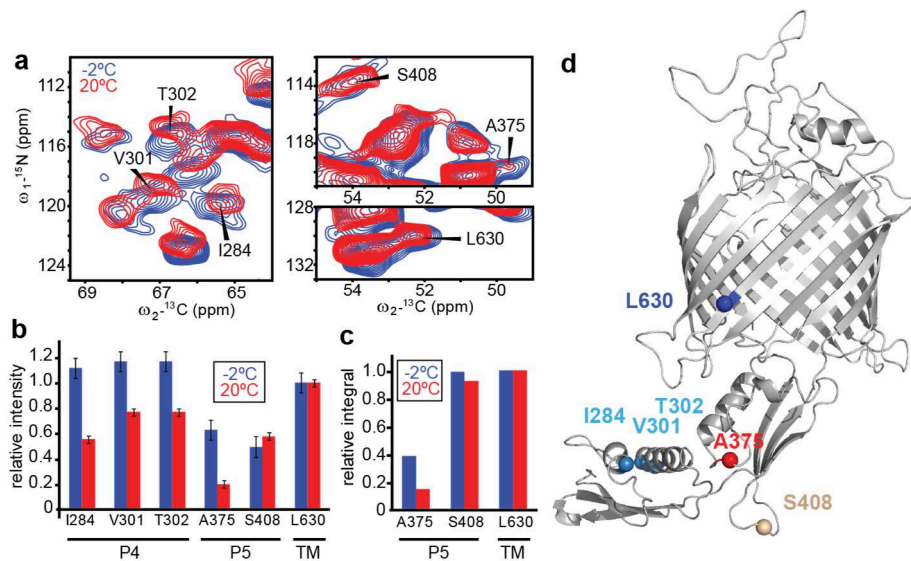

Figure 3. Local dynamics of the POTRA domains in membrane-embedded BamA. a) Extracted regions from 2D NCA experiments on ^{13}C,^{15}N-GSAVLTI labeled BamA P4P5-TM in DLPC lipid bilayers, recorded at -2 °C (blue) and 20 °C (red). b) Signal intensities from the spectra shown in a). Intensities are shown relative to those of residue L630 for both temperatures. Error bars represent the noise levels of the two experiments relative to L630 intensity. c) Peak integrals of selected residues from the experiments shown in a), relative to L630. d) BamA P4P5-TM *E. coli* homology model[41] with the residues probed in ssNMR shown as spheres, color-coded according to the dynamic behavior: least dynamic (blue); mobile only at higher temperatures (light blue); statically disordered (beige), most dynamic (red).

Mutation of E373 affects conformational equilibrium of P5

Previous *in vivo* studies have demonstrated that mutation of residue E373 on BamA P5 to lysine abolishes the interaction with BamD and impairs β-barrel assembly at normal growth temperature,[118] whereas mutation to alanine has milder effects.[69] We examined the effect of E373A and E373K mutations firstly on the P4P5 conformation in solution (Fig. 4a). The 2D ^{1}H,^{15}N HSQC spectra revealed significant ^{1}H-^{15}N chemical-shift perturbations (CSPs) in both cases, the charge-reverting E373K mutant causing the strongest effect (Fig. 4a,d). Interestingly, CSPs occurred not only for residues in close vicinity of the mutation, but all throughout the P5 domain, whereas P4 remained largely unaffected apart from residues on the interface between the two domains. Also arginine side chains from P5 and the interface showed significant CSPs (data not shown), possibly as a consequence of an altered electrostatic network (see below). Long-range effects due to a point mutation might be caused by a change in structure, yet we found very little changes in the C_α and C_β chemical shift values of the E373K mutant protein (data not shown), indicating that the P4P5 fold, a least on the level of secondary structure, remains largely preserved compared to the wild-type (WT). However, the intensities for P5 in the ^{1}H,^{15}N HSQC spectrum of the E373K mutant dropped further (data not shown), suggesting that the mutation enhances the conformational exchange.

We sought to confirm this hypothesis in the membrane-embedded BamA construct using ssNMR, where the increase in chemical exchange should result in further scaling of the dipolar couplings, reducing signal intensity. Comparison of WT and E373K mutant spectra revealed complete disappearance of a set of P5 resonances such as A363, A375 and V412 in the mutant spectra (Fig. 4b,c), whereas signal-to-noise for other correlations was similar. Note that resolved P4 resonances such as A278 (Fig. 4c) were retained in the dipolar-based ssNMR spectra of the mutant, implying that our results point to increased local dynamics within P5, whereas global motion of the POTRA domains with respect to the β-barrel is still restricted.
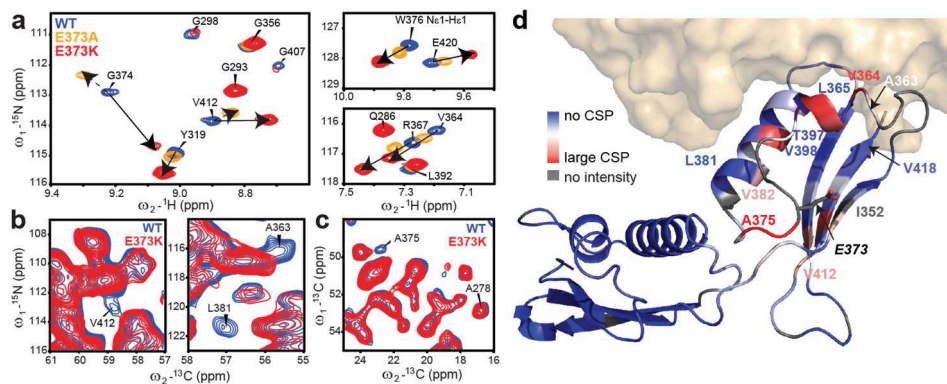
Figure 4. Mutation of E373 induces long-range perturbations and increased dynamics in P5. a) Close-ups of ^{1}H,^{15}N HSQC spectra of BamA P4P5 WT (blue), E373A (orange) and E373K (red). b,c) Details from b) 2D NCA and c) 2D ^{13}C,^{13}C ssNMR spectra of BamA P4P5-TM WT (blue) and E373K (red) labeled ^{13}C,^{15}N-GSAVLTI, highlighting P5 residues that have disappeared upon mutation and a P4 residue that remains. d) CSPs of the P4P5 E373K mutant in solution, color-coded on the crystal structure (PDB 3Q6B) with the BamA β-barrel homology model shown in surface representation.[41] Labels correspond to the residues that are not detectable in ssNMR spectra of the E373K mutant due to increased dynamics. Residue E373 is shown with the side chain as sticks and labeled in italic.

MD simulations support dynamic electrostatic network in P5

Since mutation of E373 into a positively charged residue affects the dynamic behavior of P5, we hypothesized that interchanging electrostatic interactions might play a role in the conformational exchange of P5 that we observed by NMR. For this purpose, we performed atomistic molecular dynamics (MD) simulations on BamA P4P5-TM in a DMPC bilayer and monitored the distance between the charged residues of P5 that are localized in the dynamic region (Fig. 5a). Interestingly, we observed local dynamics of electrostatic contacts around the mutation site E373 involving charged side chains of residues K351, D362 and R366 (Fig. 5b). These residues associated and dissociated to alternating pairs of salt-bridges on the sub-μs timescale, which may give rise to the μs-ms conformational exchange observed in our NMR experiments. Consistent with the results from our MD simulations, E373 and D362 were not observed in the ^{1}H,^{15}N HSQC spectrum of P4P5 due to extensive line broadening, whereas K351 and R366 belonged to the weakest signals (Fig. 2a).
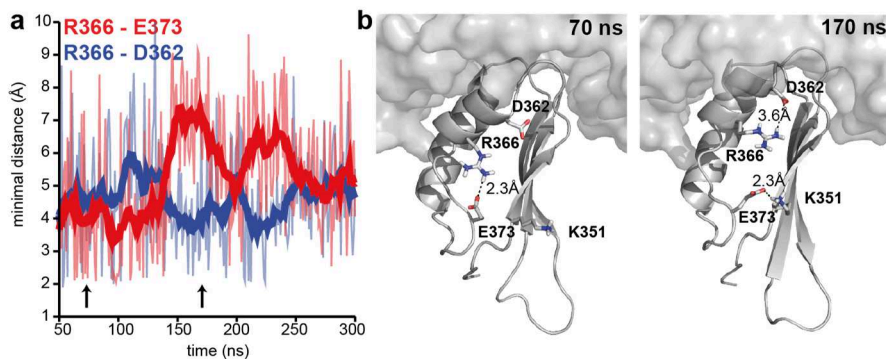
Figure 5. MD simulations of BamA P4P5-TM in a DMPC bilayer reveal a dynamic electrostatic network in P5. a) Plot of the minimal distances between the side chain carboxyl group of D362/E373 and the guanidine group of R366. The thick lines show running averages over 20 ns. b) Snapshots from the trajectory at 70 ns (left) and 170 ns (right) indicated by arrows in a), showing different electrostatic interactions. P4 is omitted for clarity.

Discussion

Folding and insertion of OMPs by the BAM complex is crucial for Gram-negative bacteria. Since no ATP is present in the periplasm, this process is thought to be driven by the gain in free energy of the folded membrane protein.[186] So far, the role of the BAM complex in mediating this process remains poorly understood, but dynamics and conformational changes within the machinery are likely to play an important role.

Using a combination of NMR, mutagenesis and MD simulations, we have shown that POTRA 5 (P5) exhibits local conformational plasticity that is present also in the context of membrane-embedded BamA. We found that P5 in isolation or attached to the β-barrel is not correctly folded, most likely due to the lack of a stabilizing interface with P4. This notion is further supported by our observation of conformational exchange processes in P5 that propagate to the P4P5 interface in the tandem construct and membrane-embedded BamA. It is interesting to note that BamA with only P5 is functional in *N. meningitidis*,[182] suggesting that the conformational equilibrium can vary among different species. Mutation of residue E373 seems to alter the conformational equilibrium of P5. The E373A mutation has only mild effects *in vivo*, leading to a weakened interaction with BamD and increased detergent-sensitivity,[69] yet in our NMR spectra a modulation on a molecular level is apparent from CSPs that extend to residues far away from the mutation. The E373K

mutant, which has a more severe phenotype *in vivo*[118] – it completely abrogates the BamD interaction and does not grow at temperatures higher than 24 °C – showed even larger CSPs. We observed that in addition to backbone amides, arginine side chains underwent large CSPs upon mutation of E373, not only in P5 but also extending to the P4P5 interface. Furthermore, the E373K mutant displayed strongly reduced signal intensities both in solution and ssNMR, implying increased dynamics. Interestingly, the residues further away from the site of mutation, such as Y319 and V364, shifted in one direction from WT to E373A to E373K (Fig. 4a), suggesting that the conformational equilibrium is driven towards one side. Correspondingly, a shift in conformational equilibrium, rather than loss of a direct contact between E373 and BamD, may correlate to the increasingly weakened BamA - BamD interaction *in vivo*.

Taken together, these findings suggest that electrostatic interactions play an important role in the conformational equilibrium of P5. In the highest resolution crystal structure of P4P5 (PDB 3Q6B), a network of electrostatic interactions exists between D362, R366, E373 and K351,[40] whereas in another crystal form (PDB 3OG5) R366 and E373 form a salt bridge, leaving D362 and K351 more exposed.[38] The B-factors of these structures do not provide any indication of dynamics, yet charged residues including K351 and E373 are involved in stabilizing crystal contacts in both cases.[40] Moreover, comparison of the two structures demonstrates that different configurations are possible. Our MD simulations indicated that the contacts between the charged residues are interchanging, which might be underlying the µs-ms conformational exchange observed in NMR. Remarkably, none of the other POTRA domains possesses such a network of charged residues in this area. Instead, they expose a hydrophobic patch in the corresponding region of the general POTRA fold (Fig. 6a, white) that is thought to allow for β-augmentation of the substrate.[36,37,39] Thus, the dynamic region that we identified is unique to the specific function of P5, and it forms a highly conserved patch on the surface of BamA (Fig. 6b, left).

Altogether, a picture emerges in which conformational flexibility occurs in all stages of β-barrel assembly. Before the substrate reaches BamA, it is bound by periplasmic chaperones that prevent aggregation and keep the substrate folding-competent (Fig. 7, 1). The complex of Skp with unfolded OMPs was thoroughly investigated before and showed a dynamic ensemble that allows for efficient release of the substrate towards the BAM complex.[14] For the other main periplasmic chaperone SurA, such studies have not yet been carried out but similar principles are expected to play a role. Next, the unfolded substrate is likely to undergo non-specific, transient interactions with the POTRA domains 1-4 by means of β-augmentation on the hydrophobic sides of these POTRA domains (Fig. 7, 2).
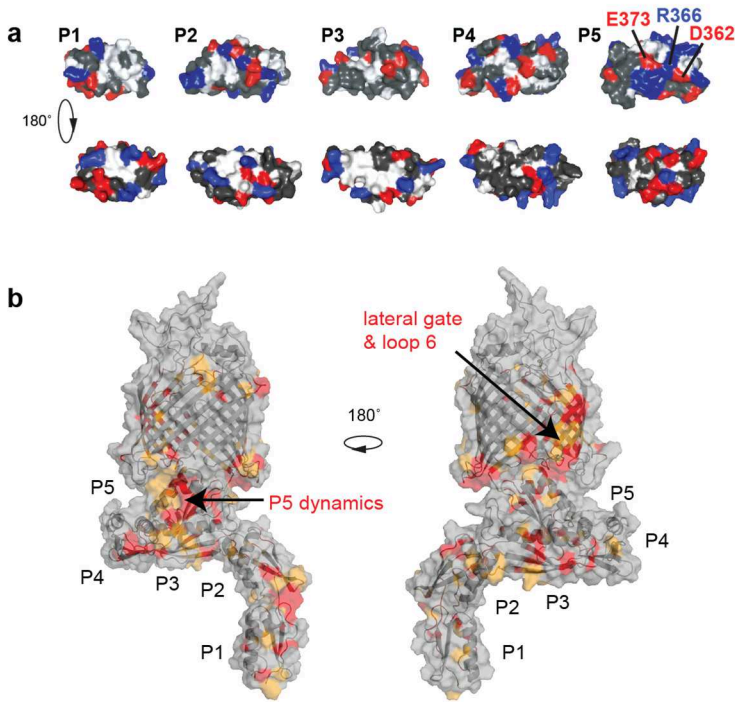
Figure 6. Conserved conformational plasticity of BamA P5. a) The sides of all POTRA domains except P5 expose a hydrophobic patch. The top row shows the side of the electrostatic network of P5 involving D362, R366 and E373. All domains are shown in the same orientation aligned on P2. White: hydrophobic, blue: positively charged, red: negatively charged. b) Sequence conservation of BamA analyzed with Consurf.[115] Conserved residues are colored on the homology model[41] in red and orange for the most and second most conserved category of residues, respectively.

Interestingly, although the angle between the P4 and P5 domains was conferred rigid from SAXS and RDC measurements,[38] we identified subtle dynamics on the interface that might be important to guide unfolded substrates to the β-barrel. Again the functional relevance is supported by conservation analysis (Fig. 6b) that shows high conservation on the interfaces of not only P4 and P5 but of all neighboring POTRA domains.

The BamA P5 domain, subsequently, is the critical hub for β-barrel assembly, because it has to interact with substrate OMPs as well as the essential lipoprotein BamD.[36,49] Our data on the E373 mutants that weaken binding of BamD *in vivo*[69,118] provide evidence for a role of P5 plasticity in this interaction. Furthermore, we can speculate that P5 is involved in recognition of the C-terminal motif of substrate OMPs,[65] transiently accommodating the negatively charged C-terminus within its dynamic electrostatic network. Interestingly, also BamD has been implied in binding

of the C-terminal motif[53,54] and further studies will have to elucidate the precise interplay between P5, BamD and the substrate (Fig. 7, 3).

Protein plasticity again plays a role in the final step of insertion into the outer membrane (Fig. 7, 4). Recently solved BamA crystal structures suggested that the β-barrel could open laterally.[41] The first and last β-strand in the crystal structure are again highly conserved, as well as extracellular loop 6 and residues from the β-barrel that coordinate it (Fig. 6b, right). In addition, interplay between BamA and the lipid bilayer is thought to locally perturb the membrane,[41,157] potentially stimulating insertion of the substrate.

In conclusion, we identified conformational plasticity in the BamA P5 domain that is compatible with a model for OMP assembly in which local dynamics are involved in substrate recognition, targeting to the membrane and insertion. Future interaction studies of BamA with substrate and the accessory lipoproteins will be crucial to obtain a picture of the β-barrel assembly machinery in action.
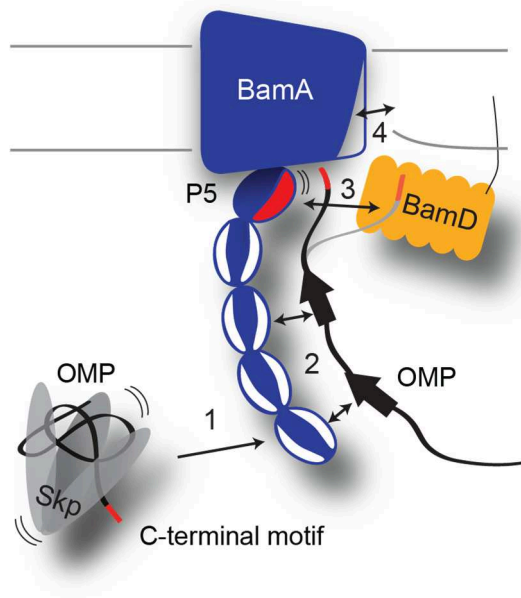

Figure 7. Schematic view of conformational plasticity in the process of β-barrel assembly. 1: Periplasmic chaperones form a dynamic complex with the unfolded OMP. 2: The unfolded OMP transiently interacts with hydrophobic patches on POTRA domains 1-4 (white). 3: Interactions occur between the substrate (possibly the C-terminal recognition motif, shown in red), P5 (dynamic region shown in red), and BamD. 4: The substrate is inserted by a dynamic interplay between the BamA β-barrel, which might open up laterally, and the lipids of the outer membrane.

CHAPTER 6

Interaction between the essential components of the β-barrel assembly machinery

Tessa Sinnige, Iva Pritisanac, Klaartje Houben, João Rodrigues, Jan Grijpstra, Rolf Boelens, Alexandre M.J.J. Bonvin, Jan Tommassen, Marc Baldus

Abstract

Folding and insertion of outer membrane proteins (OMPs) in Gram-negative bacteria is mediated by the β-barrel assembly machinery (BAM). In *Escherichia coli*, the BAM complex consists of five proteins, yet only two of them are essential: the integral membrane protein BamA and the lipoprotein BamD. It is not known what their individual roles in the process of β-barrel assembly are, nor how they interact. We used solution and solid-state NMR spectroscopy in combination with protein-protein docking to dissect the BamA-BamD interaction. We found that BamD interacted weakly with POTRA domains of BamA in solution and in lipid bilayers. However, when we co-reconstituted BamA with the three lipoproteins BamCDE, a stable complex was formed. Our results suggest that complex formation not only involves the POTRA 5 domain, but also the periplasmic loops from BamA. Protein-protein docking revealed possible modes of interaction between BamA and BamD, in which electrostatic interactions play a dominant role. In addition, a highly conserved patch of residues on the C-terminal side of BamD that has been implied in the BamA-D interaction remained accessible in the observed complexes, suggesting another functional role such as binding of the unfolded substrate.

Introduction

The β-barrel assembly machinery (BAM) that mediates folding and insertion of outer membrane proteins (OMPs) consists of five proteins in *Escherichia coli*: the integral membrane protein BamA and the lipoproteins BamBCDE.[33,49,50] Only BamA and BamD are essential,[33,51] whereas individual deletions of the other lipoproteins are tolerated, although they cause mild defects in OMP assembly and reduced antibiotic- and detergent-resistance.[33,49,50] BamA is highly conserved among Gram-negative bacteria and functional knock-outs have not been described, underlining its importance.[31,32] Homologous proteins exist even in mitochondria[34,187,188] and more distantly in chloroplasts.[189] The lipoproteins, on the other hand, are less conserved. However, homologues of BamD in Gram-negative bacteria other than *E. coli* have been identified, e.g. ComL in *Neisseria* species, although these are not always essential.[190,191]

In *E. coli*, BamA and BamD have been shown to interact via the POTRA 5 (P5) domain of BamA.[36] The mutation E373K in P5 was found to abrogate the interaction with BamD, suggesting that this residue is directly or indirectly involved in binding (see also Chapter 5).[118] The binding site on BamD seems to be located towards the C-terminal side of the protein, since replacement of 19 residues from the C-terminal end with a non-native amino acid tail resulted in loss of its interaction with BamA.[50]

In the *E. coli* BAM complex, BamD furthermore interacts with BamC and BamE, thereby linking the BamCDE sub-complex to the main, membrane-embedded component BamA.[49,50] BamA-CDE form a stable complex independent of BamB, as demonstrated both *in vivo*[49,50] and *in vitro*.[21] BamC and BamE are not essential, but they seem to stabilize the interactions within the BAM complex,[50] which might be their primary function. The interaction between BamD and BamC was elucidated from a co-crystal structure, which showed an extended interface mediated mostly by the unstructured N-terminal extension of BamC.[60] However, it remains unclear how BamD precisely interacts with BamA and to what extent the non-essential lipoproteins BamC and BamE contribute.

Moreover, the function of BamD and the reason for its essentiality are poorly understood, although the crystal structure has provided some clues.[53] BamD has a fold consisting of five TPRs (tetratricopeptide repeats) that are generally implicated in protein-protein interactions.[55] Other TPR proteins such as HOP[57] and PEX5[56] have been shown to bind to extended polypeptides, suggesting that also BamD might be involved in binding of unfolded OMPs. Interestingly, in the BamD crystal structure, the molecules are binding the C-terminal His-tag of neighboring molecules between TPRs 1 and 2, suggesting that this is a possible binding site.[53]

Thus, BamD might have a role in recognizing the C-terminus of unfolded OMPs. However, the question remains how BamD cooperates with BamA in the process of β-barrel assembly. Genetic data have suggested that BamD serves to "activate" BamA as part of a conformational cycle that BamA undergoes to fold and insert substrates,[118] but a structural basis that supports such a mechanism does not exist so far. Therefore, it is of great importance to characterize the interaction between BamA and BamD on a molecular level, which we accomplished using NMR and protein-protein docking as described in this chapter.

Results

POTRA 5 and BamD interact weakly in solution

Because the P5 domain of BamA is essential for its interaction with BamD,[36] we initially set out to find the binding site on this domain using solution NMR. Since P5 is not correctly folded without P4 (see Chapter 5), we used the POTRA 4-5 (P4P5) tandem construct. Surprisingly, although we expected the two protein segments to form a complex with 1:1 stoichiometry, solution NMR titrations of ^{15}N-labeled P4P5 with unlabeled BamD (Fig. 1a,b) yielded only very small chemical-shift perturbations (CSPs) that were not always dose-dependent (Fig. 1c). Relatively little line broadening was observed (Fig. 1d) even at higher molar ratios of BamD:P4P5. Interestingly, the residues that were slightly affected largely corresponded to the P5 residues that are intrinsically broadened by conformational exchange (see Chapter 5) and therefore difficult to observe, e.g. I352, R366, G374, A375, and F394. If these residues were to be involved in a weak interaction with BamD, a combination of exchange within the P5 domain and that between the free and BamD-bound state would be the result, which is hard to distinguish. However, formation of a tight complex between P4P5 and BamD under the conditions used in this experiment could be excluded.

To circumvent the inherent conformational exchange of the aforementioned P5 residues, we performed the reverse titration experiment in which we added unlabeled P4P5 to ^{15}N-labeled BamD (Fig. 2a). Again, certain signals seemed to undergo small CSPs as well as line broadening (Fig. 2b). Strikingly, these were also among the weaker signals in the spectrum of BamD alone, suggesting that the binding sites on P5 and BamD are both intrinsically flexible.
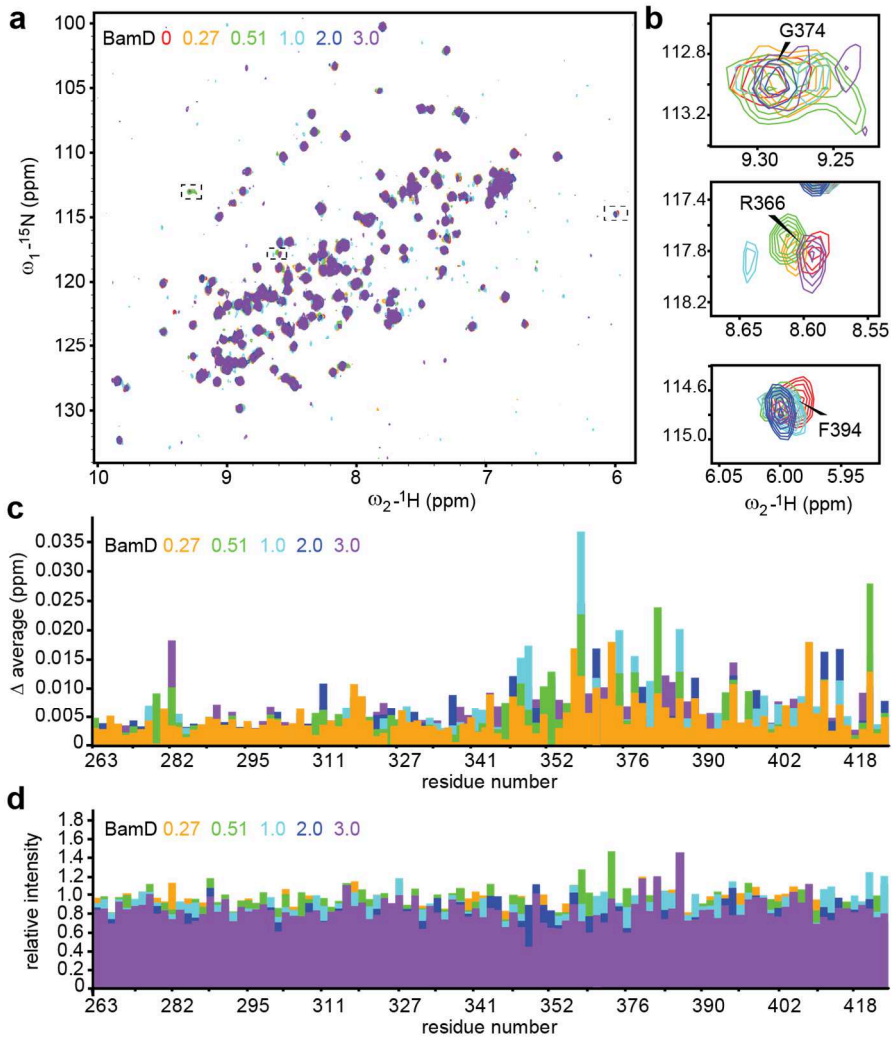
Figure 1. Titration of P4P5 with BamD shows weak interaction. a) ^{1}H,^{15}N-HSQC spectrum of ^{15}N,^{13}C BamA P4P5 titrated with different molar ratios of BamD. b) Close-ups of the spectra indicated by dashed boxes in a), showing small CSPs and line broadening for certain residues. c) Plot of combined ^{1}H,^{15}N CSP for different ratios of BamD along the sequence of BamA P4P5, calculated using Δaverage$=\sqrt{\left(\left(\Delta_{1H}\right)^2+\frac{\left(\Delta_{15N}\right)^2}{25}\right)/2}$. d) Plot of relative signal intensity for different ratios of BamD along the sequence of BamA P4P5, scaled to the spectrum of P4P5 alone.
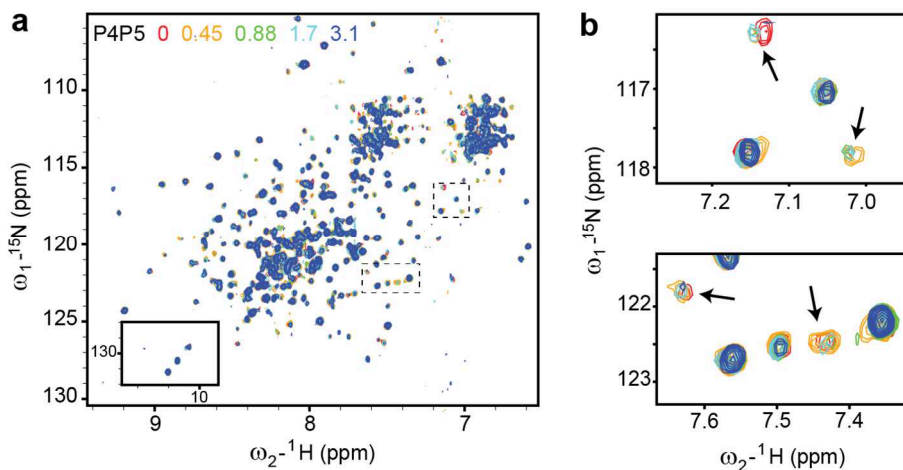
Figure 2. Titration of BamD with P4P5 shows signs of weak interaction. a) ^{1}H,^{15}N-TROSY-HSQC spectrum of ^{2}H,^{15}N,^{13}C BamD titrated with different molar ratios of BamA P4P5. b) Close-ups of the spectra indicated by dashed boxes in a), showing small CSPs and line broadening for some residues, whereas others are unaffected.

The presence of the other lipoproteins BamC and BamE could modulate the binding affinity of BamD for P5. In fact, the crystal structures of BamD in isolation and in complex with a fragment of BamC, consisting of the N-terminal domain plus an unstructured extension of BamC (designated BamC^{UN}) slightly differ. In the complex BamD-BamC^{UN}, a slight conformational change occurs in the orientation of BamD's TPR repeats with respect to its structure in isolation (Fig. 3a),[60] which might affect its interaction with BamA. Therefore, we decided to titrate P4P5 with this complex. We expected that with 47 kDa, binding of BamC^{UN}D to P4P5 would cause considerable line broadening. On the contrary, at molar ratios of up to 1:0.6 P4P5:BamC^{UN}D, no clear CSPs were detected and only the already broadened signals such as I352 and A375 seemed possibly affected (Fig. 3b). We concluded that BamC does not enhance the interaction of BamD with P4P5 in solution.
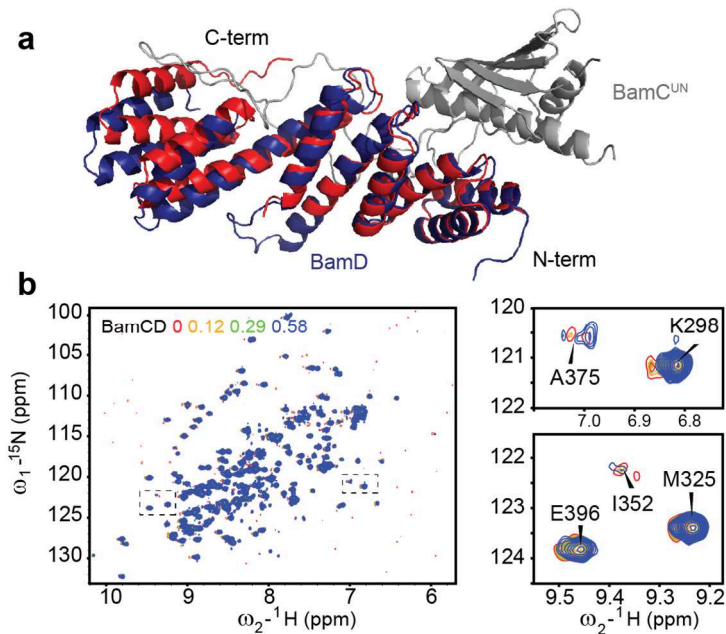
Figure 3. BamC does not affect the P4P5-BamD interaction in solution. a) Crystal structures of BamD in isolation (red, PDB 2YHC) and the complex consisting of BamD (blue) and BamC^UN (grey) containing the N-terminal domain and unfolded extension (PDB 3TGO), overlaid on the three N-terminal TPR repeats of BamD. b) ^{1}H,^{15}N-HSQC spectrum of ^{15}N,^{13}C BamA P4P5 titrated with different molar ratios of the BamCD sub-complex, consisting of BamD and BamC^UN. Boxed regions are shown in close-up on the right.

Association of BamA and BamD in lipid bilayers

Apart from protein-protein interactions, the membrane environment may well have an effect on the BamA-BamD interaction. Firstly, the BamA transmembrane (TM) domain might contribute to the binding interface with its periplasmic loops. Secondly, tethering to the lipid bilayer of both P5 attached to the BamA β-barrel and BamD with its lipid anchor pre-orients the interaction partners and increases their local concentration, which may lead to stronger binding. To investigate these phenomena, we co-reconstituted the BamA construct P4P5-TM with different ratios of lipidated BamD in DLPC lipid bilayers and recorded solid-state NMR (ssNMR) spectra on these samples. For analysis of the spectra, we made use of the solution NMR assignments that we obtained for the P4P5 construct, and which overlay well with the ssNMR spectra of membrane-embedded BamA constructs (see Chapter 5). Overall, the spectra of BamA P4P5-TM with and without BamD were similar, but a number of resonances from the P5 domain broadened significantly, suggesting

binding with intermediate exchange on the NMR timescale (Fig. 4a,b). These residues included I352, A375, S408 (C_β-C_α correlation, but not C_α-C_β) and R366/R388, which all cluster to one side of the P5 domain except for R388, which points towards the β-barrel (Fig. 4c). R366 and R388 have overlapping chemical shifts as assigned in solution NMR, so we could not exclude that R388 shifted or broadened upon binding of BamD. More likely, it has a different chemical shift in the membrane-embedded BamA construct due to its proximity to the BamA β-barrel.
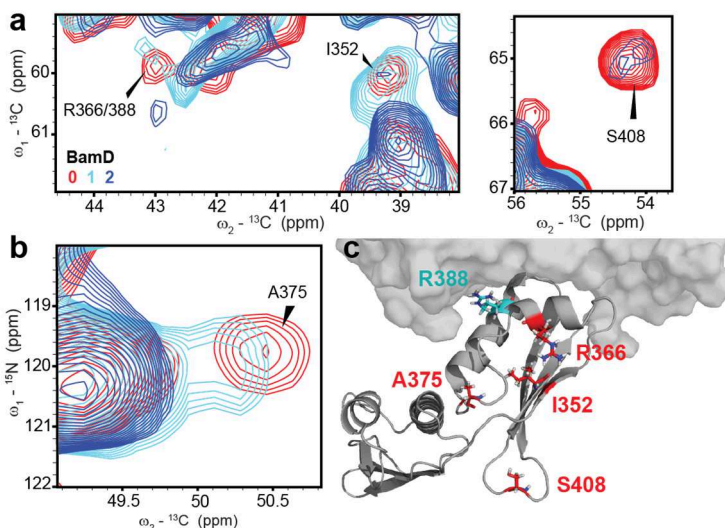

Figure 4. Close-ups of a) 2D ^{13}C,^{13}C PARIS experiments with 30 ms mixing time and b) 2D NCA experiments on U-^{13}C,^{15}N BamA P4P5-TM alone (red) and co-reconstituted with unlabeled BamD at a molar ratio of 1:1 (cyan) and 1:2 (blue) in DLPC lipid bilayers. c) Residues affected by the presence of BamD highlighted in red on the homology model of *E. coli* BamA P4P5-TM[41] with the TM domain shown in surface representation. R388 (cyan) is unlikely to be part of the BamD binding site as it is positioned towards the BamA β-barrel.

BamA and BamCDE form a stable complex in lipid bilayers

Considering that neither addition of BamC, nor anchoring to the lipid bilayer provided a stable interaction between BamA and BamD, we set out to express and purify the lipid-anchored BamCDE sub-complex. After co-reconstitution of BamA P4P5-TM (containing ^{13}C,^{15}N labeling at amino acids GSAVLTI, see Chapter 3 and 4) with unlabeled BamCDE in lipid bilayers, we observed clear CSPs of certain P5 residues in ssNMR experiments (Fig. 5), namely I352 and A375, consistent with the previous experiments (Fig. 4), as well as V364, L365 and V418 (Fig. 5a,b,d). The majority of these P5 residues belong to the ones that were found to undergo

conformational exchange in isolated BamA constructs (see Chapter 5), yet complex formation with BamCDE seemed to reduce the local dynamics. For example, A375 not only shifted, but also became more intense. Signals from several other residues that did not shift and hence are probably not directly involved in binding, including I328, S408 and V412, also became stronger (Fig. 5a), pointing towards a general stabilizing effect of the interaction with BamCDE. Note that all lipoproteins were properly incorporated in the proteoliposomes together with BamA P4P5-TM and only a small fraction of each protein was observed in the supernatant after reconstitution (Fig. 5c).
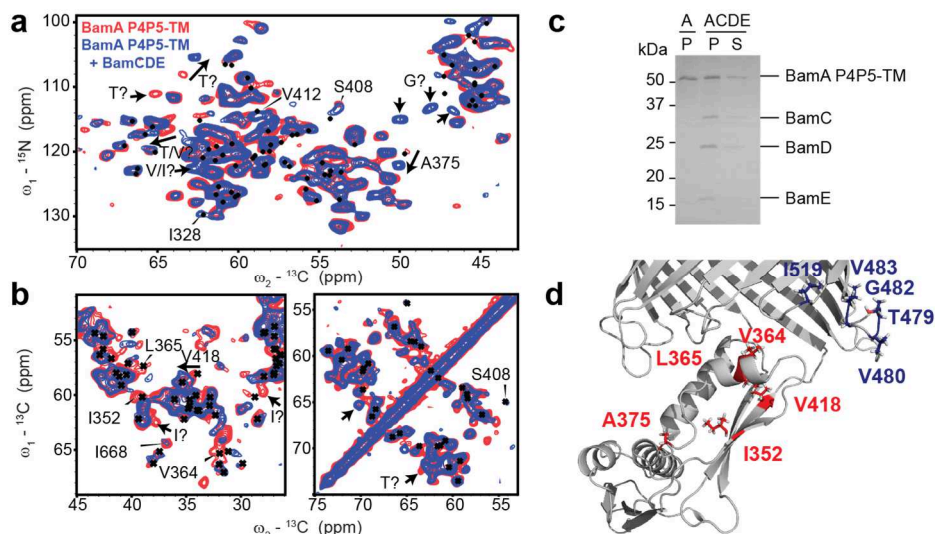

Figure 5. BamA and BamCDE form a stable complex in lipid bilayers. a) 2D NCA and b) 2D ^{13}C,^{13}C PARIS with 30 ms mixing time of BamA P4P5-TM labeled with residues GSAVLTI ^{13}C,^{15}N reconstituted in DLPC bilayers at a LPR of 50:1 in isolation (red) and co-reconstituted with BamCDE (blue). Spectra were recorded at 4 °C. Black dots indicate solution NMR assignments for P4P5. Most likely residue type based on average chemical shifts for the TM domain are labeled with question marks. c) SDS PAGE of the proteoliposomes of BamA P4P5-TM alone (A) and together with BamCDE (ACDE). P, proteoliposome pellet; S, supernatant. d) P5 residues that show CSPs in complex with BamCDE shown as red sticks on the homology model of *E. coli* BamA P4P5-TM. Residues from pL2 and pL3 that were isotope labeled in these experiments are shown in blue.

Furthermore, several correlations not belonging to P4P5, hence comprising the BamA TM domain, changed upon co-reconstitution with BamCDE (Fig. 5a,b). Since ssNMR assignments for these correlations are not yet available (see Chapter 4), it was not certain to which residues they corresponded, but probable residue types could be deduced from average BMRB chemical shifts[192] (Fig. 5a,b, labeled with

question marks). The periplasmic loops pL2 and pL3, which are closest to the binding site on P5 in the BamA structure, likely contribute to the interaction with BamD and contain residue types (Fig. 5d, blue) that correlate with the ones identified experimentally (Fig. 5a,b).

Docking reveals possible modes of interaction between BamA and BamD

Using our NMR data as input, we next performed protein-protein docking with HADDOCK[116] to obtain a model for the BamA-BamD complex. For BamD, no residue-specific information was available, so we searched for conserved patches on its surface, which often play a role in protein-protein interactions.[193] Notably, the binding site that we identified on BamA P5 is highly conserved (see Chapter 5), as well as residues from pL2 (T479, D481, G482) and pL3 (P518) (Fig 6a). On the surface of BamD, several conserved regions were found (Fig. 6b), but the binding site for BamC could be excluded. Furthermore, the C-terminal region of BamD is known to be required for its interaction with BamA.[50] Taking this into account, a likely binding site for BamA was identified (Fig. 6b, red). The residues comprising this patch (A175- R197) were defined as passive residues in the docking procedure. The BamD structure to be used in HADDOCK was taken from the co-crystal structure of the complex with BamC (PDB 3TGO) to account for the conformational change that BamC induces.
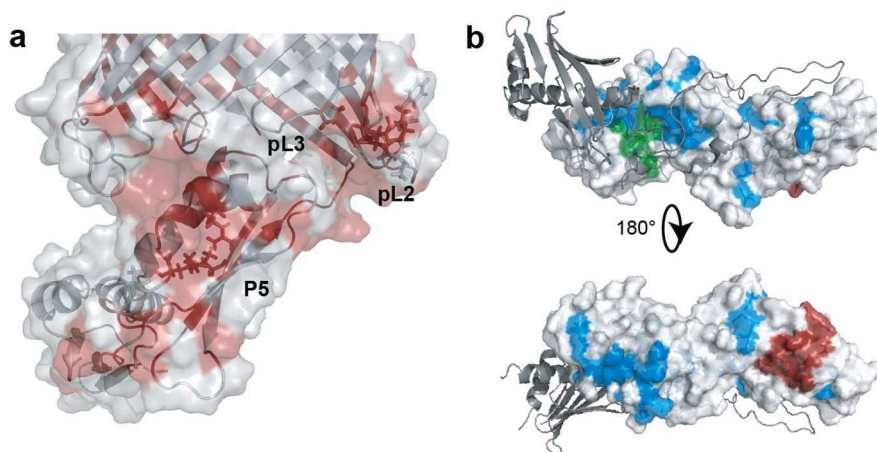
Figure 6. Conservation on the surfaces of BamA and BamD. a) Two most conserved groups of BamA residues according to Consurf[115] shown in red on the *E. coli* homology model of BamA P4P5-TM. Residues used as active in HADDOCK are shown as sticks. B) Two most conserved residue groups determined by Consurf[115] colored on the surface of BamD in complex with BamC (grey cartoon) (PDB 3TGO). Green: the postulated binding site for the C-terminus of OMPS; red: proposed interaction site for BamA; blue: other conserved residues.

First, we performed docking using only the P4P5 domains of BamA (PDB 3Q6B) and BamD. We carried out various HADDOCK runs based on the solution NMR or the ssNMR data. In both cases we also included the conserved residue E373 that is known to be crucial for correct β-barrel assembly. From the solution NMR data, we used the solvent-exposed residues that experienced significant CSPs (Y348, I352, A363, R366, E368, A375, F394; see methods in Chapter 2) as active residues to drive the docking. Random removal of 50% of the restraints was applied. From ssNMR, we combined the data from co-reconstitution with BamD (I352, A375, R366) and BamCDE (V364 in addition). L365 and V418 were not included as they are not surface-exposed. Note that the selected residues were largely a subset of the solution NMR dataset. Passive residues were automatically defined as any surface-accessible residue within a radius of 6.5 Å from the active residues. Because of the small number of ssNMR restraints, random removal of restraints was either limited to 33% or completely turned off in HADDOCK.

Irrespective of whether solution or ssNMR data were used, the acceptable solutions converged to two possible models that will hereafter be referred to as models A and B (Fig. 7a, A: light blue, B: dark blue). Both were characterized by extensive electrostatic contacts as well as hydrogen bonding, but the precise pairwise interactions between P5 and BamD residues differed (Fig. 7b), resulting in a relative translation of BamD from one model to the other. We favored model A because of the interaction between the highly conserved residues E373 from P5 and R188 from BamD. Additional salt bridges occurred between E187 from BamD with R353 and R366 from P5, as well as E183 from BamD with R350 from P5 (Fig. 7b, left). Model B was stabilized by electrostatic interactions of R366 (P5) with E187 and E183 (BamD), R353 (P5) with E183 (BamD), and K351 and R350 (P5) with D213 (BamD) (Fig. 7b, right).

However, docking of BamD to the BamA P4P5-TM construct led to different results. Residues from periplasmic loops 2 (F478-G482) and 3 (P518) were included as active in these runs, based on our ssNMR data and sequence conservation as described above (see Fig. 6a). As a consequence, BamD was drawn upwards to the periplasmic loops compared to the docking runs with isolated P4P5 (Fig. 7c), resulting in less interactions with our proposed binding site on P5. One model, hereafter referred to as model C, was selected in which the conserved residue R188 of BamD retained an interaction with D362 from P5 (Fig. 7c,d). D362 is involved in an electrostatic network with R366 and E373 within P5 (see Chapter 5), which means that these residues may also experience a change of environment even without direct binding to BamD. Further electrostatic interactions were formed by K361 from P5 with E183 and E187 from BamD. Interestingly, considering the

location of the lipid bilayer around the TM domain of BamA, model C would roughly result in alignment of BamD on the surface of the membrane (Fig. 7c).
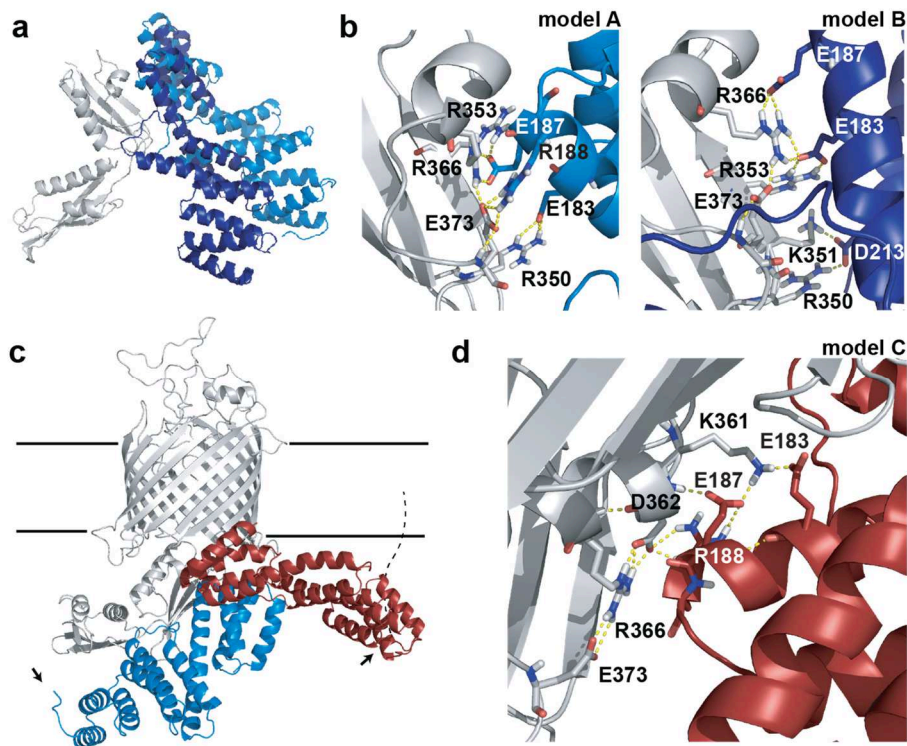
Figure 7. Docking models for the BamA-BamD interaction. a) Representative models of P4P5-BamD obtained with solution or ssNMR restraints. P4P5 is shown in grey, BamD in light blue for model A and in dark blue for model B. b) Close-ups of the P5-BamD interface from models A (left) and B (right), with residues contributing to electrostatic interactions labeled. c) Docking model C of BamA P4P5-TM (grey) with BamD (red) based on ssNMR restraints including the periplasmic loops. Model A is superimposed on the P4P5 domains with BamD in light blue. Horizontal lines indicate the estimated boundaries of the lipid bilayer, arrows the N-termini of BamD in the two models and the dashed line the hypothetical position of the 6 N-terminal residues and lipid anchor of BamD that are missing in the crystal structure for model C. d) Close-up of the P5-BamD interaction in model C, indicating the electrostatic interactions.

Discussion

BamA and BamD are the only essential components of the *E. coli* BAM complex that serves to fold and insert β-barrel proteins in the outer membrane. Crystal structures of the individual BAM components have been solved,[36-38,41,45,53,59-62] but the details of the interactions within the complex, as well as their interplay in the

process of β-barrel assembly, have remained elusive. In this work, we have studied the interaction between BamA and BamD in atomic detail using NMR spectroscopy in combination with docking.

Remarkably, although the BAM complex is a stable entity both *in vivo* and *in vitro*, we found that the interaction between BamA and BamD as isolated components is relatively weak. Solution NMR titrations of BamA P4P5 with BamD, as well as the reverse experiment, resulted in hardly detectable changes. Interestingly, the residues that seemed to shift or broaden slightly in these experiments were also the ones that suffer from intrinsic line broadening due to local conformational exchange. We previously identified this plasticity in P5 (see Chapter 5), but a similar effect was apparent in BamD (Fig. 2b).

In a more native-like environment, consisting of BamA P4P5-TM and lipid-anchored BamD reconstituted in lipid bilayers, the interaction was expected to be stronger due to a higher local protein concentration and pre-orientation of the binding partners, as well as a possible contribution of periplasmic loops from BamA to the binding interface. However, the two components still did not form a stable complex. Even a 2-fold excess of BamD caused only line broadening in ssNMR experiments but no CSPs, indicating an exchange interaction on the intermediate timescale.

Alternatively, we speculated that the other lipoproteins could affect the BamA-BamD interaction by modulating the conformation of BamD. A fragment of BamC, containing the N-terminal unfolded region and the first helix-grip domain, induces a conformational change in BamD judged from comparison of available crystal structures (Fig. 3a). However, including this BamC fragment did not increase the affinity of BamD for P4P5 in our experiments. Possibly, the conformational change in BamD does not affect its interaction with BamA, but it is also conceivable that the different conformations were caused by different crystal packing, rather than by binding of BamC. In this regard, it is interesting to note that the two helix-grip domains of BamC have been suggested to be exposed on the outside of the outer membrane,[59] which raises questions about the relevance of the BamCD structure in the absence of a membrane.

Ultimately, the key to a stable interaction between BamA and BamD turned out to be a combination of the membrane environment and the presence of the lipoproteins BamC and BamE. This is in agreement with *in vivo* data that show destabilization of the complex in deletion strains lacking BamC or BamE.[49,50] Likely, protein-lipid interactions also play a role. BamE binds negatively charged phospholipids,[62] whereas BamD has been shown to interact with both zwitterionic and negatively charged lipids.[194] Altogether this highlights the importance of performing structural studies on the BAM complex in a membrane environment. In

the current study, the BamA-CDE complex was studied in lipid bilayers composed of DLPC which is zwitterionic, but in principle the sample preparation method allows for reconstitution in membranes with different lipid compositions (see Chapter 3).

Finally, the CSPs in ssNMR spectra of the BamA-CDE complex corresponded to the same region of P5 that was slightly affected in solution NMR titrations with BamD, and that was found to experience local dynamics in previous studies (see Chapter 5). Individual docking runs using HADDOCK based on either solution or solid-state NMR data converged to two well-scoring models that were in agreement with the input data (Fig. 7a). Both of these models were characterized by extensive electrostatic interactions, but the pairwise contacts differed. However, the possibility exists that neither of these models represents the binding mode of BamA and BamD in context of the membrane-embedded BAM complex. In our ssNMR experiments, we observed CSPs stemming from the BamA TM domain, which likely corresponded to residues in the periplasmic loop that may be involved in the interaction with BamD. High conservation of a number of residues in pL2 and pL3 supported this hypothesis and docking with these residues as additional restraints led to very different models than the previous ones with only P4P5. Interestingly, model C (Fig. 7c) that was obtained in this way, showed an orientation of BamD that would be compatible with its binding to the lipid bilayer, in agreement with the *in vitro* data on BamD's interaction with phospholipids.[194]

Typically, mutagenesis studies can be applied to distinguish between possible docking models, yet the models described here were quite ambiguous in terms of residues involved in the interface. For example, mutation of R188 in BamD is likely to disrupt binding in case of model A, but also model B, whereas mutation of E183 and E187 would affect the interaction in all three proposed models. Model B however could be discriminated by mutation of D213 in BamD. On P5, the situation seems even more complex, because the charged residues are involved in a dynamic network within the domain (see Chapter 5). Residue E373 is known to be critical for binding of BamD and correct β-barrel assembly,[118] yet we have shown that mutation of this residue has long range effects on the conformational equilibrium of the P5 domain (Chapter 5), which may lead to a loss of binding even in absence of a direct contact of this residue with BamD.

It seems that more experimental data, such as assignments for the BamA TM domain (see Chapter 4), as well as for BamD, will be required to provide more restraints and guide the docking towards a reliable solution. Moreover, the current docking protocol does not take into account the specific features of membrane proteins, such as the physical barrier formed by the lipid bilayer as well as its low dielectric constant. However, even a dedicated docking method for membrane proteins would not be sufficient to obtain a molecular model for the BamA-BamD

interactions in atomistic lipid bilayers, for which molecular dynamics (MD) simulations would be required.

Strikingly, the conserved patch on the C-terminal side of BamD that we hypothesized as the BamA binding site (residues A175, E177, V181, A182, Y185, R188, A190, A193, V194, N196 and R197), was only sideways involved in the BamA-BamD interaction in our docking models. From the most conserved residues, only R188 that protrudes from one side of helix 7 was part of the interface with BamA in models A and C (Fig. 7b,d), whereas the remainder of the conserved patch remained accessible (Fig. 8). Possibly, these residues serve another purpose in the process of β-barrel assembly, such as binding to substrate OMPs. A binding site for the C-terminus of OMPs was previously proposed (Fig. 6b, green) based on the identification of a C-terminal His-tag from another molecule in the crystal lattice,[53] but the possible existence of an alternative binding site for unfolded polypeptides or of multiple binding sites can certainly not be excluded. Interestingly, whereas in models A (Fig. 8a) and B (not shown), the conserved patch of BamD was aligned parallel to P5, it was turned downwards facing the periplasm in model C. The first orientation might allow shuttling of the substrate from the POTRA domains to BamD, whereas in the latter, BamD would be accessible for direct targeting of the substrate from the periplasm.
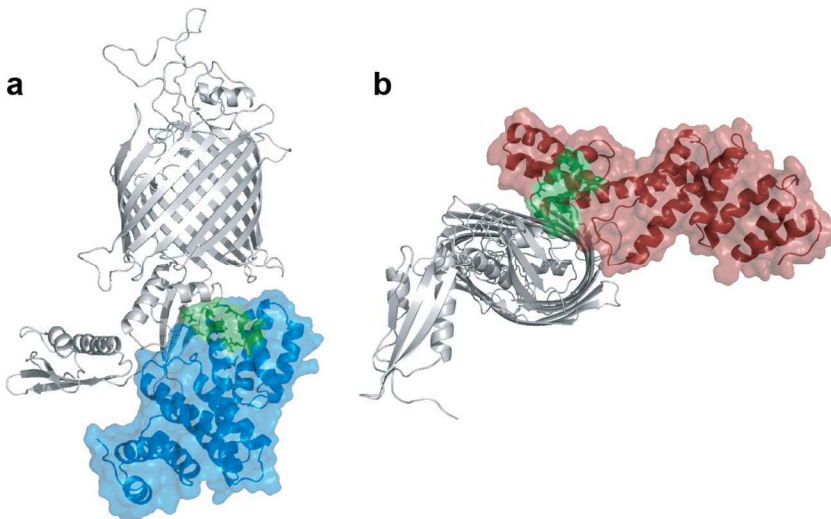
Figure 8. The conserved patch of residues in BamD shown as green sticks in the docking models. a) Model A overlaid on the *E. coli* homology model of BamA P4P5-TM, with BamD shown in blue surface representation. c) Model C seen from the periplasmic side. BamD is shown in red surface representation.

CHAPTER 7

Proton-clouds to measure long-range contacts between non-exchangeable side chain protons in solid-state NMR

Tessa Sinnige, Mark Daniëls, Marc Baldus, Markus Weingarth

Abstract

We show that selective labeling of proteins with protonated amino acids embedded in a perdeuterated matrix, dubbed 'proton-clouds', provides general access to long-range contacts between non-exchangeable side chain protons in proton-detected solid-state NMR, which is important to study protein tertiary structure. Proton-cloud labeling significantly improves spectral resolution by simultaneously reducing proton linewidth and spectral crowding despite a high local proton density in clouds. The approach is amenable to almost all canonical amino acids. Our method is demonstrated on ubiquitin and the β-barrel membrane protein BamA.

Introduction

Over the last years, solid-state NMR (ssNMR) spectroscopy has evolved as an established technique to study insoluble biomolecules, driven by impressive technical and methodological progress.[76,79,80,195-197] The use of ssNMR could be further boosted by the realization of ^{1}H-detection in solid biomolecules,[198-201] which by virtue of the high ^{1}H gyromagnetic ratio can enhance spectral sensitivity by more than one order of magnitude in comparison to heteronuclear detection. Labile amino protons can be accessed by means of perdeuteration and subsequent back-exchange of deuterons by protons,[87,198,200,202-204] possibly in combination with fast magic-angle spinning (MAS) and high magnetic fields.[205] High ^{1}H resolution has been reported with these approaches for backbone H_N in microcrystalline proteins, membrane proteins and amyloid fibrils.[86] H_N-H_N distance measurements[86,89] were shown to allow the rapid determination of protein fold.[89]

Protein structures are defined by side chain – side chain contacts[88,206,207] and protein structure determination by ^{1}H-detection hence requires access to aliphatic protons. Due to their high gyromagnetic ratio and peripheral location they are ideal to probe molecular distances in ssNMR, which has been extensively employed in heteronuclear-detected experiments.[76,79,195,208] As shown recently,[209] the use of fully protonated proteins is principally a straightforward mean to the assignment of aliphatic protons, although the measurement of long-range contacts (that is, separated by > 4 sequential residues) between non-exchangeable side chain protons in such systems is challenging due to spectral overlap and has not been reported thus far. Another approach, coined reduced adjoining protonation (RAP)[178] that relies on random incorporation of aliphatic protons in a deuterated matrix was shown to yield very high resolution. RAP was reported to give access to long-range contacts between methyl-protons. Such contacts could also be obtained using precursors with $^{13}C^1H_3$ groups.[206,88] Contacts between methyl-protons are precious structural probes. However, to determine high-resolution structures it is principally desirable to collect distance information for all side chains and chemical groups. Here we present a general avenue to measure long-range contacts between non-exchangeable protons by selective protonation of amino acid types. Such 'proton-clouds' in a perdeuterated background provide a significant improvement in ^{1}H-spectral resolution compared to fully protonated samples.

Results and discussion

Proton-cloud approach yields high-resolution ssNMR spectra

Proton-cloud samples are prepared by addition of uniformly ^{1}H,^{13}C,^{15}N-labeled amino acids to D$_2$O minimal medium containing ^{2}H,^{12}C-glucose. As a test case we prepared a sample of perdeuterated ubiquitin labeled with protonated valine and leucine (V,L ^{1}H-cloud ubiquitin). Figure 1a shows a superposition of ^{1}H-detected 2D ^{13}C,^{1}H correlation spectra of fully protonated (FP) ubiquitin (in blue, see Fig. 1b) and V,L ^{1}H-cloud ubiquitin (in red, see Fig. 1c). It is readily visible that spectral congestion is much reduced in the ^{1}H-cloud spectrum, which is a combined effect of a narrower ^{1}H linewidth (Fig. 1d), which is about a factor of two to three better than in the FP sample (0.12 - 0.26 ppm), and a stark reduction in signal crowding. The improvement in ^{1}H linewidth is most significant for protons of CH and CH$_2$ groups and almost a factor of two for methyl-protons. The spectral quality in the ^{1}H-cloud sample allowed, based on available ^{1}H solution[210] and ^{13}C solid-state[211] NMR assignments, to identify all C$_\alpha$-H$_\alpha$ groups (except for the highly mobile sites L8, L71 and L73) and most of the side chain resonances. In all ssNMR experiments the MISSISSIPI[105] scheme was used for water suppression and low-power PISSARRO decoupling was applied in both ^{13}C and ^{1}H dimensions.[103] Details of the pulse sequence are shown in Figure 2.

Proton-cloud labeling hence provides attractive spectral resolution. However, the ^{1}H linewidth in RAP samples, that was demonstrated to be on the order of 25-60 Hz,[178,212] is narrower than in ^{1}H-cloud samples, which we suspected to result from residual homogeneous line-broadening in ^{1}H-clouds.[213] Numerical simulations indeed indicate that very high magnetic fields (≥ 1000 MHz) and MAS frequencies (≥ 90 kHz) are required to suppress the local proton couplings of a single valine residue (Fig. 3a,b). Experimental data (Fig. 3c), measured at 40 – 60 kHz MAS frequency and 16.4 T (700 MHz ^{1}H frequency) magnetic field showed an approximately linear decrease in ^{1}H linewidth with increasing MAS frequency, which is in line with previous studies,[201,213] confirming the presence of homogeneous contributions. We observed broad ^{1}H line shapes at medium MAS frequencies, which further illustrates strong local dipolar couplings in ^{1}H-cloud preparations (data not shown).
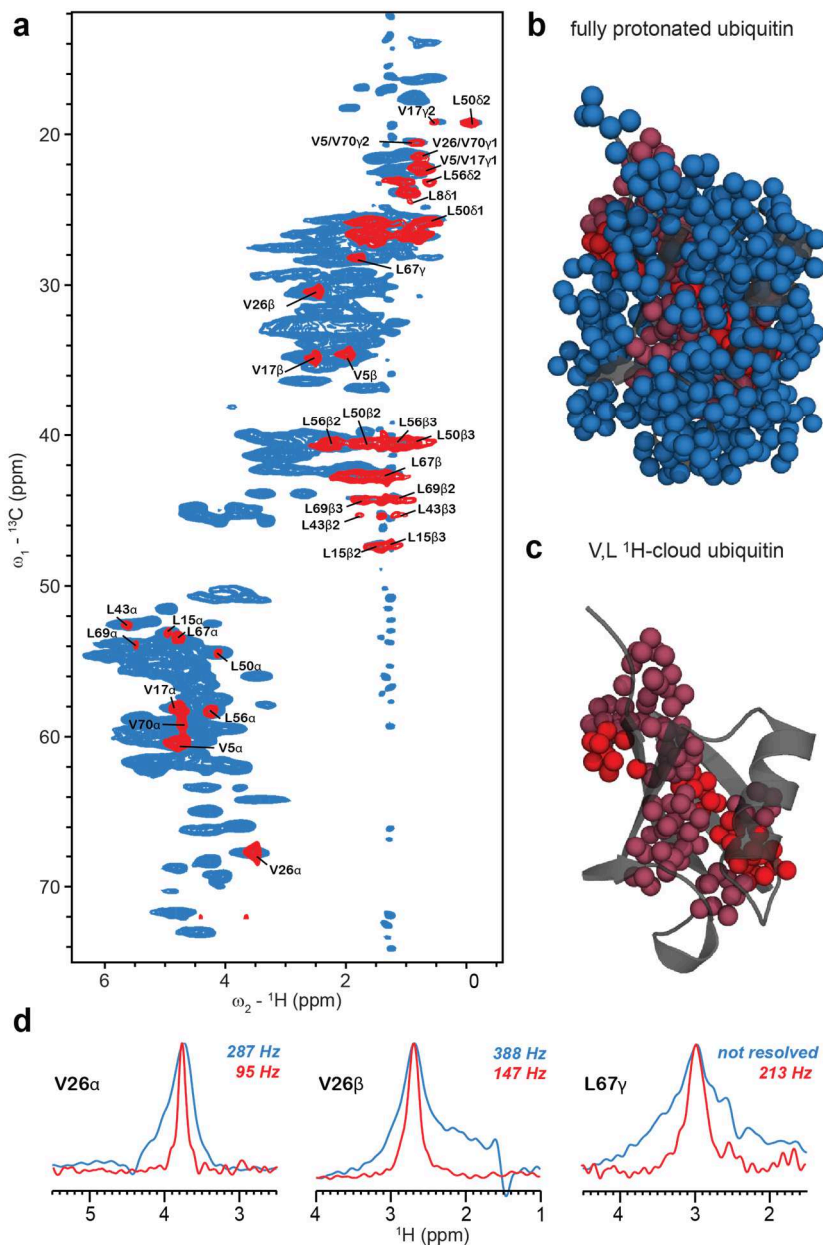
Figure 1. Proton-clouds reduce spectral congestion with respect to full protonation. a) Superposition of 2D ^{13}C,^{1}H spectra of perdeuterated V,L ^{1}H-cloud ubiquitin (red) and fully protonated (FP) ubiquitin (blue). Spectra were recorded at 18.8 T (800 MHz ^{1}H frequency) and 60 kHz MAS. The $C_\alpha H_\alpha$ region of the V,L ^{1}H-cloud spectrum is shown with a scaling factor of 3.5 with respect to the rest of the ^{1}H-cloud spectrum. b) Protonation pattern of FP ubiquitin. c) Protonation pattern of V,L ^{1}H-cloud ubiquitin. Leucine and valine residues are colored in dark and light red, respectively. d) Superposition of cross-sections extracted from a) without application of a window function in the ^{1}H dimension.

Figure 2. Pulse sequence diagrams. a) 2D ^{13}C,^{1}H experiment with phase cycling: ϕ_1 = x; ϕ_2 = y, -y; ϕ_3 = x; ϕ_4 = y, y, y, y, -y, -y, -y, -y, -y, -y, -y, -y, y, y, y, y; ϕ_5 = y, y, y, y, y, y, y, y, -y, -y, -y, -y, -y, -y, -y, -y; ϕ_6 = x, x, -x, -x, -x, -x, x, x, x, x, -x, -x, -x, -x, x, x; ϕ_7 = y; ϕ_8 = x, -x, -x, x, x, -x, -x, x, x, -x, -x, x, x, -x, -x, x. b) the 2D ^{13}C,(^{1}H)^{1}H experiment with phase cycling: ϕ_1 = x; ϕ_2 = y, -y; ϕ_3 = x; ϕ_4 = y, y, y, y, -y, -y, -y, -y, -y, -y, -y, -y, y, y, y, y; ϕ_5 = y, y, y, y, y, y, y, y, -y, -y, -y, -y, -y, -y, -y, -y; ϕ_6 = x, x, -x, -x, -x, -x, x, x, x, x, -x, -x, -x, -x, x, x; ϕ_7 = y; ϕ_8 = x; ϕ_9 = -x; ϕ_{10} = x, -x, -x, x, x, -x, -x, x, x, -x, -x, x, x, -x, -x, x. For both experimental schemes quadrature-detection in t_1 was achieved using TPPI.

Figure 3. Homogeneous contribution to the ^{1}H linewidth in ^{1}H-cloud samples. a) The ^{1}H linewidth of V17H$_\beta$ shown as a function of MAS frequency and magnetic field. Green stars indicate the experimental conditions used in Figures 1 and 5. Note that experimental and simulated values differ due to the finite proton network in simulations and possibly inhomogeneous broadening effects seen in our experiments. b) The ten-spin system used to carry out numerical simulations. c) Experimental ^{1}H linewidth as a function of the MAS frequency, measured at 700 MHz with the ubiquitin ^{1}H-cloud sample described in Figure 5.

Long-range contacts between proton-clouds in ubiquitin

In ¹H-cloud samples, the closest ¹H neighbors do usually not correspond to long-range contacts and dipolar truncation curtails the potential of first-order recoupling methods to bring about non-trivial distance-information.[214] However, the presence of pronounced residual homogeneous proton couplings at high MAS frequencies opens up the possibility to resort to second-order recoupling methods including spin diffusion. It was demonstrated that such methods are not very sensitive to dipolar truncation,[208,215,216] which provides a potential way to transfer magnetization between ¹H-clouds. We probed whether ¹H-cloud samples gave rise to long-range contacts by means of ¹H-¹H spin diffusion,[208] using a ubiquitin sample in which only ¹H-valine was supplemented to the perdeuterated medium, while leucine C_γ and C_δ positions became ¹³C labeled and protonated due to scrambling (Fig. 4). The protein was precipitated in D_2O based buffers and ²H-MPD to prevent reintroduction of fast relaxing precipitant signal during the mixing block. 2D ¹³C,(¹H)¹H spectra (Fig. 5; see Fig. 2 for details of the pulse sequence) acquired at 55 and 60 kHz MAS with 25 ms and 75 ms ¹H-¹H mixing, respectively, showed numerous long-range contacts between ¹H-clouds, spanning distances of 2 – 6 Å in the hydrophobic core of ubiquitin. Long-range contacts including H_β protons of valine (V17H_β – L56$H_{\delta1}$ and V26H_β – L43$H_{\delta1}$) and H_γ of leucine (L43H_γ - L50$_{\delta2}$ and V26H_β – L43H_γ) demonstrate that ¹H-cloud sample preparations bear the capacity to provide a general access to long-range contacts between non-exchangeable side chain protons. It is also apparent from Figure 5 that the access to chemical groups other than methyl alleviates spectral degeneracy.
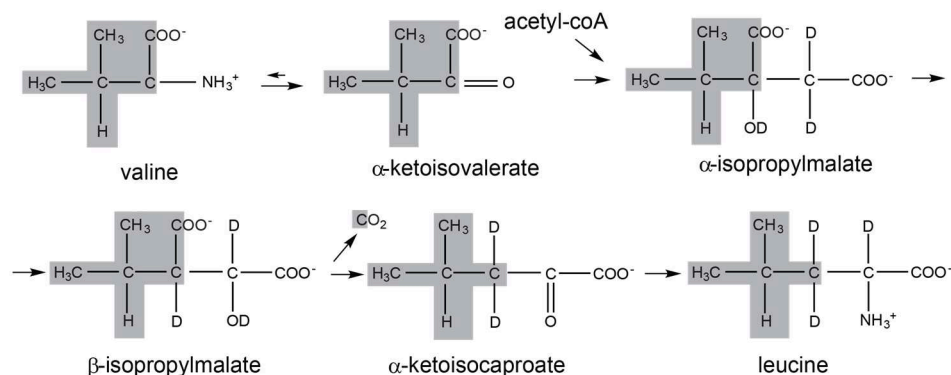

Figure 4. Scrambling of valine to leucine occurred during protein expression, causing the ¹³C labeled, protonated valine that was added to the medium to convert to leucine. Shown is the pathway in *Escherichia coli* with ¹³C and ¹H nuclei that are retained during conversion from valine to leucine shaded in grey. Note that only C_γ, $C_{\delta1}$ and $C_{\delta2}$ of leucine will be both ¹³C labeled and protonated, corresponding to the correlations detected in our NMR experiments.
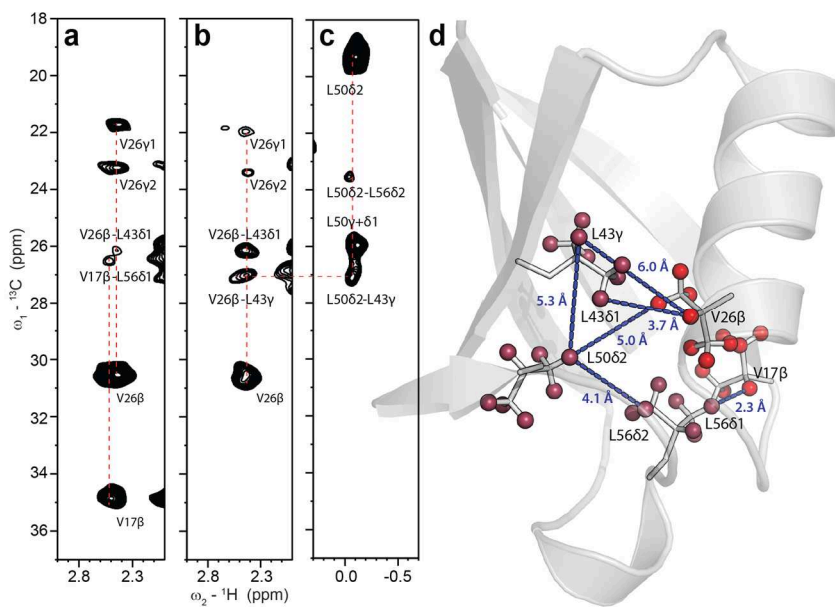
Figure 5. Long-range contacts between proton-clouds in ubiquitin. a-c) Cutouts of 2D ^{13}C,(^{1}H)^{1}H spectra, measured at a) 55 kHz and b,c) 60 kHz MAS with 25 ms and 75 ms ^{1}H-^{1}H spin diffusion mixing time, respectively, using V,L* ^{1}H-cloud ubiquitin (*only C_γ and C_δ positions of leucine were protonated and ^{13}C labeled). Long-range contacts involving a,b) V17$_\beta$ and V26$_\beta$ and c) L50$_{\delta2}$ are annotated. d) Illustration of the long-range contacts annotated in a-c) in the crystal structure of ubiquitin (PDB: 1UBQ).[217]

Proton-clouds for non-exchangeable sites in membrane proteins

^{1}H-cloud labeling does not require reprotonation of exchangeable sites, which is especially challenging for membrane proteins that are shielded by the lipid bilayer.[218] To probe the scope of ^{1}H-cloud labeling, we prepared a V,L,K ^{1}H-cloud sample of membrane-embedded BamA, which is the main component of the β-barrel assembly machinery.[41,91] We could readily identify separated spectral regions that group signature correlations of the three amino acid types in a 2D ^{13}C,^{1}H spectrum (Fig. 6a). In comparison to FP BamA, spectral congestion is much reduced in ^{1}H-cloud BamA spectrum (Fig. 6b), which, drawing on chemical shift predictions,[145,146] allows the observation of the apparent absence of low-field valine residues located in periplasmic (V480) and extracellular (V543 and V706) loops, suggesting increased mobility of these elements (Fig. 6b,c). Spectral congestion is increased in comparison to V,L ^{1}H-cloud ubiquitin, due to the residue degeneracy (31 x V, 25 x L, 21 x K) in BamA (52 kDa for the construct used in this study) and due to a residual ^{1}H linewidth of 0.3 – 0.4 ppm. Causes for the increased ^{1}H linewidth are presumably

that BamA is not microcrystalline and that the hydrophobic side chains in BamA point to the protonated membrane (Fig. 6c), which may be remedied by the use of deuterated lipids. We observed efficient ^{1}H-^{1}H magnetization transfer in V,L,K ^{1}H-cloud BamA including transfer to amino and aliphatic protons (Fig. 6e,f). This bodes well for future structural studies of BamA and other large membrane proteins, especially since the spectral quality of ^{1}H-cloud samples will considerably benefit from emerging magnetic fields (≥ 1000 MHz) and MAS frequencies (≥ 80 kHz) (see Fig. 3).
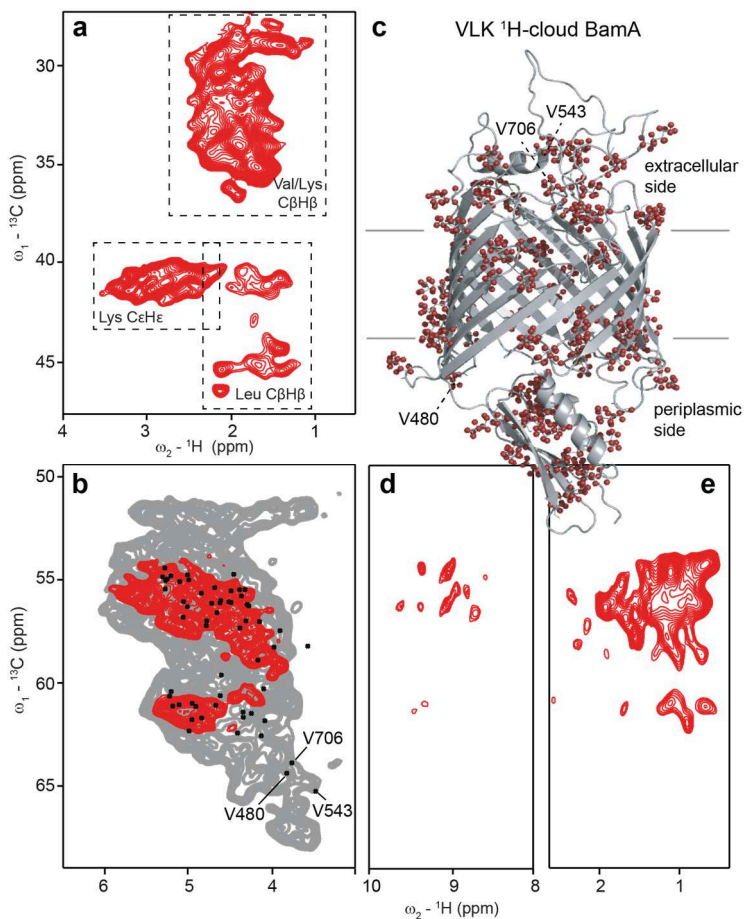
Figure 6. a) Close-up of a 2D ^{13}C,^{1}H spectrum measured at 55 kHz MAS and 800 MHz using a V,L,K ^{1}H-cloud sample of membrane protein BamA, washed in D$_2$0. b) Close-up of the same spectrum as in a) (red) superposed on that of FP BamA (grey). The FANDAS[145] spectral predictions (black crosses) were obtained with SHIFTX2[146] using the homology model for *E. coli* BamA.[41] c) Val, Leu and Lys residues highlighted on the BamA homology model by red spheres. d,e) Cutouts of a 2D ^{13}C,(^{1}H)^{1}H spectrum showing transfer from H$_\alpha$ to d) H$_N$ and e) side chain protons.

Importantly, our approach is not restricted to the presence of methyl groups but renders most canonical amino acids amenable to measure contacts between non-exchangeable protons. This includes positively charged (Lys, Arg, His) and aromatic (Phe, Trp, Tyr) amino acids, which are all devoid of methyl functions. To avoid scrambling of the labels to unintended positions, certain amino acids, such as serine, cysteine and glycine, must be simultaneously labeled, while most amino acids can be added independently to mix the desired ^{1}H-cloud pattern. Alternatively, auxotrophic strains can be used to avoid scrambling (see e.g. Ref 219). Generally, ^{1}H-cloud labeling can be readily designed based on the nature of the research problem. For structure calculations, combining restraints from different mixtures of aliphatic ^{1}H-clouds will be most informative while ^{1}H-cloud labeling with charged, polar or aromatic residues can be useful to refine protein catalytic sites or study protein – membrane interactions.[77] In conclusion, we have shown that fully protonated amino acids in a deuterated background give rise to favorable spectral resolution despite a high local proton density. Such sample preparations provide a general and straightforward approach to obtain long-range distance information between non-exchangeable protons in ^{1}H-detected experiments, which we assume to increase the use of ssNMR in structural biology.

CHAPTER 8

Discussion and perspectives

Understanding membrane protein folding

Protein folding is an essential process for every living cell. After synthesis at the ribosome, the nascent polypeptide chain has to adopt its unique three-dimensional structure that ensures its proper functioning. Importantly, protein misfolding and aggregation are associated with many diseases ranging from cancer to Alzheimer's disease. Not surprisingly, protein folding has been a widely studied subject since more than fifty years ago. However, the folding pathways of integral membrane proteins, which represent 30% of known proteomes and are important drug targets, have remained difficult to study due to the need for a membrane-mimicking environment. In addition, these systems are often large and inherently flexible. Especially the mechanisms behind the folding of β-barrel proteins that reside in the outer membranes of Gram-negative bacteria, mitochondria and chloroplasts remain poorly understood, relative to α-helical proteins from inner bacterial and eukaryotic membranes. Characterizing the process of outer membrane protein (OMP) folding in Gram-negative bacteria would not only yield fundamental insights into membrane protein folding, but may also aid the development of antibiotics targeted at this essential process.

About a decade ago, proteins involved in β-barrel assembly were first identified.[32,33,187] More recently, high-resolution crystal and solution NMR structures of all components of the archetypical Gram-negative β-barrel assembly machinery (BAM) have become available,[36-39,41,44-46,53,54,58,60-62] including the main, highly conserved integral membrane protein BamA and the lipoproteins BamBCDE, yet the mechanism by which it functions to fold and insert β-barrel proteins is still not well understood. Major questions concern dynamics and conformational changes that are likely to drive β-barrel assembly in the absence of an energy source such as ATP in the periplasm. Also, protein-lipid and protein-protein interactions within the BAM complex and with the substrate have remained elusive. Solid-state NMR (ssNMR) spectroscopy is capable of studying membrane proteins in their native environment consisting of lipid bilayers. Moreover, it does not suffer from inherent size limitations (see Chapter 1, Fig. 2 for the sizes of the BAM components). Therefore, we set out to investigate the main component BamA in lipid bilayers by ssNMR, complemented with solution NMR, electron microscopy and computational methods, with the aim to obtain insights into its dynamics and complex formation with other components of the machinery.

Implications for the mechanism of β-barrel assembly

The importance of studying BamA in a lipid environment followed immediately from our results. We found that the periplasmic POTRA domains of BamA were overall relatively rigid when the protein was reconstituted in lipid bilayers (Chapter 3), despite different conformations in crystal structures that suggested large-scale reorientations of the POTRA domains.[41] Although motion on a slower timescale, i.e. milliseconds to seconds, cannot be excluded, it is more likely that the different orientations of the POTRA domains in the crystal structures were caused by crystallographic packing, whereas in proteoliposomes they were stabilized by interactions between POTRA 5 (P5) and the BamA transmembrane (TM) domain, or perhaps by interactions between the POTRA domains and the lipid bilayer. Interestingly, molecular dynamics (MD) simulations from our lab on BamA in lipid bilayers have shown association of the POTRA domains with the phosphatidyl choline head groups (Dr. Markus Weingarth, unpublished observation), which, however, remains to be experimentally verified.

Despite the absence of overall domain motion, we did identify *local* motion in residues of the P5 domain that mapped to a highly conserved region of BamA, suggesting functional relevance (Chapter 5). Indeed, we found that this part of P5 formed the main binding site for BamD (Chapter 6), which is the only essential lipoprotein of the *E. coli* BAM complex.[49,51] Remarkably, the interaction between the two components BamA and BamD seemed relatively weak, whereas the BamCDE sub-complex stably associated with BamA in lipid bilayers. Moreover, in this complex, the local dynamics of P5 appeared restrained. This raises the possibility that the conformational plasticity of P5 that we observed in isolated BamA was merely an artifact of our *in vitro* experiments caused by absence of the lipoproteins. However, β-barrel assembly is likely a dynamic process in which the lipoproteins associate and dissociate. Dissociation of BamC and/or BamE would weaken the BamA-BamD interaction according to our findings, in agreement with published *in vivo* data.[49,50] Dissociation of BamD in turn would leave P5 free to undergo conformational exchange, which may be involved in targeting substrate OMPs to the BamA β-barrel.

Alternatively, the local conformational exchange in P5 might be directly related to the interaction with BamD. In our NMR experiments, the mutation E373K in P5, which is lethal at normal growth temperatures and abolishes the BamA-BamD interaction *in vivo*,[118] resulted in increased local dynamics (Chapter 5). E373 was found to be located in the BamD binding region and was involved in electrostatic interactions with BamD in two of the selected docking models (Chapter 6). Mutation of E373 hence could simply destabilize or destroy the BamA – BamD interface. However, another possibility for the lost ability to bind to BamD could be a change in the conformational equilibrium of the E373K mutant.

Interestingly, although the role of BamD is not known, it has been implied in substrate interactions. Previous studies suggested a conserved substrate binding site in between the N-terminal TPRs 1-3,[53] yet in our docking models of the BamA-BamD complex, a highly conserved patch on TPR 4 which we speculated to be involved in BamA binding was still exposed and hence might also serve as substrate binding site. Notably, the previously suggested N-terminal substrate binding site on BamD is occluded by BamC in the crystal structure of the BamCD complex (see Chapter 6, Fig. 6b).[60] It is quite unlikely that BamC regulates access to the substrate binding site of BamD, as it is not essential in *E. coli* and absent from many Gram-negative bacteria. Therefore a substrate binding site at TPR 4 of BamD, which is also localized more closely to BamA, may be a more logical possibility.

Another interesting, and still not resolved, feature of the β-barrel assembly machinery is the role of the long extracellular loop 6 (eL6) in BamA that carries the essential VRGF/Y motif. In crystal structures of three BamA homologues, eL6 folds back into the lumen of the β-barrel and is stabilized by various interactions, which involve other conserved motifs.[41,42] In ssNMR experiments on membrane-embedded BamA, we tentatively assigned several short stretches from eL6 (Chapter 4), confirming that also in our sample preparations this loop was relatively rigid and did not exchange between different conformations. When examining the spectra of BamA in complex with BamCDE (Chapter 6) with these assignments at hand, it appeared that resolved resonances from e.g. residue I668 from eL6 remained unchanged, suggesting that the loop was not affected by binding of the lipoproteins in our experiments.

In vivo, however, the situation may be different. A study from the Silhavy group suggested that eL6 can sample two conformations, one being buried and the other exposed, hence accessible to protease digestion and chemical labeling.[69] Based on mutagenesis data, a conformational cycle was proposed that involves exposure of eL6 driven by BamD and return to the initial buried state aided by BamE. As noted earlier, such conformational rearrangements are likely caused by interactions with the substrate and its subsequent folding and insertion. In our ssNMR experiments we observed the steady-state conformation, highlighting the necessity of studies including a substrate OMP as discussed below.

Perspectives for substrate interaction studies

Ultimately, to understand the mechanism behind β-barrel assembly, it is not sufficient to study the isolated BAM complex in membranes, but the process needs to be monitored in presence of the unfolded OMP as well. It is currently not known which regions of the BAM complex interact with the substrate, what conformational changes are induced in the BAM machinery and what pathway the substrate follows to fold.

In vitro studies have been reported in which the kinetics of OMP folding in lipid bilayers with pre-incorporated BamA or the entire BAM complex were monitored,[21,220-222] yet a molecular description of the process cannot be derived from these experiments. The ssNMR toolbox that we have developed to study BAM components in a membrane environment would in principle allow for studies in which conformational changes in the machinery induced by substrate

binding might be detectable. Furthermore, to probe where the substrate interacts, it could be labeled with a paramagnetic tag that quenches all regions of the BAM complex that it encounters. On the other hand, the substrate may also be isotope labeled and added to unlabeled BAM components to monitor its folding and insertion into the lipid bilayer, and possibly detect intermediate states.

However, the conditions for such experiments would have to be controlled very carefully. The concentrations of BAM components and substrate in the published *in vitro* studies[21,220-222] are much lower than required for NMR purposes. One problem associated with the need for higher protein concentrations is aggregation of the unfolded substrate, which might have to be circumvented by the addition of periplasmic chaperones, adding further complexity to the system. Furthermore, to achieve conditions in which the substrate can actually fold and insert into the bilayer, high lipid-to-protein concentrations are required, which is not compatible with the high protein concentrations necessary for NMR.

Moreover, to obtain a snapshot of the substrate while it is folding, it would need to be trapped e.g. by cross-linking or rapidly freezing the sample. In the latter case, one may consider to perform Dynamic Nuclear Polarization (DNP) that is done at temperatures around 100 K and yields higher sensitivity than conventional ssNMR by transferring polarization from an unpaired electron to the nucleus of interest (see e.g. Ref. 122,126,223,224), alleviating the need for high protein concentrations. However, resolution is a concern in DNP experiments on large (membrane) proteins such as BamA, due to freezing out of different conformations as well as paramagnetic quenching by the radical that has to be added to the sample. An alternative way to trap the substrate without cross-linking or freezing would be to use a mutant autotransporter as substrate. Autotransporters are virulence factors that consist of a β-barrel domain and a passenger domain that is secreted through the outer membrane (OM). Several autotransporters have been shown to depend on BamA for folding of the β-domain,[225-228] and very interestingly, it was demonstrated that translocation of the passenger domain can only occur once folding of the β-domain is complete. This means that mutants defective in secretion are trapped at the BAM complex,[227] which would make them an ideal case to study the influence of substrate binding on the conformation of BamA and the BAM complex by ssNMR.

Current developments in ssNMR

Currently, spectral resolution and the lack of extensive assignments of the BamA TM domain are the limiting factors in the study of BamA and the BAM complex by ssNMR. This is not surprising, given that a number of challenges is associated with ssNMR studies on large membrane proteins such as BamA. With the development of a protocol for reconstitution of BamA in lipid bilayers at low lipid-to-protein ratio leading to high-resolution ssNMR spectra with reasonable sensitivity, a first barrier was overcome (Chapter 3). Furthermore, solution NMR assignments from literature[39] and from our experiments (Chapter 5) could be employed to study the POTRA domains at atomic resolution. However, obtaining ssNMR assignments for the BamA TM domain has been a bottleneck. A number of tentative assignments was obtained using various labeling schemes combined with 2D and 3D ssNMR experiments (Chapter 4), yet interesting regions such as the proposed lateral gate remained devoid of any information. Whether this was caused by a lack of extensive assignments or due to protein dynamics, remains to be determined.

Fortunately, ssNMR methods are developing rapidly in all aspects, ranging from sample preparation procedures to technological and computational advances. Most promisingly, proton-detection has entered ssNMR over the recent years as a combined result of labeling schemes with high levels of deuteration, fast MAS and dedicated pulse sequences (see e.g. Ref. 86-89,177,178,202,203). We expanded the scope of proton-detected ssNMR with the proton-cloud approach described in Chapter 7 and showed that this method is applicable to large membrane proteins such as BamA. With the introduction of faster MAS and higher magnetic fields in the near future, proton-detection bears the promise to boost sensitivity as well as spectral resolution in the study of membrane proteins by ssNMR.

In addition, several computational approaches can be implemented to aid sequential assignment for large proteins in ssNMR. In the work described in this thesis we made use of the program FANDAS[145] that predicts spectra based on chemical-shift or secondary structure input and allows easy identification of signals in multi-dimensional ssNMR experiments. Furthermore, automated assignment routines that were originally developed for solution NMR have now been made available for ssNMR.[229] Also, non-uniform sampling (NUS) can be used to obtain higher signal-to-noise per unit of measurement time,[230,231] which is especially critical for higher-dimensional (e.g. 3D and 4D) experiments[88,232] that would otherwise not always be feasible for large membrane proteins. Taken together, current developments in the field of ssNMR will be crucial for

resonance assignment of large membrane proteins such as BamA, and will open up more possibilities for atomic-resolution studies on the structure and dynamics of these systems.

BAM in its native environment?

From a biological perspective, a breakthrough would be achieved if we could study BamA in its actual native environment. It is clear that the phosphatidylcholine (PC) lipids used in the studies described in this thesis are very different from the actual composition of the OM. We did show that the BamA TM domain could accommodate to lipid bilayers of different hydrophobic thicknesses and saturation levels, resulting in only slight changes in the ssNMR spectra (Chapter 3). In addition, electron microscopy showed that high concentrations of BamA caused distortions in the PC liposomes, which is compatible with other data that suggest that BamA locally induces thinning of the bilayer, which may help to insert OMPs.[41] However, PC is not a native lipid in *E. coli* membranes and the interplay between BamA and the OM remains to be determined. Attempts to reconstitute BamA in an *E. coli* polar lipid extract, which represents at least the composition of the inner leaflet of the OM, were not successful as protein was systematically detected in the supernatant after harvesting the proteoliposomes. Possibly, incorporation of BamA would be more complete at higher lipid-to-protein ratios. Still, the outer leaflet of the OM consists of lipopolysaccharide (LPS), so synthetic lipid bilayers can never represent the native environment of OMPs. Furthermore, other components such as periplasmic proteins and the peptidoglycan layer may interact with the BAM components *in vivo*.

The only way to study BamA in its native environment would be in whole *E. coli* cells or cell envelopes, such as previously demonstrated in our lab for PagL, an OMP enzyme from *Pseudomonas aeruginosa*.[81,82] However, the feasibility of this approach would depend on the levels of overexpression that can be achieved for BamA while retaining proper folding and integration into the OM. Furthermore, spectral crowding would be an issue that is already encountered in ssNMR studies on purified BamA, but would be even more pronounced in the presence of endogenous protein and lipid components. However, with dedicated labeling schemes and a particular question in mind, it might be possible to e.g. validate the overall rigidity of the POTRA domains or to study the conformation and dynamics of eL6. Finally, to return to the advances in ssNMR described above, proton-detected methods may greatly enhance the possibilities for

cellular ssNMR on systems like BamA. Especially the proton-cloud approach (Chapter 7) may be very powerful for cellular samples, as non-proteinaceous cell components would not become ^{13}C-labeled. Furthermore, the method does not rely on back-exchange of protons which might be buried in a cellular context.

Altogether, when biochemical and technological developments are combined, much more is to be expected from NMR studies on the structure and dynamics of the BAM complex, hopefully leading to a better understanding of the essential process of β-barrel folding and insertion in outer membranes.

REFERENCES

1. Anfinsen, C.B. Principles that govern the folding of protein chains. *Science* **181**, 223-230 (1973).
2. Levinthal, C. How to fold graciously. *Mossbauer spectroscopy in biological systems*, 22-24 (1969).
3. Rapoport, T.A. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* **450**, 663-669 (2007).
4. Driessen, A.J.M. & Nouwen, N. Protein translocation across the bacterial cytoplasmic membrane. *Annu. Rev. Biochem.* **77**, 643-667 (2008).
5. Huysmans, G.H.M., Baldwin, S.A., Brockwell, D.J. & Radford, S.E. The transition state for folding of an outer membrane protein. *Proc. Nat. Acad. Sci. USA* **107**, 4099-4104 (2010).
6. Huysmans, G.H.M., Radford, S.E., Baldwin, S.A. & Brockwell, D.J. Malleability of the folding mechanism of the outer membrane protein PagP: parallel pathways and the effect of membrane elasticity. *J. Mol. Biol.* **416**, 453-464 (2012).
7. Kleinschmidt, H. & Tamm, L.K. Folding intermediates of a beta-barrel membrane protein. Kinetic evidence for a multi-step membrane insertion mechanism. *Biochemistry* **35**, 1-8 (1996).
8. Kleinschmidt, J.H. & Tamm, L.K. Time-resolved distance determination by tryptophan fluorescence quenching: probing intermediates in membrane protein folding. *Biochemistry* **38**, 4996-5005 (1999).
9. Kleinschmidt, J.H. Folding kinetics of the outer membrane proteins OmpA and FomA into phospholipid bilayers. *Chem. Phys. Lipids* **141**, 30-47 (2006).
10. Paetzel, M., Dalbey, R.E. & Strynadka, N.C.J. Crystal structure of a bacterial signal peptidase in complex with a β-lactam inhibitor. *Nature* **396**, 186-190 (1998).
11. Korndörfer, I.P., Dommel, M.K. & Skerra, A. Structure of the periplasmic chaperone Skp suggests functional similarity with cytosolic chaperones despite differing architecture. *Nat. Struct. Mol. Biol.* **11**, 1015-1020 (2004).
12. Walton, T.A. & Sousa, M.C. Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Mol. Cell* **15**, 367-374 (2004).
13. Walton, T.A., Sandoval, C.M., Fowler, C.A., Pardi, A. & Sousa, M.C. The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proc. Nat. Acad. Sci. USA* **106**, 1772-1777 (2009).
14. Burmann, B.M., Wang, C. & Hiller, S. Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp. *Nat. Struct. Mol. Biol.* **20**, 1265-1272 (2013).
15. Bitto, E. & McKay, D.B. Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins. *Structure* **10**, 1489-1498 (2002).
16. Behrens, S., Maier, R., de Cock, H., Schmid, F.X. & Gross, C.A. The SurA periplasmic PPIase lacking its parvulin domains functions in vivo and has chaperone activity. *EMBO J.* **20**, 285-294 (2001).
17. Hennecke, G., Nolte, J., Volkmer-Engert, R., Schneider-Mergener, J. & Behrens, S. The periplasmic chaperone SurA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. *J. Biol. Chem.* **280**, 23540-23548 (2005).
18. Bitto, E. & McKay, D.B. The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J. Biol. Chem.* **278**, 49316-49322 (2003).
19. Bitto, E. & McKay, D.B. Binding of phage-display-selected peptides to the periplasmic chaperone protein SurA mimics binding of unfolded outer membrane proteins. *FEBS Lett.* **568**, 94-998 (2004).
20. Sklar, J.G., Wu, T., Kahne, D. & Silhavy, T.J. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev.* **21**, 2473-2484 (2007).
21. Hagan, C.L., Kim, S. & Kahne, D. Reconstitution of outer membrane protein assembly from purified components. *Science* **328**, 890-892 (2010).
22. Spiess, C., Beil, A. & Ehrmann, M. A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97**, 339-347 (1999).

23.	Bos, M.P., Robert, V. & Tommassen, J. Biogenesis of the gram-negative bacterial outer membrane. *Annu. Rev. Microbiol.* **61**, 191-214 (2007).
24.	Hagan, C.L., Silhavy, T.J. & Kahne, D. β-Barrel membrane protein assembly by the Bam Complex. *Annu. Rev. Biochem.* **80**, 189-210 (2011).
25.	Krojer, T., Sawa, J., Schafer, E., Saibil, H.R., Ehrmann, M. & Clausen, T. Structural basis for the regulated protease and chaperone function of DegP. *Nature* **453**, 885-890 (2008).
26.	Jiang, J., Zhang, X., Chen, Y., Wu, Y., Zhou, Z.H., Chang, Z. & Sui, S.-F. Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. *Proc. Nat. Acad. Sci. USA* **105**, 11939-11944 (2008).
27.	Shen, Q.-T., Bai, X.-C., Chang, L.-F., Wu, Y., Wang, H.-W. & Sui, S.-F. Bowl-shaped oligomeric structures on membranes as DegP's new functional forms in protein quality control. *Proc. Nat. Acad. Sci. USA* **106**, 4858-4863 (2009).
28.	Ge, X., Wang, R., Ma, J., Liu, Y., Ezemaduka, A.N., Chen, P.R., Fu, X. & Chang, Z. DegP primarily functions as a protease for the biogenesis of β-barrel outer membrane proteins in the Gram-negative bacterium *Escherichia coli*. *FEBS J.* **281**, 1226-1240 (2014).
29.	Volokhina, E.B., Grijpstra, J., Stork, M., Schilders, I., Tommassen, J. & Bos, M.P. Role of the periplasmic chaperones Skp, SurA, and DegQ in outer membrane protein biogenesis in *Neisseria meningitidis*. *J. Bacteriol.* **193**, 1612-1621 (2011).
30.	Chen, R. & Henning, U. A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol. Microbiol.* **19**, 1287-1294 (1996).
31.	Tommassen, J. Assembly of outer-membrane proteins in bacteria and mitochondria. *Microbiology* **156**, 2587-2596 (2010).
32.	Voulhoux, R., Bos, M.P., Geurtsen, J., Mols, M. & Tommassen, J. Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**, 262-265 (2003).
33.	Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T.J. & Kahne, D. Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**, 235-245 (2005).
34.	Gentle, I., Gabriel, K., Beech, P., Waller, R. & Lithgow, T. The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J. Cell Biol.* **164**, 19-24 (2004).
35.	Sánchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M. & Valencia, A. POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends Biochem. Sci.* **28**, 523-526 (2003).
36.	Kim, S., Malinverni, J.C., Sliz, P., Silhavy, T.J., Harrison, S.C. & Kahne, D. Structure and function of an essential component of the outer membrane protein assembly machine. *Science* **317**, 961-964 (2007).
37.	Gatzeva-Topalova, P.Z., Walton, T.A. & Sousa, M.C. Crystal structure of YaeT: conformational flexibility and substrate recognition. *Structure* **16**, 1873-1881 (2008).
38.	Gatzeva-Topalova, P.Z., Warner, L.R., Pardi, A. & Sousa, M.C. Structure and flexibility of the complete periplasmic domain of BamA: the protein insertion machine of the outer membrane. *Structure* **18**, 1492-1501 (2010).
39.	Knowles, T.J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T., Overduin, M. & Henderson, I.R. Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol. Microbiol.* **68**, 1216-1227 (2008).
40.	Zhang, H., Gao, Z.-Q., Hou, H.-F., Xu, J.-H., Li, L.-F., Su, X.-D. & Dong, Y.-H. High-resolution structure of a new crystal form of BamA POTRA4-5 from *Escherichia coli*. *Acta Crystallogr. Sect. F* **67**, 734-738 (2011).
41.	Noinaj, N., Kuszak, A.J., Gumbart, J.C., Lukacik, P., Chang, H., Easley, N.C., Lithgow, T. & Buchanan, S.K. Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501**, 385–390 (2013).
42.	Ni, D., Wang, Y., Yang, X., Zhou, H., Hou, X., Cao, B., Lu, Z., Zhao, X., Yang, K. & Huang, Y. Structural and functional analysis of the β-barrel domain of BamA from *Escherichia coli*. *FASEB J.* (2014).
43.	Dong, C., Yang, X., Hou, H.F., Shen, Y.Q. & Dong, Y.H. Structure of *Escherichia coli* BamB and its interaction with POTRA domains of BamA. *Acta Crystallogr. Sect. D Biol. Cryst.* **68**, 1134-1139 (2012).
44.	Kim, K.H. & Paetzel, M. Crystal structure of *Escherichia coli* BamB, a lipoprotein component of the β-barrel assembly machinery complex. *J. Mol. Biol.* **406**, 667-678 (2011).

45. Noinaj, N., Fairman, J.W. & Buchanan, S.K. The crystal structure of BamB suggests interactions with BamA and its role within the BAM complex. *J. Mol. Biol.* **407**, 248-260 (2011).

46. Jansen, K.B., Baker, S.L. & Sousa, M.C. Crystal structure of BamB from *Pseudomonas aeruginosa* and functional evaluation of its conserved structural features. *PloS one* **7**, e49749 (2012).

47. Heuck, A., Schleiffer, A. & Clausen, T. Augmenting β-augmentation: structural basis of how BamB binds BamA and may support folding of outer membrane proteins. *J. Mol. Biol.* **406**, 659-66 (2011).

48. Ruiz, N., Falcone, B., Kahne, D. & Silhavy, T.J. Chemical conditionality: a genetic strategy to probe organelle assembly. *Cell* **121**, 307-317 (2005).

49. Malinverni, J.C., Werner, J., Kim, S., Sklar, J.G., Kahne, D., Misra, R. & Silhavy, T.J. YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol. Microbiol.* **61**, 151-164 (2006).

50. Sklar, J.G., Wu, T., Gronenberg, L.S., Malinverni, J.C., Kahne, D. & Silhavy, T.J. Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proc. Nat. Acad. Sci. USA* **104**, 6400-6405 (2007).

51. Onufryk, C., Crouch, M.-L., Fang, F.C. & Gross, C.A. Characterization of six lipoproteins in the σ^E regulon. *J. Bacteriol.* **187**, 4552-4561 (2005).

52. Webb, C.T., Heinz, E. & Lithgow, T. Evolution of the β-barrel assembly machinery. *Trends Microbiol.* **20**, 612-620 (2012).

53. Albrecht, R. & Zeth, K. Structural basis of outer membrane protein biogenesis in bacteria. *J. Biol. Chem.* **286**, 27792-27803 (2011).

54. Sandoval, C.M., Baker, S.L., Jansen, K., Metzner, S.I. & Sousa, M.C. Crystal structure of BamD: an essential component of the β-barrel assembly machinery of Gram-negative bacteria. *J. Mol. Biol.* **409**, 348-357 (2011).

55. D'Andrea, L.D. & Regan, L. TPR proteins: the versatile helix. *Trends Biochem. Sci.* **28**, 655-662 (2003).

56. Gatto, G.J., Jr., Geisbrecht, B.V., Gould, S.J. & Berg, J.M. Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat. Struct. Mol. Biol.* **7**, 1091-1095 (2000).

57. Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F.U. & Moarefi, I. Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell* **101**, 199-210 (2000).

58. Warner, L.R., Varga, K., Lange, O.F., Baker, S.L., Baker, D., Sousa, M.C. & Pardi, A. Structure of the BamC two-domain protein obtained by Rosetta with a limited NMR data set. *J. Mol. Biol.* **411**, 83-95 (2011).

59. Webb, C.T., Selkrig, J., Perry, A.J., Noinaj, N., Buchanan, S.K. & Lithgow, T. Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *J. Mol. Biol.* **422**, 545-555 (2012).

60. Kim, K.H., Aulakh, S. & Paetzel, M. Crystal structure of the β-barrel assembly machinery BamCD complex. *J. Biol. Chem.* **286**, 39116-39121 (2011).

61. Kim, K.H., Kang, H.-S., Okon, M., Escobar-Cabrera, E., McIntosh, L.P. & Paetzel, M. Structural characterization of *Escherichia coli* BamE, a lipoprotein component of the β-barrel assembly machinery complex. *Biochemistry* **50**, 1081-1090 (2011).

62. Knowles, T.J., Browning, D.F., Jeeves, M., Maderbocus, R., Rajesh, S., Sridhar, P., Manoli, E., Emery, D., Sommer, U., Spencer, A., Leyton, D.L., Squire, D., Chaudhuri, R.R., Viant, M.R., Cunningham, A.F., Henderson, I.R. & Overduin, M. Structure and function of BamE within the outer membrane and the β-barrel assembly machine. *EMBO reports* **12**, 123-128 (2011).

63. Remaut, H., Tang, C., Henderson, N.S., Pinkner, J.S., Wang, T., Hultgren, S.J., Thanassi, D.G., Waksman, G. & Li, H. Fiber formation across the bacterial outer nembrane by the chaperone/usher pathway. *Cell* **133**, 640-652 (2008).

64. Phan, G., Remaut, H., Wang, T., Allen, W.J., Pirker, K.F., Lebedev, A., Henderson, N.S., Geibel, S., Volkan, E., Yan, J., Kunze, M.B.A., Pinkner, J.S., Ford, B., Kay, C.W.M., Li, H., Hultgren, S.J., Thanassi, D.G. & Waksman, G. Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature* **474**, 49-53 (2011).

65. Robert, V., Volokhina, E.B., Senf, F., Bos, M.P., Gelder, P.V. & Tommassen, J. Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol.* **4**, 1984-1995 (2006).

66. Eppens, E.F., Nouwen, N. & Tommassen, J. Folding of a bacterial outer membrane protein during passage through the periplasm. *EMBO J.* **16**, 4295-4301 (1997).

67. Leonard-Rivera, M. & Misra, R. Conserved residues of the putative L6 loop of *Escherichia coli* BamA play a critical role in the assembly of β-barrel outer membrane proteins, including BamA itself. *J. Bacteriol.* **194**, 4662-4668 (2012).

68. Clantin, B., Delattre, A.-S., Rucktooa, P., Saint, N., Méli, A.C., Locht, C., Jacob-Dubuisson, F. & Villeret, V. Structure of the membrane protein FhaC: a member of the Omp85-TpsB transporter superfamily. *Science* **317**, 957-961 (2007).

69. Rigel, N.W., Ricci, D.P. & Silhavy, T.J. Conformation-specific labeling of BamA and suppressor analysis suggest a cyclic mechanism for β-barrel assembly in *Escherichia coli*. *Proc. Nat. Acad. Sci. USA* **110**, 5151-5156 (2013).

70. Tamm, L.K., Hong, H. & Liang, B. Folding and assembly of beta-barrel membrane proteins. *Biochim. Biophys. Acta* **1666**, 250-263 (2004).

71. Renault, M., Cukkemane, A. & Baldus, M. Solid-state NMR spectroscopy on complex biomolecules. *Angew. Chem. Int. Ed.* **49**, 8346-8357 (2010).

72. Hong, M., Zhang, Y. & Hu, F. Membrane protein structure and dynamics from NMR spectroscopy. *Annu. Rev. Phys. Chem.* **63**, 1-24 (2012).

73. McDermott, A. Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR. *Annu. Rev. Biophys.* **38**, 385-403 (2009).

74. Ader, C., Schneider, R., Hornig, S., Velisetty, P., Wilson, E.M., Lange, A., Giller, K., Ohmert, I., Martin-Eauclaire, M.-F., Trauner, D., Becker, S., Pongs, O. & Baldus, M. A structural link between inactivation and block of a K+ channel. *Nat. Struct. Mol. Biol.* **15**, 605-612 (2008).

75. van der Cruijsen, E.A.W., Nand, D., Weingarth, M., Prokofyev, A., Hornig, S., Cukkemane, A.A., Bonvin, A.M.J.J., Becker, S., Hulse, R.E., Perozo, E., Pongs, O. & Baldus, M. Importance of lipid-pore loop interface for potassium channel structure and function. *Proc. Nat. Acad. Sci. USA* **110**, 13008-13013 (2013).

76. Lange, A., Giller, K., Hornig, S., Martin-Eauclaire, M.-F., Pongs, O., Becker, S. & Baldus, M. Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature* **440**, 959-962 (2006).

77. Weingarth, M., Prokofyev, A., van der Cruijsen, E.A.W., Nand, D., Bonvin, A.M.J.J., Pongs, O. & Baldus, M. Structural determinants of specific lipid binding to potassium channels. *J. Am. Chem. Soc.* **135**, 3983-3988 (2013).

78. Herzfeld, J. & Lansing, J.C. Magnetic resonance studies of the bacteriorhodopsin pump cycle. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 73-95 (2002).

79. Wang, S., Munro, R.A., Shi, L., Kawamura, I., Okitsu, T., Wada, A., Kim, S.-Y., Jung, K.-H., Brown, L.S. & Ladizhansky, V. Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. *Nat. Methods* **10**, 1007-1012 (2013).

80. Shahid, S.A., Bardiaux, B., Franks, W.T., Krabben, L., Habeck, M., van Rossum, B.-J. & Linke, D. Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. *Nat. Methods* **9**, 1212-1217 (2012).

81. Renault, M., Tommassen-van Boxtel, R., Bos, M.P., Post, J.A., Tommassen, J. & Baldus, M. Cellular solid-state nuclear magnetic resonance spectroscopy. *Proc. Nat. Acad. Sci. USA* **109**, 4863–4868 (2012).

82. Renault, M., Pawsey, S., Bos, M.P., Koers, E.J., Nand, D., Tommassen-van Boxtel, R., Rosay, M., Tommassen, J., Maas, W.E. & Baldus, M. Solid-State NMR spectroscopy on cellular preparations enhanced by Dynamic Nuclear Polarization. *Angew. Chem. Int. Ed.* **51**, 2998-3001 (2012).

83. Reckel, S., Lopez, J.J., Löhr, F., Glaubitz, C. & Dötsch, V. In-cell solid-state NMR as a tool to study proteins in large complexes. *ChemBioChem* **13**, 534-537 (2012).

84. Andrew, E.R., Bradbury, A. & Eades, R.G. Nuclear magnetic resonance spectra from a crystal rotated at high speed. *Nature* **182**, 1659-1659 (1958).

85. Lowe, I.J. Free Induction Decays of rotating solids. *Phys. Rev. Lett.* **2**, 285-287 (1959).

86. Linser, R., Dasari, M., Hiller, M., Higman, V., Fink, U., Lopez del Amo, J.-M., Markovic, S., Handel, L., Kessler, B., Schmieder, P., Oesterhelt, D., Oschkinat, H. & Reif, B. Proton-detected solid-state NMR spectroscopy of fibrillar and membrane proteins. *Angew. Chem. Int. Ed.* **50**, 4508-4512 (2011).

87. Zhou, D.H., Nieuwkoop, A.J., Berthold, D.a., Comellas, G., Sperling, L.J., Tang, M., Shah, G.J., Brea, E.J., Lemkau, L.R. & Rienstra, C.M. Solid-state NMR analysis of membrane proteins and protein aggregates by proton detected spectroscopy. *J. Biomol. NMR* **54**, 291-305 (2012).

88. Huber, M., Hiller, S., Schanda, P., Ernst, M., Böckmann, A., Verel, R. & Meier, B.H. A proton-detected 4D solid-state NMR experiment for protein structure determination. *ChemPhysChem* **12**, 915-918 (2011).

89. Knight, M.J., Webber, A.L., Pell, A.J., Guerry, P., Barbet-Massin, E., Bertini, I., Felli, I.C., Gonnelli, L., Pierattelli, R., Emsley, L., Lesage, A., Herrmann, T. & Pintacuda, G. Fast resonance assignment and fold determination of human Superoxide Dismutase by high-resolution proton-detected solid-state MAS NMR spectroscopy. *Angew. Chem.* **123**, 11901-11905 (2011).

90. Andronesi, O.C., Becker, S., Seidel, K., Heise, H., Young, H.S. & Baldus, M. Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **127**, 574-581 (2005).

91. Renault, M., Bos, M.P., Tommassen, J. & Baldus, M. Solid-state NMR on a large multidomain integral membrane protein: the outer membrane protein assembly factor BamA. *J. Am. Chem. Soc.* **133**, 4175-4177 (2011).

92. Hartmann, S.R. & Hahn, E.L. Nuclear double resonance in the rotating frame. *Phys. Rev.* **128**, 2042-2053 (1962).

93. Pines, A., Gibby, M.G. & Waugh, J.S. Nuclear double resonance in the rotating frame. *J. Chem. Phys.* **59**, 569-572 (1973).

94. Suter, D. & Ernst, R.R. Spin diffusion in resolved solid-state NMR spectra. *Phys. Rev. B* **32**, 5608-5627 (1985).

95. Vanderhart, D.L. Natural-abundance ^{13}C-^{13}C spin exchange in rigid crystalline organic solids. *J. Magn. Reson.* **72**, 13-47 (1987).

96. Morris, G.A. & Freeman, R. Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* **101**, 760-762 (1979).

97. de Jong, R.N., Daniels, M.A., Kaptein, R. & Folkers, G.E. Enzyme free cloning for high throughput gene cloning and expression. *J. Struct. Funct. Genomics* **7**, 109-118 (2006).

98. Zech, S.G., Wand, a.J. & McDermott, A.E. Protein structure determination by high-resolution solid-state NMR spectroscopy: application to microcrystalline ubiquitin. *J. Am. Chem. Soc.* **127**, 8618-8626 (2005).

99. Weingarth, M., Demco, D.E., Bodenhausen, G. & Tekely, P. Improved magnetization transfer in solid-state NMR with fast magic angle spinning. *Chem. Phys. Lett.* **469**, 342-348 (2009).

100. Fung, B.M., Khitrin, a.K. & Ermolaev, K. An improved broadband decoupling sequence for liquid crystals and solids. *J. Magn. Reson.* **142**, 97-101 (2000).

101. Baldus, M., Petkova, A.T., Herzfeld, J. & Griffin, R.G. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* **95**, 1197-1207 (1998).

102. Shaka, A.J., Barker, P.B. & Freeman, R. Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Reson.* **64**, 547-552 (1985).

103. Weingarth, M., Bodenhausen, G. & Tekely, P. Low-power decoupling at high spinning frequencies in high static fields. *J. Magn. Reson.* **199**, 238-241 (2009).

104. Weingarth, M., Bodenhausen, G. & Tekely, P. Probing the quenching of rotary resonance by PISSARRO decoupling. *Chem. Phys. Lett.* **502**, 259-265 (2011).

105. Zhou, D.H. & Rienstra, C.M. High-performance solvent suppression for proton detected solid-state NMR. *J. Magn. Reson.* **192**, 167-172 (2008).

106. Goddard, T.G. & Kneller, D.G. SPARKY 3, University of California, San Francisco.

107. Rao, N.S., Legault, P., Muhandiram, D.R., Greenblatt, J., Battiste, J.L., Williamson, J.R. & Kay, L.E. NMR pulse schemes for the sequential assignment of arginine side-chain Hε-Protons. *J. Magn. Reson., Ser B* **113**, 272-276 (1996).

108. 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* **44**, 213-223 (2009).

109. Tollinger, M., Skrynnikov, N.R., Mulder, F.A.A., Forman-Kay, J.D. & Kay, L.E. Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.* **123**, 11341-11352 (2001).

110. Mulder, F.A.A., Skrynnikov, N.R., Hon, B., Dahlquist, F.W. & Kay, L.E. Measurement of slow (μs-ms) time scale dynamics in protein side chains by ^{15}N relaxation dispersion NMR spectroscopy:

application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc.* **123**, 967-975 (2001).

111. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. & Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293 (1995).

112. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435-447 (2008).

113. Oostenbrink, C., Villa, A., Mark, A.E. & van Gunsteren, W.F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **25**, 1656-1676 (2004).

114. Berger, O., Edholm, O. & Jähnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* **72**, 2002-2013 (1997).

115. Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., Pupko, T. & Ben-Tal, N. ConSurf: using evolutionary data to raise testable hypotheses about protein function. *Israel J. Chem.* **53**, 199-206 (2013).

116. de Vries, S.J., van Dijk, M. & Bonvin, A.M. The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* **5**, 883-897 (2010).

117. Hubbard, S.J. & Thornton, J.M. NACCESS. (Department of Biochemistry and Molecular Biology, University College London, 1993).

118. Ricci, D.P., Hagan, C.L., Kahne, D. & Silhavy, T.J. Activation of the *Escherichia coli* β-barrel assembly machine (Bam) is required for essential components to interact properly with substrate. *Proc. Nat. Acad. Sci. USA* **109**, 3487–3491 (2012).

119. Misra, R. Assembly of the β-barrel outer membrane proteins in gram-negative bacteria, mitochondria, and chloroplasts. *ISRN Mol. Biol.* **2012**, 1-15 (2012).

120. Walther, D.M., Rapaport, D. & Tommassen, J. Biogenesis of beta-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cell. Mol. Life Sci.* **66**, 2789-2804 (2009).

121. Bennion, D., Charlson, E.S., Coon, E. & Misra, R. Dissection of β-barrel outer membrane protein assembly pathways through characterizing BamA POTRA 1 mutants of *Escherichia coli*. *Mol. Microbiol.* **77**, 1153-1171 (2010).

122. Bajaj, V.S., Mak-Jurkauskas, M.L., Belenky, M., Herzfeld, J. & Griffin, R.G. Functional and shunt states of bacteriorhodopsin resolved by 250 GHz dynamic nuclear polarization-enhanced solid-state NMR. *Proc. Nat. Acad. Sci. USA* **106**, 9244-9249 (2009).

123. Cady, S.D., Schmidt-Rohr, K., Wang, J., Soto, C.S., DeGrado, W.F. & Hong, M. Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* **463**, 689-692 (2010).

124. Ullrich, S.J., Hellmich, U.A., Ullrich, S. & Glaubitz, C. Interfacial enzyme kinetics of a membrane bound kinase analyzed by real-time MAS-NMR. *Nat. Chem. Biol.* **7**, 263-270 (2011).

125. Ni, Q.Z., Daviso, E., Can, T.V., Markhasin, E., Jawla, S.K., Swager, T.M., Temkin, R.J., Herzfeld, J. & Griffin, R.G. High frequency Dynamic Nuclear Polarization. *Acc. Chem. Res.* **46**, 1933-1941 (2013).

126. Koers, E.J., López-Deber, M.P., Weingarth, M., Nand, D., Hickman, D.T., Mlaki Ndao, D., Reis, P., Granet, A., Pfeifer, A., Muhs, A. & Baldus, M. Dynamic Nuclear Polarization NMR spectroscopy: revealing multiple conformations in lipid-anchored peptide vaccines. *Angew. Chem. Int. Ed.* **52**, 10905-10908 (2013).

127. Lorch, M., Faham, S., Kaiser, C., Weber, I., Mason, a.J., Bowie, J.U. & Glaubitz, C. How to prepare membrane proteins for solid-state NMR: A case study on the alpha-helical integral membrane protein diacylglycerol kinase from *E. coli*. *ChemBioChem* **6**, 1693-1700 (2005).

128. Iordanov, I., Renault, M., Réat, V., Bosshart, P.D., Engel, A., Saurel, O. & Milon, A. Dynamics of *Klebsiella pneumoniae* OmpA transmembrane domain: the four extracellular loops display restricted motion behavior in micelles and in lipid bilayers. *Biochim. Biophys. Acta - Biomembranes* **1818**, 2344-2353 (2012).

129. Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J.M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K.K., Lemak, A., Ignatchenko, A., Arrowsmith, C.H., Szyperski, T., Montelione, G.T., Baker, D. & Bax, A. Consistent blind protein structure generation from NMR chemical shift data. *Proc. Nat. Acad. Sci. USA* **105**, 4685-4690 (2008).

130. Seidel, K., Etzkorn, M., Schneider, R., Ader, C. & Baldus, M. Comparative analysis of NMR chemical shift predictions for proteins in the solid phase. *Solid State Nucl. Magn. Reson.* **35**, 235-242 (2009).
131. Liao, S.Y., Fritzsching, K.J. & Hong, M. Conformational analysis of the full-length M2 protein of the influenza A virus using solid-state NMR. *Protein Sci.* **22**, 1623-1638 (2013).
132. Nakamura, K. & Mizushima, S. Effects of heating in dodecyl sulfate solution on the conformation and electrophoretic mobility of isolated major outer membrane proteins from *Escherichia coli* K-12. *J. Biochem.* **80**, 1411-1422 (1976).
133. Heller, K.B. Apparent molecular weights of a heat-modifiable protein from the outer membrane of *Escherichia coli* in gels with different acrylamide concentrations. *J. Bacteriol.* **134**, 1181-1183 (1978).
134. Dekker, N., Merck, K., Tommassen, J. & Verheij, H.M. In vitro folding of *Escherichia coli* outer-membrane phospholipase A. *Eur. J. Biochem.* **232**, 214-219 (1995).
135. White, G.F., Racher, K.I., Lipski, A., Hallett, F.R. & Wood, J.M. Physical properties of liposomes and proteoliposomes prepared from *Escherichia coli* polar lipids. *Biochim. Biophys. Acta - Biomembranes* **1468**, 175-186 (2000).
136. de Leij, L. & Witholt, B. Structural heterogeneity of the cytoplasmic and outer membranes of *Escherichia coli*. *Biochim. Biophys. Acta - Biomembranes* **471**, 92-104 (1977).
137. Ward, R., Zoltner, M., Beer, L., El Mkami, H., Henderson, I.R., Palmer, T. & Norman, D.G. The orientation of a tandem POTRA domain pair, of the beta-barrel assembly protein BamA, determined by PELDOR spectroscopy. *Structure* **17**, 1187-1194 (2009).
138. Li, S., Su, Y., Luo, W. & Hong, M. Water-protein interactions of an arginine-rich membrane peptide in lipid bilayers investigated by solid-state Nuclear Magnetic Resonance spectroscopy. *J. Phys. Chem. B* **114**, 4063-4069 (2010).
139. Schneider, R., Ader, C., Lange, A., Giller, K., Hornig, S., Pongs, O., Becker, S. & Baldus, M. Solid-state NMR spectroscopy applied to a chimeric potassium channel in lipid bilayers. *J. Am. Chem. Soc.* **130**, 7427-7435 (2008).
140. Etzkorn, M., Martell, S., Andronesi, O.C., Seidel, K., Engelhard, M. & Baldus, M. Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed.* **46**, 459-462 (2007).
141. Li, Y., Berthold, D.a., Frericks, H.L., Gennis, R.B. & Rienstra, C.M. Partial ^{13}C and ^{15}N chemical-shift assignments of the disulfide-bond-forming enzyme DsbB by 3D magic-angle spinning NMR spectroscopy. *ChemBioChem* **8**, 434-442 (2007).
142. Ulmschneider, M.B. & Sansom, M.S.P. Amino acid distributions in integral membrane protein structures. *Biochim. Biophys. Acta - Biomembranes* **1512**, 1-14 (2001).
143. Higman, V.A., Flinders, J., Hiller, M., Jehle, S., Markovic, S., Fiedler, S., van Rossum, B.-J. & Oschkinat, H. Assigning large proteins in the solid state: a MAS NMR resonance assignment strategy using selectively and extensively ^{13}C-labelled proteins. *J. Biomol. NMR* **44**, 245-260 (2009).
144. Loquet, A., Lv, G., Giller, K., Becker, S. & Lange, A. ^{13}C spin dilution for simplified and complete solid-state NMR resonance assignment of insoluble biological assemblies. *J. Am. Chem. Soc.* **133**, 4722-4725 (2011).
145. Gradmann, S., Ader, C., Heinrich, I., Nand, D., Dittmann, M., Cukkemane, A., Dijk, M., Bonvin, A.J.J., Engelhard, M. & Baldus, M. Rapid prediction of multi-dimensional NMR data sets. *J. Biomol. NMR* **54**, 377-387 (2012).
146. Han, B., Liu, Y., Ginzinger, S.W. & Wishart, D.S. SHIFTX2: significantly improved protein chemical shift prediction. *J. Biomol. NMR* **50**, 43-57 (2011).
147. Gruss, F., Zähringer, F., Jakob, R.P., Burmann, B.M., Hiller, S. & Maier, T. The structural basis of autotransporter translocation by TamA. *Nat. Struct. Mol. Biol.* **20**, 1318-1320 (2013).
148. Lee, C.W.B. & Griffin, R.G. Two-dimensional ^{1}H/^{13}C heteronuclear chemical shift correlation spectroscopy of lipid bilayers. *Biophys. J.* **55**, 355-358 (1989).
149. Wang, Y. & Jardetzky, O. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci.* **11**, 852-861 (2002).
150. Kucerka, N., Liu, Y., Chu, N., Petrache, H.I., Tristram-Nagle, S. & Nagle, J.F. Structure of fully hydrated fluid phase DMPC and DLPC lipid bilayers using X-Ray scattering from oriented multilamellar arrays and from unilamellar vesicles. *Biophys. J.* **88**, 2626-2637 (2005).

151. Kucerka, N., Tristram-Nagle, S. & Nagle, J.F. Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains. *J. Membr. Biol.* **208**, 193-202 (2006).

152. Pogozheva, I.D., Tristram-Nagle, S., Mosberg, H.I. & Lomize, A.L. Structural adaptations of proteins to different biological membranes. *Biochim. Biophys. Acta - Biomembranes* **1828**, 2592-2608 (2013).

153. Tristram-Nagle, S., Liu, Y., Legleiter, J. & Nagle, J.F. Structure of gel phase DMPC determined by X-Ray diffraction. *Biophys. J.* **83**, 3324-3335 (2002).

154. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **67**, 593-656 (2003).

155. Collins, R.F., Beis, K., Dong, C., Botting, C.H., McDonnell, C., Ford, R.C., Clarke, B.R., Whitfield, C. & Naismith, J.H. The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli. Proc. Nat. Acad. Sci. USA* **104**, 2390-2395 (2007).

156. Graham, L.L., Harris, R., Villiger, W. & Beveridge, T.J. Freeze-substitution of gram-negative eubacteria: general cell morphology and envelope profiles. *J. Bacteriol.* **173**, 1623-1633 (1991).

157. Sinnige, T., Weingarth, M., Renault, M., Baker, L., Tommassen, J. & Baldus, M. Solid-State NMR studies of full-length BamA in lipid bilayers suggest limited overall POTRA mobility. *J. Mol. Biol.* **426**, 2009-2021 (2014).

158. Bax, A. & Grzesiek, S. Methodological advances in protein NMR. *Acc. Chem. Res.* **26**, 131-138 (1993).

159. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectr.* **34**, 93-158 (1999).

160. Baldus, M. Correlation experiments for assignment and structure elucidation of immobilized polypeptides under magic angle spinning. *Prog. Nucl. Magn. Reson. Spectr.* **41**, 1-47 (2002).

161. Pauli, J., Baldus, M., van Rossum, B., de Groot, H. & Oschkinat, H. Backbone and side-chain ^{13}C and ^{15}N signal assignments of the α-Spectrin SH3 domain by Magic Angle Spinning solid-state NMR at 17.6 Tesla. *ChemBioChem* **2**, 272-281 (2001).

162. Böckmann, A., Lange, A., Galinier, A., Luca, S., Giraud, N., Juy, M., Heise, H., Montserret, R., Penin, F. & Baldus, M. Solid state NMR sequential resonance assignments and conformational analysis of the 2 x 10.4 kDa dimeric form of the Bacillus subtilis protein Crh. *J. Biomol. NMR* **27**, 323-339 (2003).

163. Igumenova, T.I., McDermott, A.E., Zilm, K.W., Martin, R.W., Paulson, E.K. & Wand, A.J. Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* **126**, 6720-6727 (2004).

164. Igumenova, T.I., Wand, a.J. & McDermott, A.E. Assignment of the backbone resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* **126**, 5323-5331 (2004).

165. Seidel, K., Etzkorn, M., Heise, H., Becker, S. & Baldus, M. High-resolution solid-state NMR studies on uniformly [^{13}C,^{15}N]-labeled ubiquitin. *ChemBioChem* **6**, 1638-1647 (2005).

166. Sperling, L.J., Berthold, D.A., Sasser, T.L., Jeisy-Scott, V. & Rienstra, C.M. Assignment strategies for large proteins by Magic-Angle Spinning NMR: the 21-kDa disulfide-bond-forming enzyme DsbA. *J. Mol. Biol.* **399**, 268-282 (2010).

167. Wang, S., Shi, L., Okitsu, T., Wada, A., Brown, L.S. & Ladizhansky, V. Solid-state NMR ^{13}C and ^{15}N resonance assignments of a seven-transmembrane helical protein Anabaena Sensory Rhodopsin. *Biomol. NMR Assign.* **7**, 253-256 (2013).

168. Habenstein, B., Wasmer, C., Bousset, L., Sourigues, Y., Schütz, A., Loquet, A., Meier, B., Melki, R. & Böckmann, A. Extensive de novo solid-state NMR assignments of the 33 kDa C-terminal domain of the Ure2 prion. *J. Biomol. NMR* **51**, 235-243 (2011).

169. Seidel, K., Lange, A., Becker, S., Hughes, C.E., Heise, H. & Baldus, M. Protein solid-state NMR resonance assignments from ^{13}C,[13]C correlation spectroscopy. *Phys. Chem. Chem. Phys.* **6**, 5090-5093 (2004).

170. Loquet, A., Lv, G., Giller, K., Becker, S. & Lange, A. Resonance assignment of insoluble biological assemblies. *J. Am. Chem. Soc.*, 4722-4725 (2011).

171. Muir, T.W. Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **72**, 249-289 (2003).

172. Skrisovska, L., Schubert, M. & Allain, F.H.T. Recent advances in segmental isotope labeling of proteins: NMR applications to large proteins and glycoproteins. *J. Biomol. NMR* **46**, 51-65 (2010).

173. Michel, E., Skrisovska, L., Wüthrich, K. & Allain, F.H.T. Amino acid-selective segmental isotope labeling of multidomain proteins for structural biology. *ChemBioChem* **14**, 457-466 (2013).

174. Refaei, M., Combs, A., Kojetin, D., Cavanagh, J., Caperelli, C., Rance, M., Sapitro, J. & Tsang, P. Observing selected domains in multi-domain proteins via sortase-mediated ligation and NMR spectroscopy. *J. Biomol. NMR* **49**, 3-7 (2011).

175. Karagöz, G.E., Sinnige, T., Hsieh, O. & Rüdiger, S.G.D. Expressed protein ligation for a large dimeric protein. *Protein Eng. Des. Sel.* **24**, 495-501 (2011).

176. Minato, Y., Ueda, T., Machiyama, A., Shimada, I. & Iwaï, H. Segmental isotopic labeling of a 140-kDa dimeric multi-domain protein CheA from *Escherichia coli* by expressed protein ligation and protein trans-splicing. *J. Biomol. NMR* **53**, 191-207 (2012).

177. Reif, B. Ultra-high resolution in MAS solid-state NMR of perdeuterated proteins: implications for structure and dynamics. *J. Magn. Reson.* **216**, 1-12 (2012).

178. Asami, S., Schmieder, P. & Reif, B. High resolution ^{1}H-detected solid-state NMR spectroscopy of protein aliphatic resonances: access to tertiary structure information. *J. Am. Chem. Soc.* **132**, 15133-15135 (2010).

179. Sinnige, T., Daniëls, M., Baldus, M. & Weingarth, M. Proton clouds to measure long-range contacts between nonexchangeable side chain protons in solid-state NMR. *J. Am. Chem. Soc.* **136**, 4452-4455 (2014).

180. Dalbey, R.E., Wang, P. & Kuhn, A. Assembly of bacterial inner membrane proteins. *Annu. Rev. Biochem.* **80**, 161-187 (2011).

181. Ricci, D.P. & Silhavy, T.J. The Bam machine: A molecular cooper. *Biochim. Biophys. Acta - Biomembranes* **1818**, 1067-1084 (2012).

182. Bos, M.P., Robert, V. & Tommassen, J. Functioning of outer membrane protein assembly factor Omp85 requires a single POTRA domain. *EMBO reports* **8**, 1149-1154 (2007).

183. Tzeng, S.-R. & Kalodimos, C.G. Protein dynamics and allostery: an NMR view. *Curr. Opin. Struct. Biol.* **21**, 62-67 (2011).

184. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964-972 (2007).

185. Palmer III, A.G., Kroenke, C.D. & Loria, J.P. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.* **339**, 204-238 (2001).

186. Moon, C.P., Zaccai, N.R., Fleming, P.J., Gessmann, D. & Fleming, K.G. Membrane protein thermodynamic stability may serve as the energy sink for sorting in the periplasm. *Proc. Nat. Acad. Sci. USA* **110**, 4285-4290 (2013).

187. Kozjak, V., Wiedemann, N., Milenkovic, D., Lohaus, C., Meyer, H.E., Guiard, B., Meisinger, C. & Pfanner, N. An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane. *J. Biol. Chem.* **278**, 48520-48523 (2003).

188. Paschen, S.A., Waizenegger, T., Stan, T., Preuss, M., Cyrklaff, M., Hell, K., Rapaport, D. & Neupert, W. Evolutionary conservation of biogenesis of β-barrel membrane proteins. *Nature* **426**, 862-866 (2003).

189. Inoue, K. & Potter, D. The chloroplastic protein translocation channel Toc75 and its paralog OEP80 represent two distinct protein families and are targeted to the chloroplastic outer envelope by different mechanisms. *Plant J.* **39**, 354-365 (2004).

190. Fussenegger, M., Facius, D., Meier, J. & Meyer, T.F. A novel peptidoglycan-linked lipoprotein (ComL) that functions in natural transformation competence of Neisseria gonorrhoeae. *Mol. Microbiol.* **19**, 1095-1105 (1996).

191. Volokhina, E.B., Beckers, F., Tommassen, J. & Bos, M.P. The β-barrel outer membrane protein assembly complex of *Neisseria meningitidis*. *J. Bacteriol.* **191**, 7074-7085 (2009).

192. Ulrich, E.L., Akutsu, H., Doreleijers, J.F., Harano, Y., Ioannidis, Y.E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C.F., Tolmie, D.E., Kent Wenger, R., Yao, H. & Markley, J.L. BioMagResBank. *Nucleic Acids Res.* **36**, D402-D408 (2008).

193. de Vries, S.J. & Bonvin, A.M.J.J. How proteins get in touch: Interface prediction in the study of biomolecular complexes. *Curr. Protein Peptide Sci.* **9**, 394-406 (2008).

194. Sharma, M., Patel, G.J. & Kleinschmidt, J.H. Folding and lipid membrane interactions of BamD, an essential component of the β-barrel assembly machine from *Escherichia Coli*. in *2014 Biophysical Society Meeting* Vol. 106 265a (Biophysical Journal, 2014).

195. Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A.B., Riek, R. & Meier, B.H. Amyloid fibrils of the HET-s(218–289) prion form a β solenoid with a triangular hydrophobic core. *Science* **319**, 1523-1526 (2008).

196. Loquet, A., Sgourakis, N.G., Gupta, R., Giller, K., Riedel, D., Goosmann, C., Griesinger, C., Kolbe, M., Baker, D., Becker, S. & Lange, A. Atomic model of the type III secretion system needle. *Nature* **486**, 276-279 (2012).

197. Park, S.H., Das, B.B., Casagrande, F., Tian, Y., Nothnagel, H.J., Chu, M., Kiefer, H., Maier, K., De Angelis, A.a., Marassi, F.M. & Opella, S.J. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature* **491**, 779-783 (2012).

198. Reif, B., Jaroniec, C.P., Rienstra, C.M., Hohwy, M. & Griffin, R.G. ^{1}H-^{1}H MAS correlation spectroscopy and distance measurements in a deuterated peptide. *J. Magn. Reson.* **151**, 320-327 (2001).

199. Paulson, E.K., Morcombe, C.R., Gaponenko, V., Dancheck, B., Byrd, R.A. & Zilm, K.W. Sensitive high resolution inverse detection NMR spectroscopy of proteins in the solid state. *J. Am. Chem. Soc.* **125**, 15831-15836 (2003).

200. Chevelkov, V., Rehbein, K., Diehl, A. & Reif, B. Ultrahigh resolution in proton solid-state NMR spectroscopy at high levels of deuteration. *Angew. Chem. Int. Ed.* **45**, 3878-3881 (2006).

201. Zhou, D.H., Shah, G., Cormos, M., Mullen, C., Sandoz, D. & Rienstra, C.M. Proton-detected solid-state NMR spectroscopy of fully protonated proteins at 40 kHz magic-angle spinning. *J. Am. Chem. Soc.* **129**, 11791-801 (2007).

202. Schanda, P., Meier, B.H. & Ernst, M. Quantitative analysis of protein backbone dynamics in microcrystalline ubiquitin by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **132**, 15957-15967 (2010).

203. Ward, M.E., Shi, L., Lake, E., Krishnamurthy, S., Hutchins, H., Brown, L.S. & Ladizhansky, V. Proton-detected solid-state NMR reveals intramembrane polar networks in a seven-helical transmembrane protein proteorhodopsin. *J. Am. Chem. Soc.* **133**, 17434-17443 (2011).

204. Knight, M.J., Pell, A.J., Bertini, I., Felli, I.C., Gonnelli, L., Pierattelli, R., Herrmann, T., Emsley, L. & Pintacuda, G. Structure and backbone dynamics of a microcrystalline metalloprotein by solid-state NMR. *Proc. Nat. Acad. Sci. USA* **109**, 11095-11100 (2012).

205. Lewandowski, J.R., Dumez, J.N., Akbey, Ü., Lange, S., Emsley, L., Dumez, J.-n. & Oschkinat, H. Enhanced resolution and coherence lifetimes in the solid-state NMR spectroscopy of perdeuterated proteins under ultrafast magic-angle spinning. *J. Phys. Chem. Lett.* **2**, 2205-2211 (2011).

206. Gardner, K.H. & Kay, L.E. Production and incorporation of ^{15}N, ^{13}C, ^{2}H (^{1}H-δ1 methyl) isoleucine into proteins for multidimensional NMR studies. *J. Am. Chem. Soc.* **119**, 7599-7600 (1997).

207. Asami, S. & Reif, B. Proton-detected solid-state NMR spectroscopy at aliphatic sites: application to crystalline systems. *Acc. Chem. Res.* **46**, 2089-2097 (2013).

208. Lange, A., Luca, S. & Baldus, M. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J. Am. Chem. Soc.* **124**, 9704-9705 (2002).

209. Marchetti, A., Jehle, S., Felletti, M., Knight, M.J., Wang, Y., Xu, Z.Q., Park, A.Y., Otting, G., Lesage, A., Emsley, L., Dixon, N.E. & Pintacuda, G. Backbone assignment of fully protonated solid proteins by ^{1}H detection and ultrafast magic-angle-spinning NMR spectroscopy. *Angew. Chem. Int. Ed.* **51**, 10756-10759 (2012).

210. Cornilescu, G., Marquardt, J.L., Ottiger, M. & Bax, A. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **120**, 6836-6837 (1998).

211. Schubert, M., Manolikas, T., Rogowski, M. & Meier, B.H. Solid-state NMR spectroscopy of 10% ^{13}C labeled ubiquitin: spectral simplification and stereospecific assignment of isopropyl groups. *J. Biomol. NMR* **35**, 167-173 (2006).

212. Asami, S., Szekely, K., Schanda, P., Meier, B.H. & Reif, B. Optimal degree of protonation for ^{1}H-detection of aliphatic sites in randomly deuterated proteins as a function of the MAS frequency. *J. Biomol. NMR* **54**, 155-168 (2012).

213. Zorin, V.E., Brown, S.P. & Hodgkinson, P. Origins of linewidth in ^{1}H magic-angle spinning NMR. *J. Chem. Phys.* **125**, 144508 (2006).

214. Bayro, M.J., Huber, M., Ramachandran, R., Davenport, T.C., Meier, B.H., Ernst, M. & Griffin, R.G. Dipolar truncation in magic-angle spinning NMR recoupling experiments. *J. Chem. Phys.* **130**, 114506 (2009).

215. Grommek, A., Meier, B.H. & Ernst, M. Distance information from proton-driven spin diffusion under MAS. *Chem. Phys. Lett.* **427**, 404-409 (2006).

216. Weingarth, M., Masuda, Y., Takegoshi, K., Bodenhausen, G. & Tekely, P. Sensitive ^{13}C-^{13}C correlation spectra of amyloid fibrils at very high spinning frequencies and magnetic fields. *J. Biomol. NMR* **50**, 129-136 (2011).

217. Vijay-Kumar, S., Bugg, C.E. & Cook, W.J. Structure of ubiquitin refined at 1.8Å resolution. *J. Mol. Biol.* **194**, 531-544 (1987).

218. Shi, L., Kawamura, I., Jung, K.H., Brown, L.S. & Ladizhansky, V. Conformation of a seven-helical transmembrane photosensor in the lipid environment. *Angew. Chem. Int. Ed.* **50**, 1302-1305 (2011).

219. Lin, M.T., Sperling, L.J., Frericks Schmidt, H.L., Tang, M., Samoilova, R.I., Kumasaka, T., Iwasaki, T., Dikanov, S.A., Rienstra, C.M. & Gennis, R.B. A rapid and robust method for selective isotope labeling of proteins. *Methods* **55**, 370-378 (2011).

220. Hagan, C.L. & Kahne, D. The reconstituted *Escherichia coli* Bam complex catalyzes multiple rounds of β-barrel assembly. *Biochemistry* **50**, 7444-7446 (2011).

221. Patel, G.J. & Kleinschmidt, J.H. The lipid bilayer-inserted membrane protein BamA of *Escherichia coli* facilitates insertion and folding of Outer Membrane Protein A from its complex with Skp. *Biochemistry* **52**, 3974–3986 (2013).

222. Gessmann, D., Chung, Y.H., Danoff, E.J., Plummer, A.M., Sandlin, C.W., Zaccai, N.R. & Fleming, K.G. Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proc. Nat. Acad. Sci. USA* **111**, 5878–5883 (2014).

223. Salnikov, E., Rosay, M., Pawsey, S., Ouari, O., Tordo, P. & Bechinger, B. Solid-state NMR spectroscopy of oriented membrane polypeptides at 100 K with signal enhancement by dynamic nuclear polarization. *J. Am. Chem. Soc.* **132**, 5940-5941 (2010).

224. Linden, A.H., Franks, W.T., Akbey, Ü., Lange, S., van Rossum, B.-J. & Oschkinat, H. Cryogenic temperature effects and resolution upon slow cooling of protein preparations in solid state NMR. *J. Biomol. NMR* **51**, 283-292 (2011).

225. Ieva, R. & Bernstein, H.D. Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proc. Nat. Acad. Sci. USA* **106**, 19120-19125 (2009).

226. Ieva, R., Tian, P., Peterson, J.H. & Bernstein, H.D. Sequential and spatially restricted interactions of assembly factors with an autotransporter beta-domain. *Proc. Nat. Acad. Sci. USA* **108**, 383-391 (2011).

227. Sauri, A., Soprova, Z., Wickström, D., de Gier, J.-W., Van der Schors, R.C., Smit, A.B., Jong, W.S.P. & Luirink, J. The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease. *Microbiology* **155**, 3982-3991 (2009).

228. Lehr, U., Schütz, M., Oberhettinger, P., Ruiz-Perez, F., Donald, J.W., Palmer, T., Linke, D., Henderson, I.R. & Autenrieth, I.B. C-terminal amino acid residues of the trimeric autotransporter adhesin YadA of *Yersinia enterocolitica* are decisive for its recognition and assembly by BamA. *Mol. Microbiol.* **78**, 932-946 (2010).

229. Schmidt, E., Gath, J., Habenstein, B., Ravotti, F., Székely, K., Huber, M., Buchner, L., Böckmann, A., Meier, B.H. & Güntert, P. Automated solid-state NMR resonance assignment of protein microcrystals and amyloids. *J. Biomol. NMR* **56**, 243-254 (2013).

230. Paramasivam, S., Suiter, C.L., Hou, G., Sun, S., Palmer, M., Hoch, J.C., Rovnyak, D. & Polenova, T. Enhanced sensitivity by nonuniform sampling enables multidimensional MAS NMR spectroscopy of protein assemblies. *J. Phys. Chem. B* **116**, 7416-7427 (2012).

231. Lin, E.C. & Opella, S.J. Sampling scheme and compressed sensing applied to solid-state NMR spectroscopy. *J. Magn. Reson.* **237**, 40-48 (2013).

232. Franks, W.T., Kloepper, K.D., Wylie, B.J. & Rienstra, C.M. Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins. *J. Biomol. NMR* **39**, 107-131 (2007).

SUMMARY

The β-barrel assembly machinery (BAM) is a protein complex in the outer membrane of Gram-negative bacteria that mediates folding and insertion of outer membrane proteins (OMPs) (**Chapter 1**). These OMPs fulfill critical roles for survival of the bacteria: they serve e.g. as porins that allow diffusion of nutrients into the periplasm, as transporters for active uptake of nutrients, as virulence factors and as enzymes that function in maintenance of the outer membrane itself. Therefore it is not surprising that the main component of the BAM complex, BamA, is highly conserved in Gram-negative bacteria and also has homologues in the outer membranes of mitochondria and chloroplasts. The structures of the BAM components are known, yet it is unclear how BamA functions and what role the lipoproteins BamB-E play. Solid-state NMR (ssNMR) spectroscopy is very suitable to study membrane proteins in their native environment, i.e. in lipid bilayers. In this thesis, we described ssNMR studies on the dynamics of BamA and its interactions with the lipid bilayer and the lipoproteins BamCDE. The experimental procedures were reported in **Chapter 2**.

BamA consists of a membrane-embedded β-barrel domain and five periplasmic POTRA domains, which have been shown to play a role in substrate interactions. In **Chapter 3**, we set out to investigate global motion of the POTRA domains in context of full-length BamA reconstituted in proteoliposomes. Using dipolar-based as well as scalar-based pulse sequences, in combination with different BamA constructs, we discovered that POTRA 1-5 were relatively rigid and did not experience global motion on the µs-ms timescale or faster. In addition, we studied the interplay between BamA and the lipid bilayer. We found that the structure and dynamics of BamA reconstituted in lipids of different chain lengths and saturation levels were highly similar, the only variation lying in certain residues in β-strand conformation. Apparently, conformational plasticity of the BamA β-barrel allows it to accommodate to different lipid bilayers. Interestingly, electron microscopy on the ssNMR samples revealed notches in the BamA proteoliposomes, suggesting that BamA can distort the lipid bilayer, which might play a role in substrate insertion into the membrane.

To obtain more residue-specific information on the transmembrane (TM) domain of BamA, ssNMR assignments are a prerequisite, which is not trivial due to its large size of 390 residues. In **Chapter 4** we described different reverse and forward labeling strategies leading to the first chemical shift assignments for the membrane-embedded BamA TM domain. Interestingly, we succeeded to tentatively

assign a number of residues from the long extracellular loop 6. This loop contains a conserved motif that is essential for β-barrel assembly and it has been speculated to undergo conformational changes. Possibly, these only occur upon binding of the substrate. On the other hand, we did not succeed to assign residues at the proposed lateral gate between β-strands 1 and 16, suggesting that this region may be subject to dynamics.

In **Chapter 5**, we went on to investigate local dynamics within the POTRA 5 (P5) domain, which is positioned at the interface of the BamA β-barrel and the lipid bilayer and is required for interaction with BamD. Solution NMR spectra of the isolated P5 domain compared to the POTRA 4-5 (P4P5) tandem construct revealed that P5 depends on P4 for its correct folding. This finding was confirmed in ssNMR studies on membrane-embedded BamA constructs. Moreover, we found that a highly conserved patch of residues on P5 experienced local conformational exchange on the μs-ms timescale, as evidenced by relaxation dispersion measurements in solution and analysis of signal intensities in dipolar-based ssNMR experiments. Interestingly, mutation of the conserved residue E373 led to wide-spread chemical shift perturbations in solution, whereas in ssNMR we found an increase in local dynamics of the E373K mutant. *In vivo* this mutant has been shown deleterious at normal growth temperatures and unable to bind to BamD, suggesting that the plasticity of P5 might play a role in the interaction with BamD and, possibly, the substrate.

In addition to BamA, BamD is essential for viability in *Escherichia coli*, yet its function is unclear. BamD binds BamA through the P5 domain, but the molecular basis for the interaction has not been revealed. In **Chapter 6**, we investigated the interaction between these two essential components of the β-barrel assembly machinery using a combination of NMR spectroscopy and docking. Solution NMR titrations of P4P5 with BamD and the BamC^{UN}D complex did not yield evidence for a strong interaction. We speculated that tethering of the proteins to the membrane by means of the β-barrel of BamA and the lipid anchor of BamD might be crucial for the interaction. When we co-reconstituted BamA and BamD in lipid bilayers, certain residues on P5 showed line broadening in the ssNMR spectra, indicative of conformational exchange on an intermediate NMR timescale. Only when we co-reconstituted BamA with the BamCDE lipid-anchored sub-complex, a strong interaction was observed. It was mediated by P5 residues that mapped to the dynamic POTRA 5 region that we identified before. Additional shifts were compatible with residues from the BamA transmembrane (TM) domain that reside in periplasmic loops close to POTRA 5. Docking yielded several possible models for the BamA-BamD interaction in which electrostatic interactions played a dominant role. Remarkably, a conserved patch on the surface of BamD that we expected to

form the BamA binding site remained accessible in the selected docking models, suggesting that it might be involved in another process, i.e. substrate binding.

Evidently, ssNMR methods need to be further developed in order to advance the study of large membrane proteins such as BamA. Proton-detected methods bear the promise of higher sensitivity and resolution. However, thus far methods have relied on very sparse labeling schemes with protonation exclusively at amide or methyl sites. Hence, most side chains as well as contacts among them, which are crucial for three-dimensional structure determination, remain inaccessible. In **Chapter 7**, we introduced a new approach termed "proton-clouds", in which fully protonated amino acids are incorporated in a deuterated background. We first demonstrated the method on ubiquitin and showed that the proton line widths are significantly better than for a fully protonated sample. In addition, we could employ the remaining proton-proton dipolar couplings for magnetization transfer by means of spin diffusion, which yielded a number of contacts among the proton-clouds. As a proof of principle, we prepared a proton-cloud sample of membrane-embedded BamA and demonstrated efficient magnetization transfer. Further improvements in sample preparation, e.g. reconstitution in deuterated lipids, and the use of higher dimensional experiments are expected to extend the application of proton-clouds to the study of large membrane proteins such as BamA.

In **Chapter 8**, the results were jointly discussed in the context of existing literature and perspectives for future research were presented, including substrate interaction studies and investigation of the BAM complex in its native cellular environment.

SAMENVATTING

De β-barrel assembly machinery (BAM) is een eiwitcomplex in het buitenmembraan van Gram-negatieve bacteriën, dat een rol speelt in de vouwing en insertie van buitenmembraan-eiwitten (outer membrane proteins, OMPs) (**Hoofdstuk 1**). Deze OMPs zijn essentieel voor het overleven van de bacterie: ze vormen bijvoorbeeld poriën waardoor voedingsstoffen het periplasma in kunnen diffunderen, transporteiwitten voor de actieve opname van voedingsstoffen, virulentiefactoren en enzymen die het buitenmembraan zelf kunnen modificeren. De belangrijkste component van het BAM complex, BamA, is dan ook een sterk geconserveerd eiwit in Gram-negatieve bacteriën en er zijn zelfs homologe eiwitten in de buitenmembranen van mitochondria en chloroplasten geïdentificeerd. De structuren van alle BAM onderdelen zijn bekend, maar desondanks begrijpt men niet hoe BamA functioneert en welke rol de gelipideerde eiwitten BamB-E spelen. Vaste-stof NMR (solid-state NMR, ssNMR) spectroscopie is erg geschikt om membraaneiwitten in hun natieve omgeving, bestaande uit een lipide bilaag, te bestuderen. In dit proefschrift werd ssNMR onderzoek beschreven naar de dynamica van BamA en de interacties van BamA met de lipide bilaag en de gelipideerde eiwitten BamCDE. De experimentele procedures zijn opgenomen in **Hoofdstuk 2**.

BamA bestaat uit een membraan-geïntegreerde β-barrel en vijf periplasmatische POTRA domeinen, die mogelijk een rol spelen in substraat-interacties. In **Hoofdstuk 3** onderzochten we de globale beweging van de POTRA domeinen in de context van het gehele BamA-eiwit gereconstitueerd in proteoliposomen. Met behulp van dipolair- en scalair-gebaseerde pulssequenties, gecombineerd met verschillende BamA-constructen, ontdekten we dat POTRA 1-5 relatief rigide waren en niet als geheel bewogen op een tijdsschaal van μs-ms of sneller. Verder bestudeerden we het samenspel tussen BamA en de lipide bilaag. We constateerden dat de structuur en dynamica van BamA in bilagen gevormd door lipiden van verschillende lengte en verzadiging vergelijkbaar waren. Slechts enkele residuen in β-strand-conformatie waren variabel. Blijkbaar is de β-barrel van BamA in staat zich aan de lipide bilaag aan te passen door een zekere mate van plasticiteit. Elektronenmicroscopie toonde inkepingen in de BamA proteoliposomen, wat erop wijst dat BamA de lipide bilaag kan verstoren. Mogelijk is dit een mechanisme om het substraat in het membraan te kunnen inserteren.

Om meer residu-specifieke informatie over het transmembraandomein (TM domein) van BamA te verkrijgen, zijn toekenningen van de ssNMR resonanties

noodzakelijk. Gezien de grootte van het domein (390 residuen) is dit niet eenvoudig. In **Hoofdstuk 4** beschreven we verschillende labeling-strategieën waarin gelabelde of ongelabelde aminozuren werden toegevoegd, waarmee we de eerste chemical shifts voor het membraan-geïnserteerde BamA TM domein konden toekennen. We slaagden erin om een aantal residuen van de lange extracellulaire lus 6 toe te kennen. Deze lus bevat een geconserveerd motief dat essentieel is voor β-barrel assembly en zou conformationele veranderingen kunnen ondergaan. Misschien gebeurt dit enkel onder invloed van het substraat en daarom niet in onze experimenten. Aan de andere kant konden we geen residuen toekennen in het gebied van de veronderstelde laterale opening tussen β-strands 1 en 16, wat eventueel door dynamica veroorzaakt zou kunnen zijn.

In **Hoofdstuk 5** vervolgden we het onderzoek naar dynamica en richtten we ons op het POTRA 5 (P5) domein, dat zich op de rand van de BamA β-barrel en het membraan bevindt en nodig is voor de interactie met BamD. Uit de vergelijking van vloeistof NMR spectra van geïsoleerd P5 en het POTRA 4-5 (P4P5) tandem construct, leidden we af dat P5 van P4 afhangt voor correcte vouwing. Met ssNMR experimenten aan membraan-geïntegreerde BamA constructen konden we dit bevestigen. Verder ontdekten we lokale conformationele uitwisseling op een μs-ms tijdsschaal van residuen in P5 die samen een sterk geconserveerd oppervlak op het domein vormen. Bewijs hiervoor kwam van relaxatie-dispersie metingen in vloeistof NMR en analyse van de signaalintensiteiten in dipolair-gebaseerde ssNMR experimenten. Mutatie van het geconserveerde residu E373 veroorzaakte verstrekkende verstoringen van de chemical shifts in vloeistof NMR, terwijl de E373K mutant in ssNMR de lokale dynamica leek te versterken. Deze mutant kan *in vivo* niet overleven onder normale groeitemperatuur en verliest de binding met BamD, hetgeen suggereert dat de plasticiteit van P5 een rol zou kunnen spelen in de interactie met BamD en wellicht ook met het substraat.

Naast BamA is BamD essentieel voor het overleven van *Escherichia coli*. De functie van dit eiwit is echter onbekend. BamD bindt aan BamA via het P5 domein, maar een moleculaire omschrijving van deze interactie ontbreekt. In **Hoofdstuk 6** onderzochten we de interactie tussen de twee essentiële componenten van de β-barrel assembly machinery met behulp van NMR spectroscopie en docking. Vloeistof NMR titraties van P4P5 met BamD en het BamC^{UN}D complex leverden geen bewijs voor een sterke interactie. We vermoedden dat binding van de eiwitten aan het membraan, door middel van de β-barrel van BamA en het lipide-anker van BamD, wellicht cruciaal kon zijn voor de interactie. Daarom co-reconstitueerden we BamA en BamD in lipide bilagen, maar dit leidde slechts tot verbreding van enkele P5 signalen in het ssNMR spectrum, wat op een interactie met uitwisseling op een intermediaire tijdsschaal duidt. Pas toen we BamA met het gelipideerde BamCDE

sub-complex co-reconstitueerden, vonden we een sterke interactie waarbij residuen van P5 betrokken waren die deel uitmaakten van het dynamische oppervlak dat we eerder hadden aangetoond, alsook residuen van het BamA TM domein die waarschijnlijk van de periplasmatische lussen afkomstig waren. Door gebruik te maken van eiwit-eiwit docking verkregen we mogelijke modellen voor de BamA-BamD interactie, waarin elektrostatische interacties dominant bleken. Het viel ons op dat een geconserveerd oppervlak op BamD, waarvan we verwachtten dat het bij de binding aan BamA betrokken zou zijn, grotendeels toegankelijk bleef in de docking-modellen. De mogelijkheid bestaat dat dit bij een ander proces, bijvoorbeeld interactie met het substraat, betrokken is.

Het is duidelijk dat ssNMR methoden verder ontwikkeld moeten worden om voorwaarts te komen in het onderzoek aan grote membraaneiwitten zoals BamA. Proton-gedetecteerde methoden bieden in principe de mogelijkheid om gevoeligheid en resolutie te vergroten. Tot dusver hingen deze methoden echter af van zeer schaarse labeling, waarbij enkel de amiden of methylgroepen geprotoneerd zijn. Dit houdt in dat de meeste zijketens en contacten tussen zijketens niet toegankelijk zijn, terwijl dit essentieel is voor driedimensionale structuurbepaling. In **Hoofdstuk 7** introduceerden we een nieuwe methode, "proton-clouds" genaamd, waarin volledig geprotoneerde aminozuren opgenomen worden in een gedeutereerde achtergrond. We demonstreerden de methode aan de hand van ubiquitine en toonden aan dat de protonenresolutie duidelijk beter was dan voor een volledig geprotoneerd monster. Verder konden we van de resterende dipolaire koppelingen gebruik maken voor magnetisatie-overdacht tussen de protonen via spindiffusie, waarmee we diverse contacten tussen de proton-clouds konden waarnemen. Om aan te tonen dat de methode ook voor grote membraaneiwitten geschikt is, maakten we een proton-clouds monster van BamA in lipiden en namen ook daar efficiënte magnetisatie-overdracht waar. Verdere ontwikkelingen in het prepareren van de monsters, bijvoorbeeld met gedeutereerde lipiden, en het gebruik van drie- of vierdimensionale experimenten zullen de toepassing van proton-clouds voor het bestuderen van grote membraaneiwitten verder uitbreiden.

In **Hoofdstuk 8** bespraken we de resultaten gezamenlijk in het kader van bestaande literatuur. Ook werden perspectieven voor toekomstig onderzoek gepresenteerd, zoals het karakteriseren van substraat-interacties en het bestuderen van het BAM complex in een natieve, cellulaire omgeving.

ACKNOWLEDGEMENTS

A PhD is like a marathon, my supervisor told me. Many people helped me to get to the finish line. First of all, I would like to thank Marc Baldus, my supervisor, for his creativity and enthusiasm and for giving me the freedom to explore the project and different techniques. Thanks to the other professors in the group, Rolf Boelens and Alexandre Bonvin, for the chance to use solution NMR and computational methods. I would also like to thank my previous mentors, Gregg Siegal, Stefan Rüdiger and Matthias Rief, for teaching me and inspiring me to continue in science.

Thanks to the members of the solid-state NMR group: Markus, I learned a lot from you about NMR and you always made time to help me out and to discuss the project. Elwin, thanks for fun in the office. Lindsay, thanks for good advice on pretty much anything. Eline, thanks for motivation to do sports and for training us for the Singelloop. Lindsay and Eline, I also enjoyed our movie nights and organizing the borrels together. Sasha, we had good times on BAM Friday! Mohammed, thanks for joining our NMR club and for bringing good luck. Deni, you can do it! Klaartje, thanks a lot for everything you did for the project; it made a big difference. Also I really appreciated your positive contribution to the group. Iva, the time that you were around was definitely the best. Your project turned into a large part of the work described in this thesis and I am very glad that we have kept in touch.

The rest of the NMR group: Mark, I would have been lost in the wet lab without your cloning skills. Gert, thank you too for wet lab advice. Johan, without your support we could forget about doing any NMR! Hans and Hugo, thanks for help with solution NMR. Klemen, thanks for (mental) support in the wet lab. Koen, thanks for the amazing JACS cover design! Maryam, it was always nice to chat with you, and thanks for the souvenir from Iran. My computational friends: Panos, with you around there was never a dull moment! João, thanks for help with docking and MD, for tons of chocolate, but mostly for being my friend in the lab. Anna, it was nice to hang out together in- and outside of the lab. Adrien, Mikaël and Gydo, and also Rama, Abhi and Sabine, playing badminton was a lot of fun. And finally Barbara, I could always rely on you for administrative affairs, especially towards the end.

Thanks to our collaborator Prof. Jan Tommassen for insights into microbiology, and to his group members Jan Grijpstra and Frank Beckers for technical assistance.

I was lucky to be part of the International Max Planck Research School in Chemical Biology in Dortmund. Many thanks to Prof. Martin Engelhard, Prof. Roger Goody and Christa Hornemann for giving me the chance to join the meetings and for funding.

Thanks to the Bijvoet PhD students for scientific and social exchange during our AIO evenings, and especially to the other representatives Sara, Deniz, Jonas, Nick and Nelleke; organizing the puzzle treasure hunt was so much fun! Also I discovered that crystallographers are not that bad: thanks to Remco for our (yet to become) successful collaboration, Camilla for Italian lessons, Revina for being my favorite conference room mate and Louris for a great concert.

The "NMR helpdesk", Ole, Koen, Manvendra, Bas, Merijn, Nan, Yuliya and Brijith, thanks for keeping in touch and supporting each other to survive a PhD (or postdoc) in NMR!

My "master friends": Steven, Caspar, Timo, João, Marcelo, Marti and Sacha, thanks for our dinners and drinks, and our attempts not to talk about work too much…!

Eva en Nadine, mijn paranimfen, hoe kan het ook anders?! Het is alweer een tijd geleden dat we op dezelfde plek studeerden, maar ik heb nog steeds het idee dat we het samen doen. Bedankt voor jullie steun en gezelligheid.

Last but not least, mijn familie, bij wie ik altijd terecht kan. Tomas, bedankt voor je technische adviezen en ook voor je zakelijke kijk op de wetenschap; ik heb er heus wat van opgestoken! Mam, zonder jouw chemische tips was mijn eiwit blijven precipiteren en had dit proefschrift er misschien wel heel anders uitgezien, maar het zijn boven alles het doorzettingsvermogen en zelfvertrouwen dat je me hebt meegegeven en je onvoorwaardelijke steun, waarmee ik deze promotie heb kunnen volbrengen.

LIST OF PUBLICATIONS

T. Sinnige, K. Houben, I. Pritisanac, M. Weingarth, M. Renault, M Daniëls, R. Boelens, A.M.J.J. Bonvin, J. Tommassen, M. Baldus (2014) Conformational plasticity of the POTRA5 domain in the outer membrane protein assembly factor BamA. *Submitted*

T. Sinnige, M. Baldus (2014) Solid-state NMR resonance assignments of membrane-embedded BamA: Implications for dynamics within the transmembrane domain. *Submitted*

T. Sinnige, M. Weingarth, M. Renault, L. Baker, J. Tommassen, M. Baldus (2014) Solid-state NMR studies of full-length BamA in lipid bilayers suggest limited overall POTRA mobility. *J. Mol. Biol.* **426**(9), 2009–2021.

T. Sinnige, M. Daniëls, M. Baldus, M. Weingarth (2014) Proton-clouds to measure long-range contacts between nonexchangeable side chain protons in solid-state NMR. *J. Am. Chem. Soc.* **136** (12), 4452–4455.

G.E. Karagöz, T. Sinnige, O. Hshie, S.G.D. Rüdiger (2011) Expressed protein ligation for a large dimeric protein. *Protein Eng. Des. Sel.* **24**(6), 495-501.

G.E. Karagöz, A.M.S. Duarte, H. Ippel, C. Uetrecht, T. Sinnige, M. van Rosmalen, J. Hausmann, A.J.R. Heck, R. Boelens and S.G.D. Rüdiger (2010) The N-terminal domain of human Hsp90 triggers binding to the co-chaperone p23. *Proc. Nat. Acad. Sci. USA* **108**(2), 580-585.

T. Sinnige, G.E. Karagöz and S.G.D. Rüdiger (June 2010) Protein folding and chaperones. Encyclopedia of Life Sciences 2010, John Wiley & Sons, Ltd: Chichester, http://www.els.net/

CURRICULUM VITAE

Tessa Sinnige was born on 8 August 1987 in Leiden, the Netherlands. She followed a bachelor in Life Science & Technology at Leiden University and the Technical University of Delft, with a final research project in the group of Dr. Gregg Siegal at Leiden University on the interaction between RFC and DNA ligase I, which she studied using NMR. She continued with a master in Biomolecular Sciences at Utrecht University, where she studied conformational changes of the chaperone Hsp90 in the group of Dr. Stefan Rüdiger, followed by a single molecule study on knotted proteins at the TU Munich under supervision of Prof. Dr. Matthias Rief. In 2010, Tessa started her PhD studies in the NMR department at Utrecht University in the group of Prof. Dr. Marc Baldus, where the work described in this thesis took place.