SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
BIOPHYSICS
STRUCTURAL BIOLOGY
MICROBIOLOGY
BACTERIA

Received
30 July 2012

Accepted
15 October 2012

Published
12 November 2012

Correspondence and requests for materials should be addressed to B.-J.v.R. (brossum@fmp-berlin.de)

Assignment and secondary structure of the YadA membrane protein by solid-state MAS NMR

Shakeel A. Shahid[1,2], Stefan Markovic[1], Dirk Linke[2] & Barth-Jan van Rossum[1]

[1]Leibniz-Institut für Molekulare Pharmakologie (FMP); Robert-Rössle-Straße 10, 13125 Berlin, Germany, [2]Max Planck Institute for Developmental Biology; Spemannstraße 35, 72076 Tübingen, Germany.

We report the complete solid-state MAS NMR resonance assignment of a medium-sized, trimeric membrane protein, YadA-M. The protein YadA (*Yersinia* adhesin A) is an important virulence factor of enteropathogenic *Yersinia* species (such as *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*). YadA is localized on the bacterial cell surface and is involved in adhesion to host cells and tissues. It is anchored in the outer membrane by a transmembrane anchor domain (YadA-M). This domain hosts the so-called autotransporter function of YadA: it transports its own N-terminal domain through the outer membrane. The assignment is based on a dataset that consisted of several MAS NMR correlation spectra, recorded on a single, uniformly ^{13}C, ^{15}N- labelled microcrystalline preparation. Except for the single C-terminal residue and the mobile strep tag, we were able to completely assign YadA-M. From this, secondary structure elements were predicted, which, combined with several long-range interstrand restraints, yielded the architecture of the β-sheet.

YadA is a virulence factor found in the outer membrane of pathogenic *Yersinia* species, such as *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*[1]. This membrane protein is a prototypical member of a group of non-fimbrial, non-pilus adhesins, named trimeric autotransporter adhesins (TAAs)[2]. Its transmembrane anchor domain (hereafter called YadA-M) forms a highly stable trimeric β-barrel[3]. Upon membrane insertion which is dependent on the Bam machinery[4], the YadA anchor domain functions as an autotransporter: it transports its own N-terminal domain through the membrane, a process that can be slowed down and even inhibited by single point mutations in the transport pore[5]. After completion of this transport, the three N-termini build a stalk with a sticky head group that can adhere to extracellular matrix components such as collagen on the host cell surface[6]. This adherence is one of the first steps in the infection pathway, resulting in several enteric food-borne diseases in the case of *Yersinia*, ranging from enterocolitis, acute enteritis, diarrhoea, mesenteric lymphadenitis, septicaemia, to autoimmune disorders such as reactive arthritis[1]. The autotransport mechanism is currently poorly understood; knowledge about the structure and dynamics of the domains involved in the autotransport will contribute to a better understanding of this process and will shed light on the pathogenic function of TAAs in *Yersiniae* as well as in other species.

Membrane proteins are generally difficult to purify and to crystallize, which renders them difficult to be studied with well-established techniques for structure determination, such as X-ray crystallography and solution NMR spectroscopy[7]. Indeed, attempts to prepare high-quality single crystals of YadA-M diffracting well enough for structure studies with X-ray crystallography were thus far unsuccessful[3]. However, during our efforts to optimize crystallization conditions, protocols were found to prepare microcrystalline material in large amounts. Such 'messy' crystals are highly suitable for structure studies with solid-state magic-angle spinning (MAS) NMR, since this technique neither relies on availability of high-quality crystals, nor on rapid molecular tumbling. In fact, the high degree of structural order of the bulk material within the microcrystals aids to reduce inhomogeneous broadening of the NMR lines[8]. In this work, we report on the ^{13}C and ^{15}N chemical shift assignment of the membrane anchor domain of YadA by solid-state MAS NMR. From the chemical shifts, secondary structure elements were predicted, which, in combination with long range magnetization transfers, yielded the topology of the beta-sheet. A full structure calculation is presented in a separate paper[9]. The YadA-M construct used in these studies consists of the trimeric membrane anchor domain and the first part of the stalk, and contains 105 amino-acid residues per monomer. We were able to sequentially assign YadA-M almost completely, using a single, uniformly ^{13}C, ^{15}N labelled micro-crystallized sample made from detergent-extracted protein as described in[3].

The chemical-shift assignment was an essential first step towards the structure calculation of YadA-M[9] (PDB entry 2LME). This study briefly encompasses the strategies and experiments used to achieve the complete chemical shift assignment of a membrane protein using uniformly ^{13}C, ^{15}N labelled microcrystalline material. We believe that the scenario in which high-quality crystals are not available and protocols for the preparation of micro-crystals do exist, is not unique for our YadA-M system; rather, it is a situation that is commonly encountered in structural biology and puts the strategy described in this paper in a broader perspective.

Results

Spectral quality. The homotrimer YadA-M consist of three protomers of 105 residues each, of which the first fifteen residues are part of a highly flexible strep-tag introduced to aid in protein purification[3]. Figure 1a shows a one-dimensional (1D) ^{1}H-^{13}C cross-polarization (CP) MAS NMR spectrum recorded from uniformly ^{13}C, ^{15}N-labelled YadA-M. The spectrum is well-dispersed, which is illustrated by the many fine details in the generally poorly resolved carbonyl region around 172 ppm (shown as inset in Fig. 1a). The overall high dispersion is partly the result from the secondary structure of TAA membrane anchors, which for YadA-M is about half α-helix / half β-sheet[3]. In addition, Fourier-transform infrared studies on YadA-M have shown that the main part of the protein is extremely rigid, leading to a very high spectral resolution also in FTIR[3]. The (micro-) crystalline environment and the overall rigidity of the protein will contribute to the structural homogeneity and hence, to a narrow line width. In contrast, the strep-tag is known to be highly mobile (colour-coded in grey on the primary sequence in Fig. 1b)[3]. As a result, it can be expected that the MAS NMR signals of residues in the strep-tag will be very weak or not observed at all. The presumption that all strong signals arise from residues in the rigid part of the protein was proven during the course of the assignment procedure by the self-consistency of the assignment.

As a quality check and to monitor the integrity of the sample, 1D spectra were recorded directly before and after any multidimensional experiment. During the course of the work, no indication for change or degradation of the sample was observed. Sample integrity was also monitored by following the Cε-Cγ correlation of the (single) methionine in the sequence (M96), which did not show any sign of oxidation (See Supplementary Fig. S1 online). Hence, these findings imply that the YadA-M trimer is remarkably stable over a longer period of time.

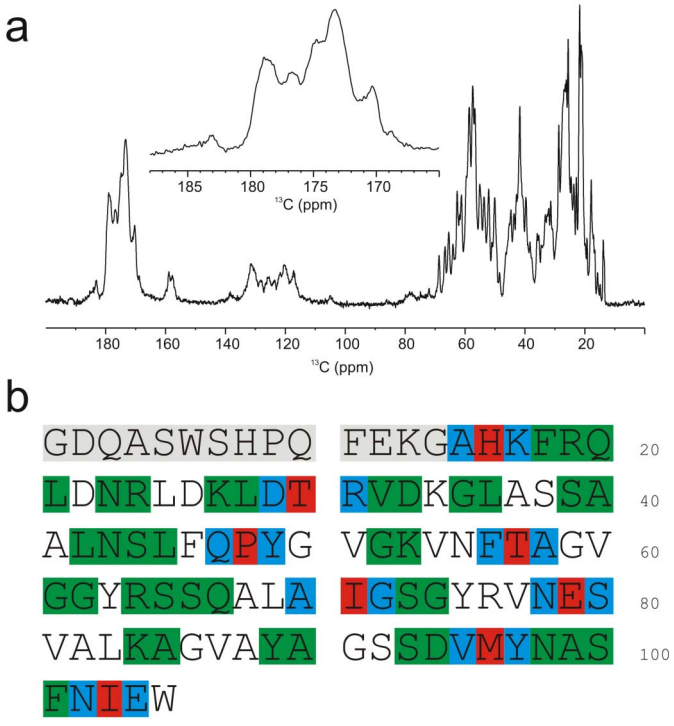

Figure 1 | 1D ^{13}C spectrum and YadA-M sequence. The spectrum (a) is recorded at a field of 900 MHz, a spinning frequency of 15 kHz and at a temperature of 275 K, using a standard ^{1}H-^{13}C cross polarization (75–100% ramp on ^{13}C)[27]. The inset shows the carbonyl region. The starting points of the sequential assignment are colour coded on the primary sequence of YadA-M; residues that appear only once (H16, P48, M96) or twice (T30 and T57; I71 and I103) are coloured red, their sequential neighbours blue, and unique sequential pairs green.

Assignment strategies - Identification of spin systems. The sequence-specific assignment was achieved using a dataset consisting of several two- and three-dimensional (2D and 3D, respectively) homo- and heteronuclear correlation experiments (for a detailed description of the experiments, see the supplementary material – Supplement Table 1). In a first step, 2D carbon-carbon correlation experiments were analysed to search for 'fingerprint patterns' of the various types of amino acids. In a second step, sequential relationships are established by linking ^{13}C resonances to the backbone-amide ^{15}N using heteronuclear, multidimensional experiments.

Figure 2 shows the aliphatic and carbonyl regions of a 2D ^{13}C-^{13}C homonuclear correlation spectrum of YadA-M. The correlations were obtained by using a DARR/RAD scheme with 25 ms mixing[10,11]. The spectrum is fairly well resolved and several amino-acid types can be directly identified upon their characteristic correlation pattern and unique chemical shifts, such as threonines, serines and isoleucines; other residues give fingerprint patterns in the more congested regions of the spectrum and are not readily identified due to ambiguity; still, in many cases, side-chain correlation patterns can be traced putatively and their assignment is achieved by checking for self-consistency in the different spectra of the dataset. Glycines do not have side-chains, hence, they do not provide cross peaks in the aliphatic region. However, they are readily identified upon their unique upfield-shifted Cα-C` correlations (cf. Fig. 2b).

Using unique doublets and triplets for assignment. The identification of fingerprint patterns yields residue-specific, but mostly not sequence-specific assignments. However, when these patterns are distinctive either on basis of well-dispersed chemical shifts or arise from residues that are unique or appear in a very little number, they provide good starting points for the sequential assignment. YadA-M has a highly repetitive primary sequence of mainly alanines (13), glycines (12), serines (12), leucines (8), valines (8), asparagines (6), aspartates (6), lysines (6), arginines (5), glutamines (5), tyrosines (5) and phenylalanines (5); together these residues constitute 91 out of 105 in total for YadA-M. We found eight unique spin systems and used them as starting points for sequential assignment. This is illustrated in Figure 1b, which shows the primary sequence of YadA-M; colour-coded in red are observable amino acids that appear either only once (H16, P48, M96) or twice in the sequence (T30 and T57; I71 and I103). From these starting points, the neighbouring residues can be assigned (colour-coded blue in Fig. 1b). There are three glutamic acids in YadA-M (E12, E79 and E104). E12 resides in the strep-tag and is not observed. Since E104 is identified by its neighbouring I103, the remaining glutamic acid (E79) is considered unique and colour-coded red as well. In addition, pair-wise residues that form unique doublets with at least one of the members easily identifiable are coloured green. We found 19 of such pairs, which additionally provided unambiguous starting points for the sequential assignment.

Sequence specific assignment. For the sequence-specific assignment of the residues, it is mandatory to know the backbone ^{15}N chemical shift for each residue. For this purpose, band-selective heteronuclear ^{13}C-^{15}N correlation spectra are recorded. Here, the

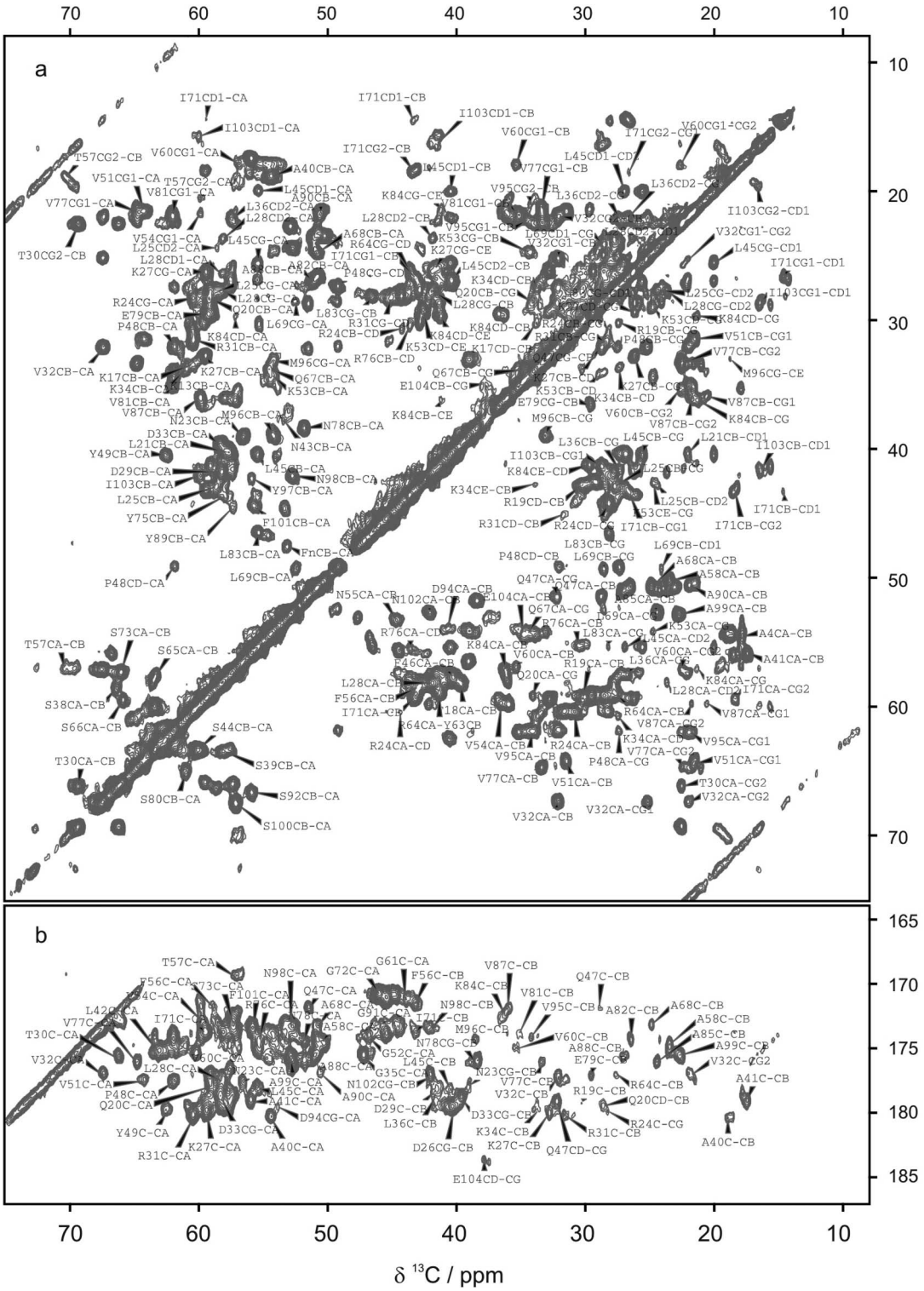

Figure 2 | 2D ^{13}C-^{13}C DARR spectrum. Contour plot of a 2D DARR experiment[10,11], recorded at a field of 900 MHz, a spinning frequency of 12 kHz and at temperature of 275 K. Magnetization between ^{13}C spins was exchanged by use of a 25 ms DARR mixing period. The aliphatic region (a) and carbonyl region (b) are shown together with chemical shift assignments of YadA-M.

essential building block is a selective polarization transfer step between the amide ^{15}N and the ^{13}Cα (which is called NCA transfer) or the C' (NCO transfer)[12], achieved with adiabatic cross polarization optimized for either the NCA condition or the NCO condition[13]. The NCA transfer is int**ra**residual, i.e., it connects between $^{15}N_i$ and $^{13}C\alpha_i$ of the same residue; the NCO transfer is int**er**residual and involves a crossing of the peptide bond to connect $^{15}N_i$ with $^{13}C'_{i-1}$. Figure 3 shows a 2D NCA (right) and NCO spectrum (left). In the NCA spectrum, certain residues give rise to fingerprint signals. For instance, glycines and alanines provide easily identifiable correlations due to their upfield shifted ^{13}Cα signals, and upfield (glycine) or downfield (alanine) shifted ^{15}N signal. Valines are characterized by a downfield shifted ^{13}Cα and ^{15}N resonance. Serines have an upfield ^{15}N signal and an intermediate ^{13}Cα signal (between 55–60 ppm). These identification 'rules' can only be applied in one direction: whereas it is true that in the typical glycine, alanine or valine NCA

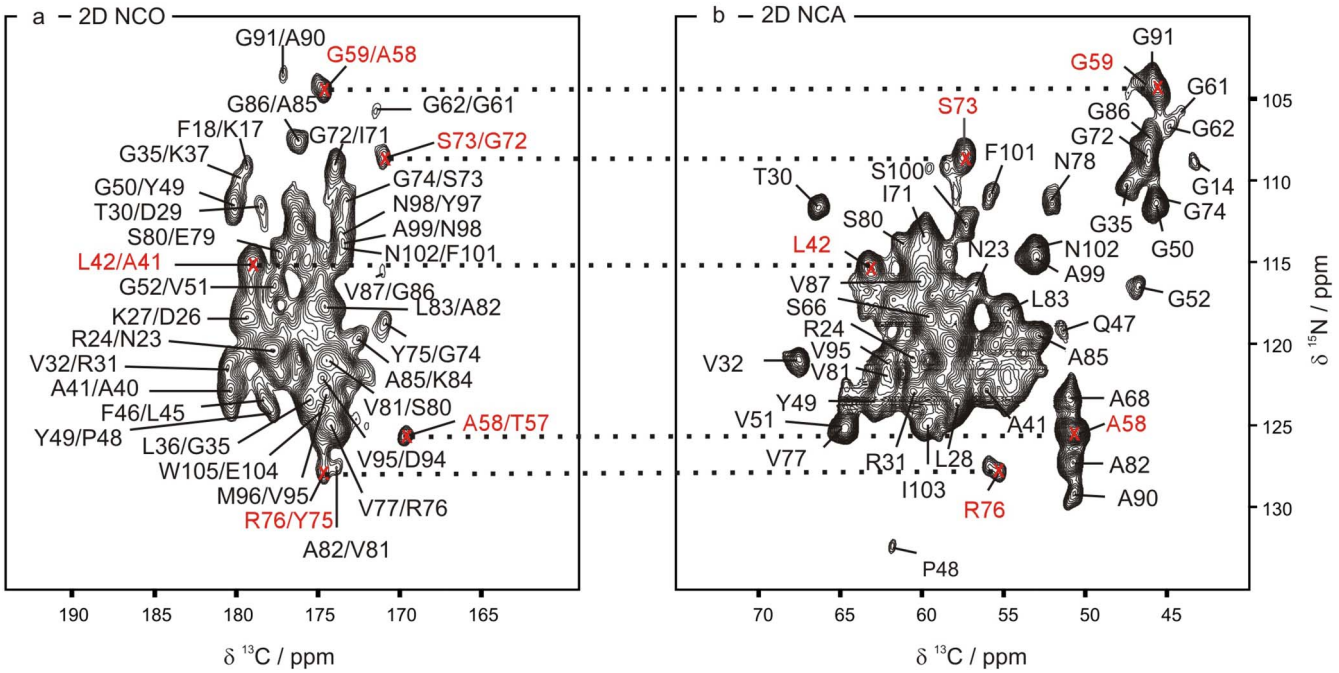

Figure 3 | 2D ^{15}N-^{13}C NCA and 2D NCO spectra. Contour plots of a 2D NCO (a) and NCA spectrum (b), recorded at a field of 400 MHz, a spinning frequency of 8 kHz and at a temperature of 275 K. Polarization exchange between ^{1}H and ^{15}N was obtained by use of a 1 ms standard cross-polarization (CP) period; selective transfer from ^{15}N to ^{13}C' (a) or ^{13}Cα (b) was achieved by a 4 ms adiabatic CP[13]. Chemical shift assignments for YadA-M are indicated in the figure.

region no other residues are found, the other way round it is not true, as demonstrated by the shifts of e.g. G52, A41, A99 and V87. In the NCO spectrum (Fig. 3, left), carbonyl spins are correlated with backbone nitrogen spins. Note that in the NCO experiment no fingerprint patterns are found, since the correlations connect between spins of different amino acids. The two spectra are linked by the ^{15}N frequency of the backbone amide, providing sequential connectivities. Several examples to illustrate this pair-wise linking are highlighted in red in Figure 3.

In general, assignment walks can only be established if the NCA and NCO correlations are 'dressed' with the shifts of the side-chain spins, which is usually done by extending the NCA or NCO experiments with a non-specific ^{13}C-homonuclear transfer step (e.g., by employing a proton-driven spin diffusion (PDSD) transfer step[14], or DAR/RADD mixing). These experiments are referred to as either NCACX or NCOCX (where CX denotes 'any' carbon)[15]. The assignment procedure is then to identify the same side chain resonance pattern correlated to its own backbone ^{15}N (NCACX) and the backbone ^{15}N of the next residue in the sequence (NCOCX). These experiments are best recorded as 3D data to lift overlap between side-chain patterns by the additional ^{15}N frequency. In this respect, the NCOCX correlation is more dispersed than the NCACX correlation, since the amide ^{15}N and the ^{13}C' signals do not cluster in fingerprint regions as is the case for the NCA correlations. In Figure 4, several 2D strips extracted from a 3D NCACX (red contours) and a 3D NCOCX (cyan contours) are shown. Both 3D spectra were recorded with 35 ms DARR mixing following the adiabatic ^{15}N-^{13}C transfer to exchange polarization between backbone and side chains. The strips illustrate the assignment procedure. In strip **A** from the NCACX experiment, a threonine spin system is identified based on its fingerprint pattern; the threonine can either be T30 or T57, as we have only two threonines in the sequence. The 3D NCOCX spectrum is scanned for the same pattern of side chain signals and at ^{15}N=122.7 ppm, a similar cross peak pattern is found (shown in strip **B**). In the next step, a strip is extracted at ^{15}N = 122.7 ppm from the 3D NCACX experiment (strip **C**). In this strip, a cross peak pattern is found that is attributed to an arginine rather than to an alanine, from which the sequential connection can be established between T30 and R31 rather than between T57 and A58. To confirm this assignment, the next residue should be a valine (V32), of which the nitrogen frequency can be found in the NCOCX experiment at 121.0 ppm (strip **D**). Indeed, the cross peak pattern observed at a ^{15}N frequency of 121.0 ppm in the NCACX spectrum is highly characteristic for a valine (cf. strip **E**) and thus confirms the sequential assignments made. By searching for the V32 side chain signals in the NCOCX, the ^{15}N shift of D33 is found at 121.6 ppm (cf. strip **F**). This procedure is continued and the stretch of amino acids from T30 to K34 is assigned as shown in strips **A–J** in Figure 4.

Assignment using selective transfers. In cases of repetitive spin pairs or when cross peaks appear in the more congested regions of the spectra, multiple assignment possibilities may exist. In order to be able to assign residues unambiguously, additional spectra were recorded. For example, a helpful addition to the dataset was a 3D NCACB spectrum, where selective exchange between ^{13}Cα and ^{13}Cβ was achieved with a DREAM transfer scheme[15,16]. In this spectrum, Cα-Cβ correlations of Ala, Val, Lys, Leu and Asp could be unambiguously resolved from overlap with correlations from other residues. Along the same lines, two 2D ^{13}C correlation spectra were recorded, employing a DREAM transfer scheme optimized for the Cα-Cβ and Cβ-Cγ spectral regions (Supplementary Fig. S2 online). The advantage of the DREAM experiment is that it can be set up so that it predominantly provides transfers between spins that are directly connected; in our experience, transfer events over two bonds occur less frequently and are mainly observed for spin pairs where at least one of the two carbon atoms is only weakly coupled to ^{1}H. Examples are two-bond couplings involving carbons in methyl groups (like leucine $C\delta_1$-$C\delta_2$, leucine Cβ-$C\delta_{1/2}$, valine $C\gamma_1$-$C\gamma_2$ and isoleucine Cβ-Cδ) or carbons in flexible side chains (like lysine Cε-Cγ and arginine Cδ-Cβ). Such occasional two-bond transfers can be easily distinguished from one-bond transfers since both have opposite sign in cross-peak intensity. Note that in highly congested

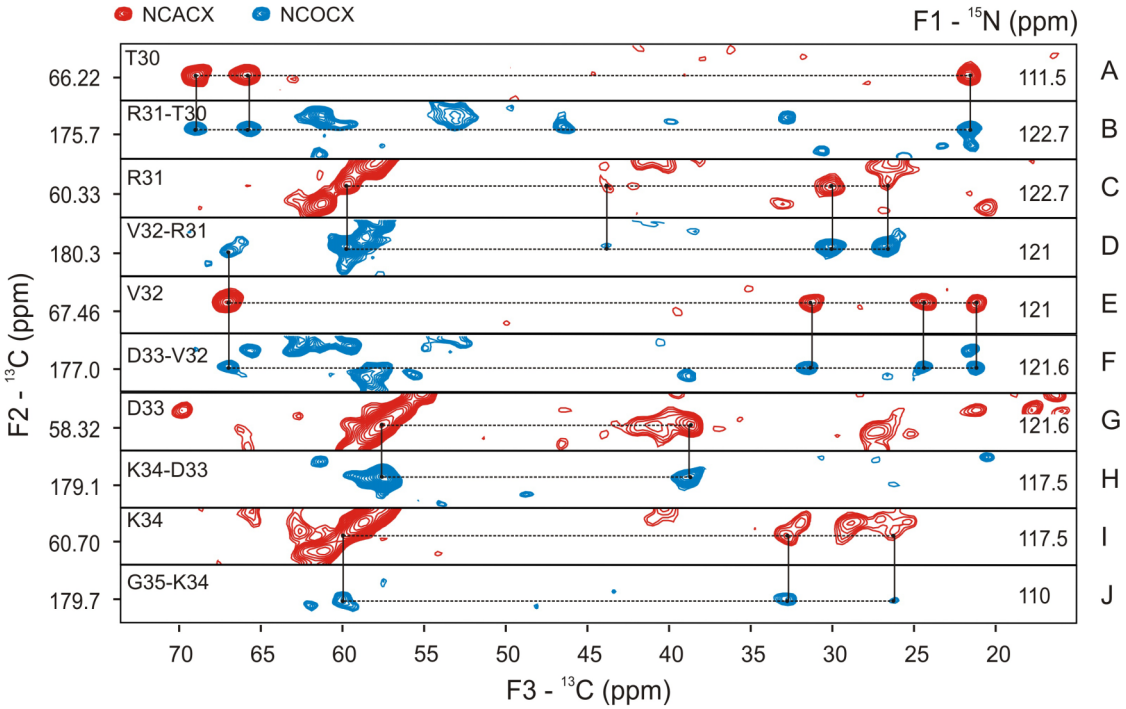

Figure 4 | Strips extracted from 3D NCACX and 3D NCOCX experiments, illustrating the sequential correlation strategy. The 3D experiments were recorded at a field of 600 MHz, a MAS frequency of 10 kHz and at a temperature of 275 K. Initial ^{15}N polarization was created with a standard ^{1}H-^{15}N CP (2 ms, 75–100% ramp on ^{1}H), followed by selective transfer from ^{15}N to ^{13}C' or ^{13}Cα using an adiabatic CP. A PDSD (proton-driven spin diffusion)[14,30] sequence of 35 ms was applied to exchange magnetization between the ^{13}C spins.

regions of the spectrum a partial cancellation of cross peaks may occur. In the 2D Cα-Cβ DREAM spectrum, correlations between spins that are more than one bond away are virtually suppressed (Supplementary Fig. S2 online). This is instrumental to distinguish Cα-Cβ cross peaks of valines, lysines, arginines and glutamic acids from two-bond Cα-Cγ correlations that appear in the same spectral region (e.g. from leucines, lysines and arginines). By optimizing the DREAM transfer around a more upfield frequency, one-bond transfers, mainly between ^{13}Cβ and ^{13}Cγ carbons, can be distinguished from two-bond transfers involving methyl groups (Supplementary Fig. S3 online). This is particularly useful for distinguishing Cγ-C$\delta_{1/2}$ and Cδ_1-Cδ_2 resonances of leucines.

Consistency check of the sequence specific assignment. In addition, we recorded 3D NCACX and 3D NCOCX with longer mixing times (100 and 200 ms PDSD transfer). Controversially, we found that such spectra helped to remove ambiguity, even though the amount of cross peaks is higher. The explanation for this is, whereas in NCACX with short CC mixing times magnetization is only exchanged intra-residually ('forward transfer'), with longer mixing times the ^{13}Cα also exchanges with the backbone and side-chain spins of the previous residue in the sequence ('backward transfer'); this provides a great means for checking the self-consistency of the sequential assignment without need to revert to the intermediate $^{15}N_i$ and $^{13}C'_{i-1}$ shifts. In complete analogy, NCOCX spectra with longer CC mixing times do not only provide backward transfer between ^{13}C' and the side chain of the same residue, but also forward transfer, directly linking $^{13}C'_{i-1}$ with the side chain of residue *i*. An example of multiple transfer pathways mutually linking the pair S44-L45 is shown in Figure 5. In the top part of the figure, magnetization transfer events are shown schematically. Red arrows represent the transfer of ^{13}C polarization within a spin system, blue arrows the exchange between residues, using the ^{13}Cα (left) or ^{13}C' (right) as starting point. The sequential assignment of the residue pair S44 and L45 is shown, aided by inter-residue correlations. In the NCACX experiment recorded with long PDSD mixing time of 200 ms, polarization can be transferred from ^{13}Cα of residue *i* to residue *i*-1. This is shown schematically in the top-left part of Figure 5 where blue arrows indicate the backward transfer from Cα of L45 to C′, Cα and Cβ of S44 (labelled with 3 and 4). The corresponding cross peaks are labelled accordingly in strip **B**. In contrast, in the NCACX experiment recorded with a short mixing time of 35 ms, only intra-residue cross-peaks for L45 are observed (strip **A**). Likewise, in the NCOCX experiment with long PDSD mixing, polarization can be exchanged between $^{13}C'_{i-1}$ and residue *i* (forward transfer, Fig. 5 top right, blue arrows). This is demonstrated in strip **D** which shows both intra-residue (labelled with 5) and inter-residue correlations (labelled 6 and 7). Interestingly, in the NCOCX obtained with a short mixing of 35 ms, apart from strong intra-residue cross peaks, weak transfer from S44 C' to L45 Cα can be observed (cf. strip **C**).

Finally and along the same lines, a further powerful check for the consistency of the sequence-specific assignment is achieved by analysis of direct correlations between the ^{13}Cα spins. Also here, this transfer does not involve the intermediate backbone amide and carbonyl ^{13}C' spins; hence, any potential mistake made during the alignment procedure of strips as described above would lead to inconsistent results. Sequential Cα-Cα correlations are readily observed in ^{13}C-^{13}C spin diffusion spectra with extended DARR or PDSD mixing times (typically 100–200 ms). As an example, the sequential walk connecting the A82-A90 subsequence via the Cα's is shown in Supplementary Figure S3 online. Since the relevant information is obtained from a relatively narrow spectral region, these experiments perfectly lend themselves for *J*-decoupling in the indirect dimension[17]. Plots of the Cα-Cα region in a standard and *J*-decoupled DARR spectrum are included in Supplementary Figure S3 online.

REDOR and methyl filtered schemes. Prolines are difficult to assign on the basis of 2D or 3D NCACX experiments because the lacking amide proton renders the first ^{1}H- ^{15}N CP very inefficient. In the 2D

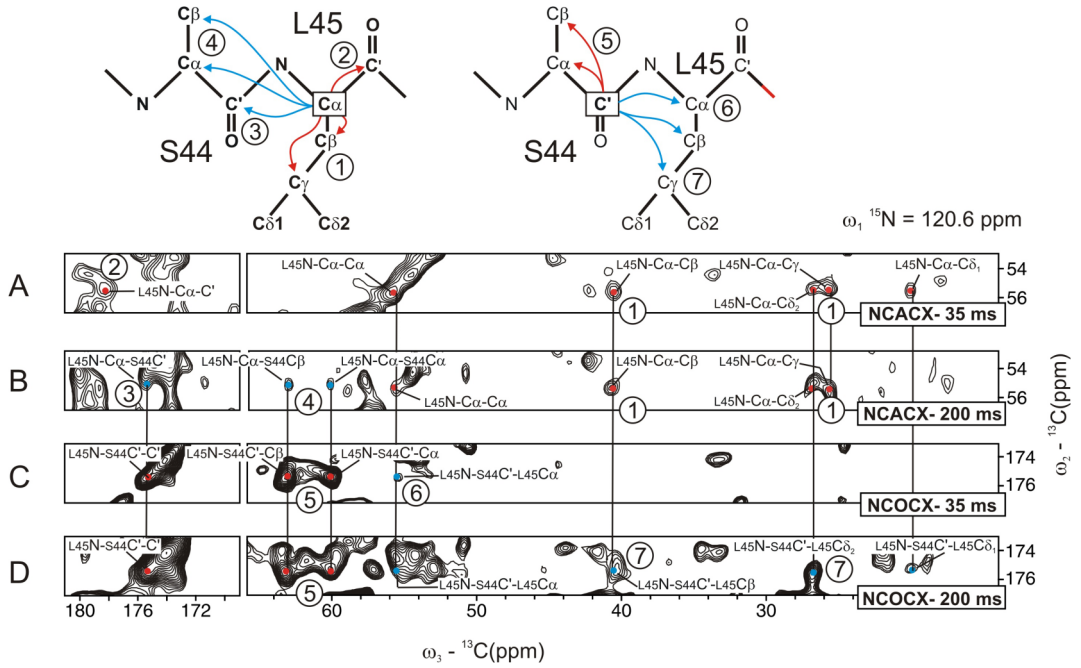

Figure 5 | Schematic representation of magnetization transfer (top); red arrows represent intra-residue transfer of ^{13}C polarization, blue arrows inter-residue ^{13}C exchange, starting from ^{13}Cα (left) or ^{13}C' (right). 2D strips extracted from 3D NCACX (strips **A** and **B**) and 3D NCOCX spectra (strips **C** and **D**, bottom); the 3D spectra were recorded with proton-driven spin diffusion mixing of 35 ms (strips **A** and **C**) and 200 ms (strips **B** and **D**).

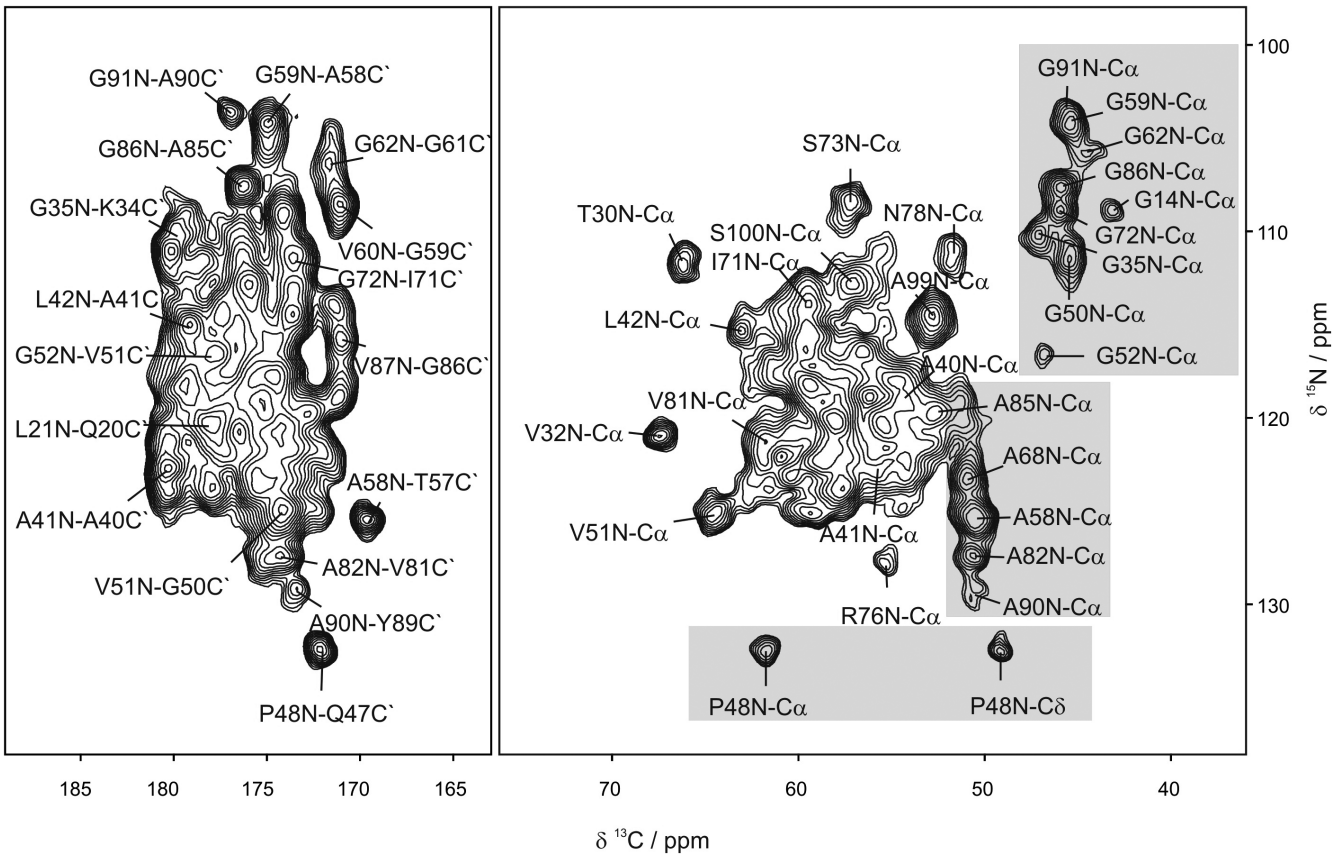

Figure 6 | 2D ^{15}N-^{13}C TEDOR experiment. Contour plot of the NCO-region (left panel) and NCA region (right panel) of a TEDOR (transferred echo double resonance)[18,19] experiment, recorded at a field of 9.4 T, 8 kHz MAS and at a temperature of 275 K. The experiment contains two REDOR (rotational echo double resonance)[36] mixing periods of 1 ms each, to transfer magnetization back-and-forth between carbon and nitrogen. The gray boxes highlight the fingerprint regions for ^{15}N-^{13}C correlations involving glycines, alanines and prolines.

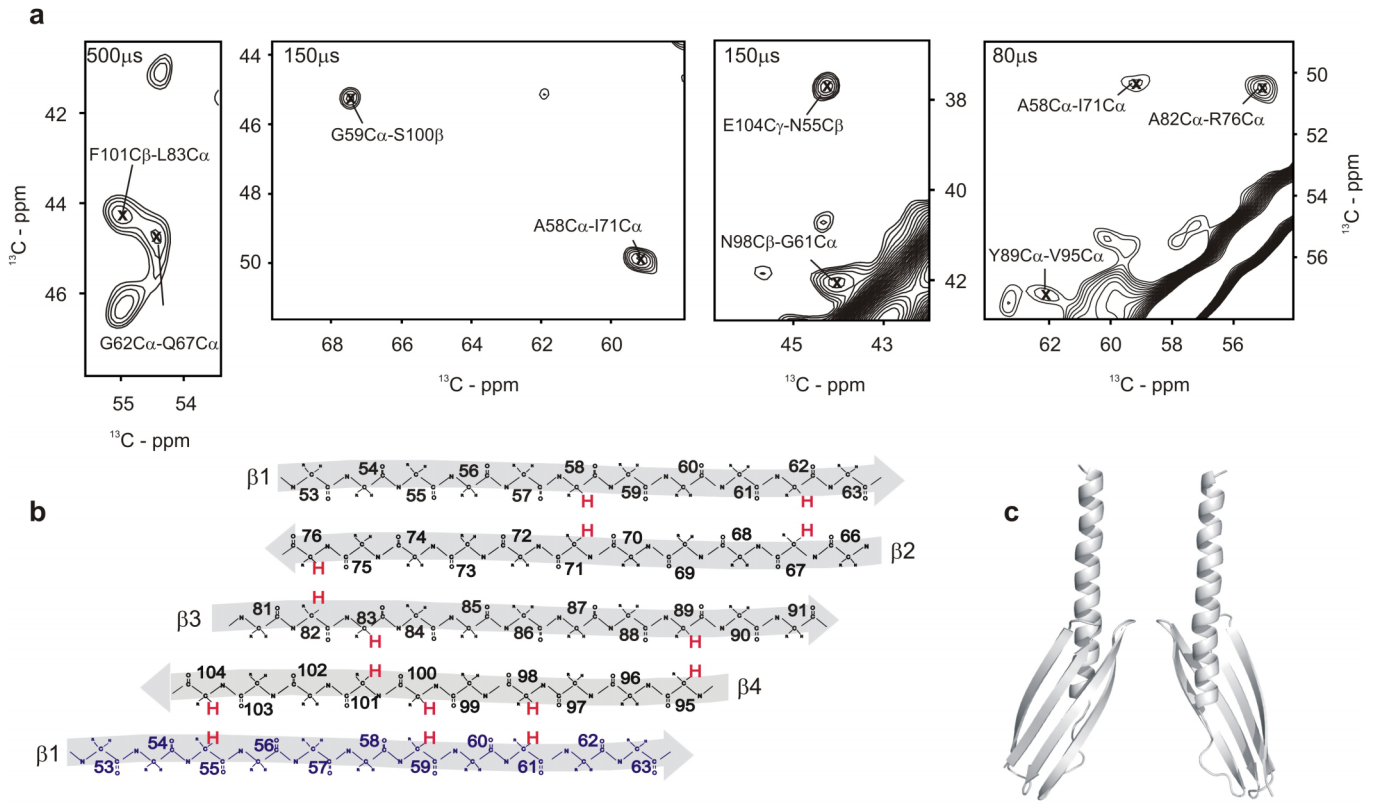

Figure 7 | Interstrand correlations and β-sheet topology of YadA-M. (a) 2D contour plots from 2D CHHC spectra recorded on YadA-M at different mixing times. Cross peaks that were unambiguously attributed to interstrand correlations are indicated. (b) Anti-parallel β-strands in YadA-M. Cross peaks indicated in (a) that define the topology of the β-strands are highlighted in red ("**H**"). The β1 strand from the adjacent monomer is shown (in blue) to be able to indicate correlations between strands β1 and β4. (c) Ribbon model of the YadA-M monomer[9].

NCA experiment shown above, only a relatively weak cross peak with $^{13}C\alpha$ is observable for P48 (cf. Fig. 3b); however, in the 2D NCO experiment (Fig. 3a), and in the 2D or 3D NCACX and NCOCX experiments (data not shown), no side-chain correlations involving P48 were observed. Prolines provide, however, a highly characteristic correlation pattern in 2D ^{15}N-^{13}C TEDOR experiments[18,19], consisting of pair-wise cross peaks involving the Cα and Cδ in the α-region with the downfield-shifted backbone amide. As shown in Figure 6, a proline fingerprint pattern appears at a ^{15}N shift of 132.2 ppm, which is assigned to P48. Note that the other proline in the sequence (P9) resides in the mobile strep-tag and is indeed not observed in Figure 6.

The last experiment in the dataset was a methyl-filtered post-C7 experiment. Methyl filtering can be achieved spectroscopically by a straightforward combination of cross-polarization with phase inversion (CPPI) and a dipolar dephasing step[20]. The filter selects carbon atoms that are relatively weakly coupled to ^{1}H, primarily targeting at the methyl-carbons, but also non-protonated carbonyls and carbons in flexible side chains. In combination with the Post-C7[21], the spectrum is highly instrumental for the assignment and/or corroboration of the assignment of methyl-containing amino acids (Ala, Ile, Leu, Val, and Thr). More specifically, it helped to unambiguously identify T57, which has an unusual upfield-shifted $^{13}C\alpha$; as a result, the Cα-Cβ cross peak of T57 is found in the serine Cα-Cβ fingerprint region and as such was not identified straightforwardly (cf. Fig. 2 and Supplementary Fig. S4 online). Additionally, in the carbonyl part of Post-C7, Cβ-Cγ correlations for Asn, Asp and Cγ-Cδ correlations for Gln, Glu can be readily identified.

Topology of beta sheet. The backbone ^{13}C and ^{15}N chemical shifts obtained from the sequential assignment can be assessed to predict secondary structure elements. For this, we used the program TALOS+[22]. According to the TALOS+ prediction, YadA-M consists of an N-terminal α-helix (residues H16-L45) and four β-strands (β1: K53-Y63; β2: S66-R76; β3: V81-G91; β4: V95-E104) connected by relatively tight beta-turns. The residues F46-G52 that connect the helix with β1 are predicted as an unstructured loop region. In addition to the TALOS+ predictions, we recorded CHHC experiments to obtain interstrand correlations. The interstrand distance between protons on adjacent alpha carbons in an anti-parallel β-sheet is approximately 2.3Å. We recorded 2D CHHC spectra using various mixing times and found several correlations that could be unambiguously assigned as interstrand cross peaks (Fig. 7a). Taking into account that the loops connecting the β-strands are fairly short, correlations between β1-β2, β2-β3 and β3-β4 are considered intramolecular. The β-sheet topology consisting of a 4-stranded, antiparallel β-sheet which is in agreement with TALOS+ predictions and the observed inter-strand cross peaks is shown schematically in Figure 7b and as ribbon model in Figure 7c. We also observed cross peaks between β1 and β4. In principle these correlations could be intramolecular, but this would result in a highly unlikely narrow beta-barrel with perpendicular β-strands. Rather, we attribute correlations between strands 1 and 4 to intermolecular interactions between the protomers forming the trimer. The relative position of strand 4 of one monomer with respect to strand 1 of a second monomer defines the shear number of the closed β-barrel. For more details and the full structure calculations, see reference[9].

Discussion

We report the solid-state MAS NMR resonance assignment of a medium-sized trimeric membrane protein, YadA-M. From evaluation of the dataset as described above, we could unambiguously

assign all residues in the subsequence 14–104 of YadA-M (BMRB entry 18108). The residues that we were unable to assign are those in the strep-tag. It has been shown by subtilisin treatment that the strep-tag is extremely mobile and renders fair degree of flexibility to the N-terminal residues[3]. As a consequence, most of the residues in the strep-tag were not observed in any of the spectra of our dataset, while some others gave very weak and incomplete cross peak patterns. For instance, as noted above, only one strong proline signal set was detected, whereas the proline that resides in the strep-tag was not observed (cf. Fig. 6). In addition, in ^{1}H-^{13}C INEPT spectra, signals are observed that cannot be attributed to the detergent but can be tentatively assigned to the strep-tag residues (Supplementary Fig. S5 online). The assigned residues in the subsequence 14–104 can completely account for the cross peaks detected in the various spectra of the dataset; the sequential assignment is highly self-consistent (e.g., cf. Fig. 5) and no strong cross peaks were 'left over' as unassigned, which strongly supports our initial presumption that the observable signals predominantly arise from the rigid part of the protein. Also for the rigid part, we found that for several regions cross peaks appeared systematically weaker, suggesting a higher degree of (local) mobility. Among these are the subsequence K53-V54-N55, the region around A37 and the C-terminus (W105). Although a tentative assignment for W105 was possible, the sequential link was poor and we do not consider W105 as part of the assigned residues.

Despite the highly repetitive primary sequence of YadA-M, several unique spin systems and a fair number of pair-wise unique spin-systems were found that served as starting points of the assignment (cf. Fig. 1b). We found that glycines and alanines, even though they are relatively abundant in YadA-M, proved very useful for sequential assignment. Their backbone ^{13}Cα and ^{15}N are readily identifiable from 2D NCA type spectra (cf. Fig. 3) and can be used as sequential linkers in a parallel evaluation of 3D NCACX and NCOCX spectra. Serine residues, on the other hand, turned out to be less helpful; although serines are generally easily identifiable by virtue of their downfield Cβ chemical shift, six out of the 12 serines in YadA-M form SS pairs, which made it difficult to link them sequentially.

At all stages of the assignment procedure, the choice of mixing time played a crucial role for deriving the required information. In the initial steps of the chemical shift assignment, identification of spin systems required relatively short mixing times, typically 10-20 ms for PDSD, to ensure predominantly intra-residue spectral cross peaks. For establishing sequential links, a PDSD or DARR mixing scheme with 100 ms proved to be sufficient. Still longer mixing times (200 ms) were useful in a final stage of the assignment process to cross check the assignment for consistency. Selective transfers using DREAM mixing in 3D NCACB, 2D CACB and 2D CBCG helped identification of several spin systems by reducing the information content of spectra. Amino-acid residues with methyl groups i.e., Ala, Leu, Iso, Val and Thr were assigned/confirmed with the help of a double-quantum methyl filtered Post-C7 experiment. In the carbonyl part of the Post-C7 experiment, presence of Cβ-Cγ correlations for Asn, Asp and Cγ-Cδ correlations for Gln, Glu made this spectrum extremely useful. Aromatic amino-acid residues generally show relatively weak intra- and inter-residue cross-peak intensities owing to their aromatic rings which can act as magnetization sink[23]. However, they provide highly characteristic cross peaks which allow their unambiguous assignment. Phenylalanines were identified on the basis of their aromatic Cγ cross peaks around 140–142 ppm. Tyrosines and arginines show their Cζ cross peaks around 159 ppm. In turn, tyrosines can be distinguished from arginines by virtue of their aromatic Cε1 and Cε2 cross peaks (116–117 ppm).

YadA is a trimeric autotransporter protein. Its transmembrane domain is homologous to other transmembrane β-barrel proteins from Gram-negative bacteria[24]. This large and diverse family of proteins has some conserved features also seen in YadA: it consists of an even number of amphipathic, antiparallel β-strands, and both ends of each strand are frequently made from aromatic residues that are known to interact with the water-lipid interface to anchor the proteins in the lipid membrane[25]. In contrast to most other β-barrel proteins where the barrel structure is produced from a single polypeptide chain, YadA is a homotrimer where all three subunits contribute four β-strands to the barrel.

In summary, we report the complete solid-state MAS NMR resonance assignment of a medium-sized trimeric membrane protein, YadA-M, based on a single, uniformly ^{13}C, ^{15}N-labelled, microcrystalline preparation. The assignment was achieved on basis of a dataset that consisted of several homo- and heteronuclear MAS NMR correlation spectra. We were able to unambiguously assign the residues 14-104 of YadA-M, which covers the entire membrane anchor domain. The residues that we were unable to assign are those in the strep-tag and the C-terminal residue. Since the strep-tag is not part of the native protein and has no physiological function, we can conclude that, except for W105, we were able to assign the entire physiologically relevant part of the anchor domain. The full chemical-shift assignment forms an important requirement and first step towards the structure calculation of YadA-M[9]. The micro-crystalline material used for our studies essentially resulted as "byproduct" from our (non-successful) crystallization efforts. This is a scenario that is not uncommon in structural biology, since membrane proteins are generally difficult to crystallize and it can well happen that, despite many efforts, at the end of the day at best poor-reflecting crystals or micro-crystals are obtained; such preparations are not suitable for X-ray diffraction studies, but still form perfect samples for solid-state NMR studies. Hence, failure of sample preparation for one technique should not be the end of the road but rather can provide invaluable samples for other techniques, and we believe that the current work could trigger many future solid-state MAS NMR investigations and offer new perspectives in structural biology, especially for membrane proteins where poor crystal quality is the rule rather than the exception.

Methods

Sample preparation. YadA-M was expressed in a full medium with ^{13}C/^{15}N (BioExpress, Cambridge Isotope Laboratories), using the expression, detergent extraction and purification procedure described in[3]. The large-scale protein crystallization was done by dialysis to remove excess detergent, as described in more detail in[9].

NMR spectroscopy. All solid-state MAS NMR experiments were performed at 275 K using samples from a single ^{13}C, ^{15}N uniformly labelled YadA-M batch preparation. Data were obtained on 400, 600 and 900 MHz AVANCE spectrometers (Bruker, Karlsruhe, Germany). Spectrometers were equipped with double (^{1}H/^{13}C) and triple-resonance (^{1}H/^{13}C/^{15}N) CP/MAS probes (Bruker). The temperature was set to 275 K. Except otherwise mentioned, magnetization transfer from ^{1}H to the ^{13}C or ^{15}N spins was achieved with ramped cross-polarization (CP)[26,27]; the ramp was kept relatively shallow (typically between 75 and 100% of the spin-lock field strength). Typical recycle delays between individual transients were kept between 2.7–3.0 s, to avoid sample heating. High-power proton decoupling with radio-frequency (rf) field strengths of 75-90 kHz using the TPPM[28] or SPINAL-64 scheme[29] was applied during evolution and detection periods. Typical ^{1}H and ^{13}C π/2 pulse lengths were 3.0 and 3–4 μs, respectively, typical ^{15}N π/2 pulses were longer, 6–7 μs.

Homonuclear 2D ^{13}C-^{13}C correlation spectra with dipolar-assisted rotational resonance (RAD/DARR)[10,11] proton-driven spin diffusion (PDSD)[14,30] mixing schemes and DREAM mixing[15,16] were recorded at 900 MHz. For the RAD experiments, the following conditions were used: experiments were obtained at 10 kHz MAS; initial ^{13}C magnetization was created using a 3.0 ms ramped CP (75–100%) with average rf-field strengths of 63 and 53 KHz, for ^{1}H and ^{13}C, respectively. Mixing times of 15, 25, 50 and 100 ms were used, in order to probe residue specific and sequential transfers in different distance ranges. A 3.0 s recycle delay was used. Acquisition times were typically 18.0 and 9.6 ms in t_2 and t_1, respectively. The PDSD experiments were recorded under similar conditions, but with a CP contact of 1.0 ms (75–100% ramp on carbons), and acquisition times in the direct and indirect dimensions of 15.0 and 4.0 ms, respectively. PDSD mixing periods ranged between 15 ms and 100 ms. The 2D ^{13}C-^{13}C DREAM spectra were obtained at 13 kHz MAS; a CP contact of 1 ms (75–100% ramp on ^{13}C) was used. During the DREAM mixing period of 1.5 ms, the carrier was placed either between the Cα and Cβ regions (at ~50 ppm, Cα-Cβ spectrum) or between the Cβ and Cγ regions (at ~30 ppm, Cβ-Cγ spectrum). Acquisition times in the direct and indirect dimensions were 15.0 and 8.2 ms. The experiments were

recorded with 1024 increments, 48 scans per increment and a recycle delay of 3.0 s, yielding a total experimental time of 41 hours per spectrum.

2D NCA and NCO experiments were recorded at 400 MHz, using a MAS frequency of 8 kHz. Initial magnetization was created using a 1 ms CP transfer from ^{1}H to ^{15}N (75–100% on ^{15}N). Selective transfer from ^{15}N to ^{13}C was achieved using an adiabatic CP transfer of 4 ms. The effective acquisition times were 20 ms and 12.8 ms for ^{13}C and ^{15}N, respectively. In total 128 increments were recorded in the ^{15}N dimension, with 64 scans per increment. Using a recycle delay of 3.0 s, the experimental time per spectrum was about 7 hours.

3D NCACX and 3D NCOCX[31] experiments were recorded at 600 MHz, at a MAS frequency of 10 kHz. A CP contact of 2 ms was employed, with 60 and 43 kHz rf-field on ^{1}H and ^{15}N, respectively, with a 75–100% ramp on protons. Polarization was transferred from ^{15}N to ^{13}Cα or ^{13}C′ with an adiabatic CP contact of 4 ms, with rf-fields of 35 and 25 kHz on ^{13}C and ^{15}N, respectively. For the NCACX spectra, PDSD mixing times of 35, 100, 200 and 500 ms were used; for the NCOCX spectra, PDSD mixing times of 35 and 200 ms were applied. The 3D spectra were obtained as data matrices of 1372 x 48 x 32, with effective acquisition times of 12 ms, 5.2 ms and 6.4 ms for ^{13}C, ^{13}Cα/^{13}C′ and ^{15}N, respectively. For each increment, 104 scans were averaged with a 2.8 s recycle delay. Hence, the total experimental time for each 3D spectrum amounted 124 hours. The 3D NCACB experiment was recorded under similar conditions, however with a DREAM mixing scheme of 3 ms to selectively exchange between Cα and Cβ[15,16]. The NCACB experiment was recorded as data matrix of 2484 x 64 x 48, with effective acquisition times of 20.0, 5.1 and 5.4 ms for ^{13}Cβ, ^{13}Cα and ^{15}N, respectively. A recycle delay of 3 s was used, and 56 scans were recorded per increment, leading to a total acquisition time of 144 hours.

The ^{15}N-^{13}C TEDOR experiment (transferred echo double resonance) was recorded at a field of 9.4 T, at 8 kHz MAS and at a temperature of 275 K. The two REDOR (rotational echo double resonance) mixing periods were 1.0 ms each. Carbon and nitrogen 180° pulses were 6.2 μs and 13.4 μs, respectively. Initial ^{13}C magnetization was created with a 2.5 ms ramped ^{1}H-^{13}C cross polarization (75–100% ramp on the ^{13}C channel). During mixing and acquisition, ^{1}H-heteronuclar decoupling was applied with a TPPM scheme[28], at a moderately high ^{1}H rf field strength of 75 kHz. The spectrum was recorded with 224 scans, 128 increments and effective acquisition times of 22 ms and 16 ms, in the ^{13}C and ^{15}N dimensions, respectively. With a recycle delay of 2.8 s, the experimental time for the 2D experiment amounted 23 hours.

^{13}C-^{13}C Methyl-filtered Post-C7 experiments were obtained at 900 MHz, at a spinning frequency of 8 kHz. Magnetization is prepared by using a methyl filter[20]; for this, a long ^{1}H-^{13}C cross polarization (CP) contact of 2 ms is directly followed by a short polarization inversion (PI) of 70 μs, achieved with a 180° phase change of the ^{1}H spin-lock pulse (CPPI)[32]. A polarization inversion period of 60 μs was found to be sufficient for 'depolarizing' non-methyl, protonated nuclei[33,34]. Following this preparation, a post-C7 scheme was used to obtain a 2D ^{13}C-^{13}C correlation spectrum (2Q/1Q)[21]. Acquisition time in the direct dimension was 12.0 ms, in the indirect (2Q) dimension 280 experiments with an increment of 32 μs were recorded.

2D CHHC[35] spectra were recorded at 700 MHz. The initial polarization was transferred from ^{1}H to ^{13}C using 1.5 ms CP with average rf-nutation frequencies of 60 and 50 kHz, on protons and carbons, respectively. SPINAL-64 decoupling with an rf-field strength of 85 kHz was applied during evolution and detection periods. Acquistion times in the direct and indirect dimensions were 12 and 7 ms with a 3.0 s recycle delay between individual transients. The second and third rectangular CP contacts were kept short at 60 μs. The ^{1}H-^{1}H mixing times used were 35, 50, 80, 150, 200, 300 and 500 μs.

1. Leo, J. C. & Skurnik, M. Adhesins of human pathogens from the genus Yersinia. *Adv. Exp. Med. Biol.* **715**, 1–15 (2011).
2. Linke, D., Riess, T., Autenrieth, I. B., Lupas, A. & Kempf, V. A. J. Trimeric autotransporter adhesins: variable structure, common function. *Trends Microbiol.* **14**, 264–270 (2006).
3. Wollmann, P., Zeth, K., Lupas, A. N. & Linke, D. Purification of the YadA membrane anchor for secondary structure analysis and crystallization. *Int. J. Biol. Macromol.* **39**, 3–9 (2006).
4. Lehr, U. *et al.* 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).
5. Grosskinsky, U. *et al.* A conserved glycine residue of trimeric autotransporter domains plays a key role in Yersinia adhesin A autotransport. *J. Bacteriol.* **189**, 9011–9019 (2007).
6. Leo, J. C. *et al.* First Analysis of a Bacterial Collagen-Binding Protein with Collagen Toolkits: Promiscuous Binding of YadA to Collagens May Explain How YadA Interferes with Host Processes. *Infect. Immun.* **78**, 3226–3236 (2010).
7. Smith, S. O. & Peersen, O. B. Solid-State Nmr Approaches for Studying Membrane-Protein Structure. *Annu. Rev. Bioph. Biom.* **21**, 25–47 (1992).
8. Martin, R. W. & Zilm, K. W. Preparation of protein nanocrystals and their characterization by solid state NMR. *J. Magn. Reson.* **165**, 162–174 (2003).
9. Shahid, S. A. *et al.* Membrane protein structure by solid-state NMR: insights into the autotransport mechanism of YadA. *Nature methods*, DOI: 10.1038/nmeth (2012).
10. Takegoshi, K., Yano, T., Takeda, K. & Terao, T. Indirect high-resolution observation of 14N NMR in rotating solids. *J. Am. Chem. Soc.* **123**, 10786–10787 (2001).
11. Morcombe, C. R., Gaponenko, V., Byrd, R. A. & Zilm, K. W. Diluting abundant spins by isotope edited radio frequency field assisted diffusion. *J. Am. Chem. Soc.* **126**, 7196–7197 (2004).
12. 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).
13. Hediger, S., Meier, B. H., Kurur, N. D., Bodenhausen, G. & Ernst, R. R. Nmr Cross-Polarization by Adiabatic Passage through the Hartmann-Hahn Condition (APHH). *Chem. Phys. Lett.* **223**, 283–288 (1994).
14. Szeverenyi, N. M., Sullivan, M. J. & Maciel, G. E. Observation of Spin Exchange by Two-Dimensional Fourier-Transform C-13 Cross Polarization-Magic-Angle Spinning. *J. Magn. Reson.* **47**, 462–475 (1982).
15. Pauli, J., Baldus, M., van Rossum, B., de Groot, H. & Oschkinat, H. Backbone and side-chain 13C and 15N signal assignments of the alpha-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 Tesla. *Chembiochem* **2**, 272–281 (2001).
16. Verel, R., Baldus, M., Ernst, M. & Meier, B. H. A homonuclear spin-pair filter for solid-state NMR based on adiabatic-passage techniques. *Chem. Phys. Lett.* **287**, 421–428 (1998).
17. Straus, S. K., Bremi, T. & Ernst, R. R. Resolution enhancement by homonuclear J decoupling in solid-state MAS NMR. *Chem. Phys. Lett.* **262**, 709–715 (1996).
18. Jaroniec, C. P., Tounge, B. A., Herzfeld, J. & Griffin, R. G. Frequency selective heteronuclear dipolar recoupling in rotating solids: accurate (13)C-(15)N distance measurements in uniformly (13)C,(15)N-labeled peptides. *J. Am. Chem. Soc.* **123**, 3507–3519 (2001).
19. Jaroniec, C. P., Filip, C. & Griffin, R. G. 3D TEDOR NMR experiments for the simultaneous measurement of multiple carbon-nitrogen distances in uniformly (13)C,(15)N-labeled solids. *J. Am. Chem. Soc.* **124**, 10728–10742 (2002).
20. Jehle, S. *et al.* Spectral editing: selection of methyl groups in multidimensional solid-state magic-angle spinning NMR. *J. Biomol. Nmr.* **36**, 169–177 (2006).
21. Hohwy, M., Jakobsen, H. J., Eden, M., Levitt, M. H. & Nielsen, N. C. Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence. *J. Chem. Phys.* **108**, 2686–2694 (1998).
22. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS plus : a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. Nmr.* **44**, 213–223 (2009).
23. Hiller, M. *et al.* [2,3-C-13]-labeling of aromatic residues-getting a head start in the magic-angle-spinning NMR assignment of membrane proteins. *J. Am. Chem. Soc.* **130**, 408 (2008).
24. Remmert, M., Biegert, A., Linke, D., Lupas, A. N. & Soding, J. Evolution of outer membrane beta-barrels from an ancestral beta beta hairpin. *Mol. Biol. Evol.* **27**, 1348–1358 (2010).
25. Korkmaz, F., Koster, S., Yildiz, O. & Mantele, W. The role of lipids for the functional integrity of porin: an FTIR study using lipid and protein reporter groups. *Biochemistry-Us.* **47**, 12126–12134 (2008).
26. Schaefer, J. & Stejskal, E. O. C-13 Nuclear Magnetic-Resonance of Polymers Spinning at Magic Angle. *J. Am. Chem. Soc.* **98**, 1031–1032 (1976).
27. Metz, G., Wu, X. L. & Smith, S. O. Ramped-Amplitude Cross-Polarization in Magic-Angle-Spinning Nmr. *J. Magn. Reson. A.* **110**, 219–227 (1994).
28. Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V. & Griffin, R. G. Heteronuclear Decoupling in Rotating Solids. *J. Chem. Phys.* **103**, 6951–6958 (1995).
29. Sinha, N. *et al.* SPINAL modulated decoupling in high field double- and triple-resonance solid-state NMR experiments on stationary samples. *J. Magn. Reson.* **177**, 197–202 (2005).
30. Suter, D. & Ernst, R. R. Spin diffusion in resolved solid-state NMR spectra. *Phys. Rev. B Condens. Matter.* **32**, 5608–5627 (1985).
31. Castellani, F., van Rossum, B. J., Diehl, A., Rehbein, K. & Oschkinat, H. Determination of solid-state NMR structures of proteins by means of three-dimensional 15N-13C-13C dipolar correlation spectroscopy and chemical shift analysis. *Biochemistry-Us.* **42**, 11476–11483 (2003).
32. Wu, X. L., Burns, S. T. & Zilm, K. W. Spectral Editing in Cpmas Nmr - Generating Subspectra Based on Proton Multiplicities. *J. Magn. Reson. A.* **111**, 29–36 (1994).
33. Opella, S. J., Frey, M. H. & Cross, T. A. Detection of Individual Carbon Resonances in Solid Proteins. *J. Am. Chem. Soc.* **101**, 5856–5857 (1979).
34. Opella, S. J. & Frey, M. H. Selection of Non-Protonated Carbon Resonances in Solid-State Nuclear Magnetic-Resonance. *J. Am. Chem. Soc.* **101**, 5854–5856 (1979).
35. Lange, A. *et al.* A concept for rapid protein-structure determination by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **44**, 2089–2092 (2005).
36. Gullion, T. & Schaefer, J. Rotational-Echo Double-Resonance Nmr. *J. Magn. Reson.* **81**, 196–200 (1989).

Acknowledgements

This work was supported by contract research 'Forschungsprogramm Methoden für die Lebenswissenschaften' of the Baden-Württemberg Stiftung to DL, DFG SFB 740 to B.J.v.R., and by institutional funds of the Leibniz Society and the Max Planck Society. The authors wish to thank Hartmut Oschkinat and Andrei Lupas for continuing support.

Author contribution

SAS, SM and BJvR conducted the majority of experiments. SAS analyzed and assigned the spectra. DL prepared the sample. SAS and BJvR wrote the manuscript. All authors read and approved the manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interest: The authors declare no competing financial interest.

How to cite this article: Shahid, S.A., Markovic, S., Linke, D. & van Rossum, B. Assignment and secondary structure of the YadA membrane protein by solid-state MAS NMR. *Sci. Rep.* **2**, 803; DOI:10.1038/srep00803 (2012).