Transmembrane segment enhanced labeling as a tool for the backbone assignment of α-helical membrane proteins

Sina Reckel*, Solmaz Sobhanifar*, Birgit Schneider*, Friederike Junge*, Daniel Schwarz*, Florian Durst*, Frank Löhr*, Peter Güntert*†, Frank Bernhard*, and Volker Dötsch*‡§

*Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance and ‡Cluster of Excellence Macromolecular Complexes, University of Frankfurt, Max-von-Laue Strasse 9, 60438 Frankfurt am Main, Germany; and †Frankfurt Institute for Advanced Studies, Ruth-Moufang-Strasse 1, 60438 Frankfurt am Main, Germany

Recent advances in cell-free expression protocols have opened a new avenue toward high-resolution structural investigations of membrane proteins by x-ray crystallography and NMR spectroscopy. One of the biggest challenges for liquid-state NMR-based structural investigations of membrane proteins is the significant peak overlap in the spectra caused by large line widths and limited chemical shift dispersion of α-helical proteins. Contributing to the limited chemical shift dispersion is the fact that ≈60% of the amino acids in transmembrane regions consist of only six different amino acid types. This principle disadvantage, however, can be exploited to aid in the assignment of the backbone resonances of membrane proteins; by ^{15}N/^{13}C-double-labeling of these six amino acid types, sequential connectivities can be obtained for large stretches of the transmembrane segments where number and length of stretches consisting exclusively of these six amino acid types are enhanced compared with the remainder of the protein. We show by experiment as well as by statistical analysis that this labeling scheme provides a large number of sequential connectivities in transmembrane regions and thus constitutes a tool for the efficient assignment of membrane protein backbone resonances.

cell-free expression | NMR | stable isotope labeling | protein structure analysis | peak overlap

Structure determination of membrane proteins is currently one of the biggest challenges in the field of structural biology. Despite significant technical improvements in the three main techniques used to determine high-resolution structures of proteins—x-ray crystallography, NMR spectroscopy, and electron microscopy—the hydrophobic nature of membrane proteins still poses tremendous difficulties. One of the biggest hurdles is to establish methods for obtaining the necessarily large amounts of homogenous and pure membrane proteins solubilized in detergent micelles. Many different expression systems have been tested and established including bacterial, yeast, insect, and mammalian cells. However, only in a few cases have these expression systems yielded expression levels that are comparable to those of soluble proteins. In particular, eukaryotic membrane proteins are difficult to obtain in the amounts necessary for structural studies, which is also reflected by the fact that fewer than 10 structures of eukaryotic membrane proteins have been determined so far.

As an alternative to cellular expression systems, cell-free expression systems have been developed and established as a standard source for soluble protein expression in many laboratories (1–3). Recently, we have shown that cell-free expression systems can also be used for the expression of high amounts of membrane proteins and that, in many cases, the cell-free method is superior to cellular-based systems in terms of obtaining solubilized membrane proteins (4–6). By adding detergents directly to the reaction mixture, cell-free expression systems are capable of directly producing micelle-solubilized membrane proteins that do not have to be extracted from a membrane or refolded from insoluble inclusion bodies (5). Functional studies with several membrane proteins obtained by cell-free expression have demonstrated that this expression method yields functional proteins (4, 7–11). In these studies membrane proteins obtained by cell-free expression, such as EmrE (4, 9) or eukaryotic organic cation and anion transporters of the SLC22 protein family (11), have shown biochemical behavior virtually identical to samples expressed in cells.

For structure determination by NMR spectroscopy cell-free expression systems provide additional unique advantages. The first step in the structure determination of both soluble and membrane proteins is always the sequential assignment of the backbone resonances for which many different types of multi-dimensional experiments are available. This traditional approach, however, is limited in cases of severe resonance overlap as encountered in unfolded proteins or α-helical membrane proteins. Although the high internal flexibility of an unfolded polypeptide chain leads to slow relaxation, thereby enabling the use of pulse sequences with more and longer transfer delays for resolving overlap (12), for large membrane proteins solubilized in micelles, only the most basic experiments yield enough sensitivity. In a recent publication we demonstrated that the resulting overlap problem can be overcome, allowing for the sequential backbone assignment of membrane proteins >200 aa (13, 14). Our strategy was based on a two-step approach in which we first use standard triple-resonance experiments to obtain as many assignments as possible and then use a double-labeling scheme to select amino acid pairs with a two-dimensional version of the HNCO experiment to sequence-specifically identify additional amino acids (15–17). By using our cell-free expression system we have established a combinatorial labeling scheme that allows us to obtain the amino acid type selective labeled samples in basically any combination.

This method works most efficiently if as many of the backbone resonances as possible have already been assigned with the nonselective triple-resonance experiments. Approaches trying to use exclusively combinatorial amino acid type selective labeling have revealed that, on average, only 50–60% of the amino acid pairs within a protein are unique, with the rest occurring twice or even more times in a protein sequence (16). [Under the

FLAP 161 aa

MDQETVGNVVLLAIVTLISVVQNGFFAHKVEHESRTQNGRSFQRTGTLAFERVYTANQNCVDAYPTFLAVLWSAGLLCSQVPAAFAGLMYLFVRQKYFVGY

LGERTQSTPGYIFGKRIILFLFLMSVAGIFNYYLIFFFGSDFENYIKTISTTISPLLLIS

Presenilin - CTF 176 aa

MVWLVNMAEGDPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGL

GDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI

DTehA 219 aa

MQSDKVLNLPAGYFGIVLGTIGMGFAWRYASQVWQVSHWLGDGLVILAMIIWGLLTSAFIARLIRFPHSVLAEVRHPVLSSFVSLFPATTMLVAIGFVPW

FRPLAVCLFSFGVVVQLAYAAWQTAGLWRGSHPEEATTPGLYLPTVANNFISAMACGALGYTDAGLVFLGAGVFSWLSLEPVILQRLRSSGELPTALRT

SLGIQLAPALVACSAWLSVN

Yfik 195 aa

MTPTLLSAFWTYTLITAMTPGPNNILALSSATSHGFRQSTRVLAGMSLGFLIVMLLCAGISFSLAVIDPAAVHLLSWAGAAYIVWLAWKIATSPTKEDGL

QAKPISFWASFALQFVNVKIILYGVTALSTFVLPQTQALSWVVGVSVLLAMIGTFGNVCWALAGHLFQRLFRQYGROLNIVLALLLVYCAVRIFY

Endothelin B Receptor 442 aa

MQPPPSLCGRALVALVLACGLSRIWGEERGFPPDRATPLLQTAEIMTPPTKTLWPKGSNASLARSLAPAEVPKGDRTAGSPPRTISPPPCQQGPIEIKETF

KYINTVVSCLVFVLGIIGNSTLLRIIYKNKCMRNGPNILIASLALGDLLHIVIDIPINVYKLLAEDWPFGAEMCKLVPFIQKASVGITVLSLCALSIDRY

RAVASWSRIKGIGVPKWTAVEIVLIWVVSVVLAVPEAIGFDIITMDYKGSYLRICLLHPVQKTAFMQFYKTAKDWWLFSEYFCLPLAITAFFYTLMTCEM

LRKKSGMQIALNDHLKQRREVAKTVFCLVLVFALCWLPLHLSRILKLTLYNQNDPNRCELLSFLLVLDYIGINMASLNSCINPIALYLVSKRFKNCFKSC
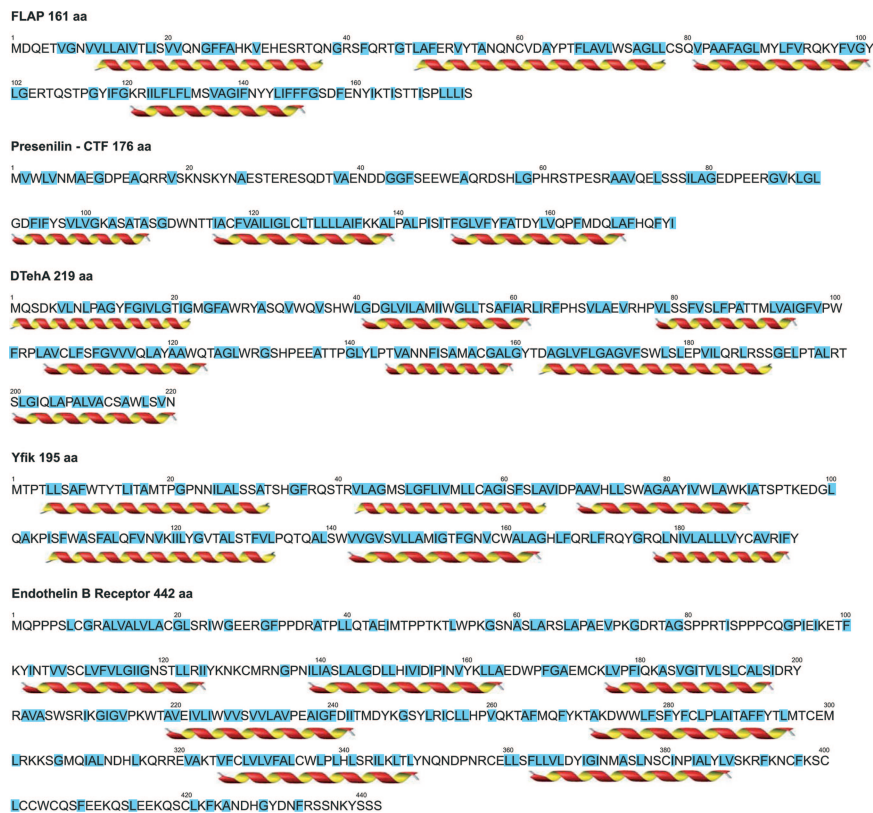
LCCWCQSFEEKQSLEEKQSCLKFKANDHGYDNFRSSNKYSSS

Fig. 1. Amino acid sequences of the five α-helical membrane proteins, FLAP, presenilin1-CTF, TehA (1–219), YfiK, and endothelin B receptor. The six amino acid types AFGILV are highlighted in blue. The locations of the transmembrane helices are indicated below the sequences.

simplifying assumption that all 20 aa types are distributed uniformly in a protein sequence of length L, a given dipeptide occurs uniquely with probability $p(L) = (1 - 1/20^2)^{L-2}$, e.g., with a probability of 78% in a 100-residue protein, and 54% in a protein of 250 residues. In a real protein in which the different amino acids are distributed unevenly, the uniqueness is on average even lower.] This degeneracy in the sequence mapping of amino acid pairs precludes the exclusive use of this technique for obtaining nearly complete backbone assignments. In combination with nonselective standard NMR experiments it is nonetheless a robust way to obtain the additional information. In larger membrane proteins or membrane proteins with extensive resonance overlap, however, obtaining the necessary sequential assignments remains very challenging. Here, we demonstrate that additional labeling schemes can be used to simplify the NMR spectra of membrane proteins, thus providing a new assignment tool for the structural investigation of large and complex membrane proteins.

Results

The difficulty in assigning membrane proteins arises from the generally broad resonance lines, the limited chemical shift dispersion, and the fact that the transmembrane helices are predominantly composed of only a small number of amino acid types: The six amino acids, alanine (A), phenylalanine (F), glycine (G), isoleucine (I), leucine (L), and valine (V), account for >60% of the residues in the transmembrane helices (18) (Fig. 1), whereas they represent only one third of the residues in other regions. This principle disadvantage can, however, be exploited to simplify the NMR spectra of membrane proteins and to accelerate the backbone assignment process. Exclusive $^{13}C/^{15}N$ labeling of the amino acid types above (AFGILV) is expected to significantly reduce peak overlap while preserving a high degree of information for the backbone assignment of the transmembrane segments (Table 1). Because the labeling will occur predominantly in transmembrane segments, we refer to this approach as the transmembrane segment-enhanced labeling

Table 1. Overview of proteins

Protein	Origin	Family/function	Size*	%TMS[†]	% AFGILV in TMS[‡]	% AFGILV entire protein[§]
FLAP	Human	MAPEG	161	63	59	51
Presenilin1-CTF	Human	Membrane-bound protease	176	36	68	45
TehA	Bacterial	Heavy metal resistance protein	219	64	63	56
YfiK	Bacterial	Cysteine exporter	195	73	65	56
Endothelin B receptor	Human	GPCR	422	40	64	43

*Number of amino acids.
[†]Percentage of amino acids in the transmembrane helices.
[‡]Percentage of the amino acids AFGILV in the transmembrane helices.
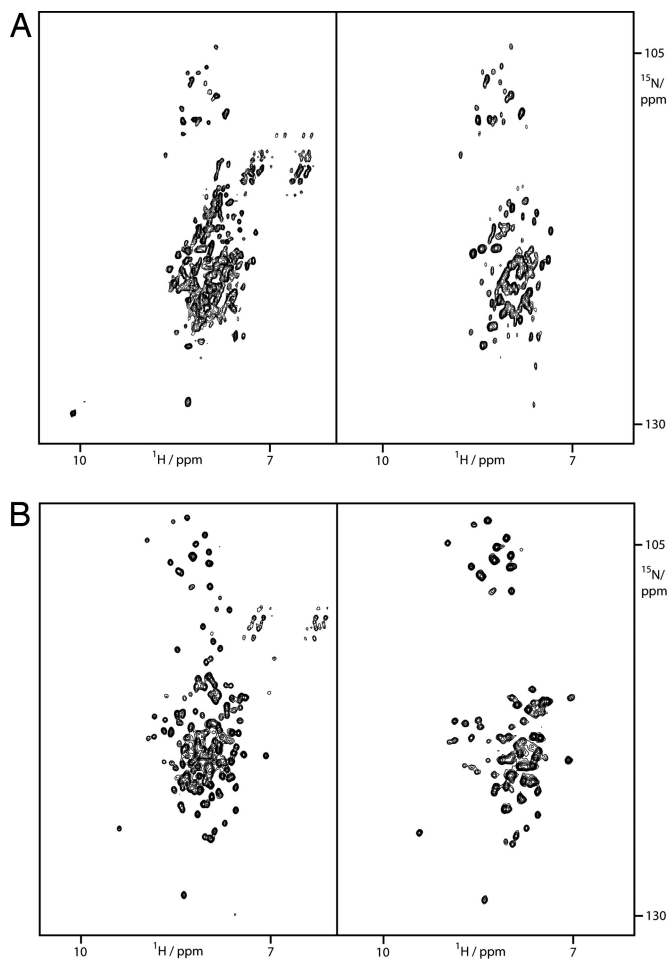[§]Percentage of the amino acids AFGILV in the entire protein sequence.

Fig. 2. Comparison of [^{15}N, ^{1}H]-TROSY spectra of uniformly ^{15}N-labeled (*Left*) and TMS-labeled (*Right*) of FLAP (*A*) and presenilin1-CTF (*B*).

Fig. 3. Percentage of coverage of protein amino acid sequences by segments composed exclusively of the six amino acid types AFGILV. The solid red curve gives the coverage of the complete amino acid sequences of the five helical membrane proteins FLAP, YfiK, presenilin1-CTF, endothelin B receptor, and TehA. The upper and lower broken red curves show the coverage of the transmembrane helices and of the other regions, respectively, in these five membrane proteins. For comparison, the coverage of the complete amino acid sequences of the three β-sheet membrane proteins OmpA, OmpG, and OmpX, are shown in blue, and the coverage of the complete amino acid sequences of the five soluble proteins DFPase, thioredoxin, ubiquitin, the DNA-binding domain of p53, and the Rel homology domain of NFκB are shown in black.

(TMS-labeling) scheme. The reduction of peak overlap by TMS-labeling can be seen in Fig. 2 in which TROSY spectra of uniformly and TMS-labeled samples are compared for two integral membrane proteins, the 5-lipoxygenase-activating protein (FLAP) and the C-terminal fragment of presenilin1 (presenilin1-CTF). FLAP is a member of the MAPEG family (membrane-associated proteins in eicosanoid and glutathione metabolism). It has 161 aa that fold into four transmembrane helices (19). Presenilin1-CTF is the C-terminal 176-aa natural cleavage fragment of the membrane-integrated protease presenilin1, which plays a major role in processing the amyloid precursor protein (APP).

TMS-labeling would not provide any advantage, if these six amino acid types were not predominately clustered in the transmembrane helices, where stretches consisting exclusively of the six amino acids AFGILV are frequently found. This clustering provides a significantly increased number of sequential connectivities in the transmembrane regions compared with loops connecting the transmembrane helices and soluble proteins where these six amino acids are distributed more evenly throughout the sequence. This is illustrated in Fig. 3 for five α-helical integral membrane proteins, FLAP, YfiK, presenilin1-CTF, endothelin B receptor, and TehA, in which between 54 and 85% of the amino acids AFGILV are located in transmembrane segments. The difference between the transmembrane segments of membrane proteins, on the one hand, and other regions of membrane proteins and soluble proteins, on the other hand,
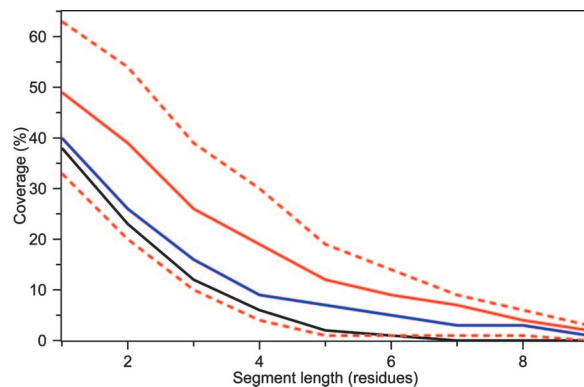
becomes even more pronounced when stretches consisting of several of these amino acids are considered. Fig. 3 compares the percentage of the entire protein sequence consisting of stretches of increasing length made up exclusively of these six amino acids AFGILV among the five membrane proteins mentioned above and the five soluble proteins: DFPase, thioredoxin, ubiquitin, the DNA-binding domain of p53, and the RHD (Rel homology domain) region of NFκB. In particular, longer stretches of these six amino acids occur significantly more often and hence cover a larger percentage of the entire sequence in helical membrane proteins than in others. This holds especially for the transmembrane helices where TMS-labeled stretches of three or more amino acid residues cover almost 40% of the sequence. Interestingly, Fig. 3 also shows that the amino acid composition of the β-barrel membrane proteins OmpA, OmpX, and OmpG is more similar to the composition found in soluble proteins than in α-helical membrane proteins.

To investigate the benefits of TMS-labeling we measured HNCA and HNCOCA experiments on both uniformly and TMS-labeled samples of the two proteins, FLAP and presenilin1-CTF. Fig. 4 compares for each protein a ^{13}C$^{\alpha}$, ^{1}H^{N} plane taken at the same ^{15}N chemical shifts as the corresponding plane from an HNCA experiment measured with the TMS-labeled sample. These spectra demonstrate that the reduced overlap from the TMS-labeled samples can be exploited to obtain sequential connectivities that cannot be extracted from the uniformly labeled sample because of severe spectral overlap. In the spectra of the C-terminal fragment of presenilin1, 40 of 42 expected sequential connectivities could be identified in the TMS-labeled sample. In the case of FLAP, 50 of 51 expected sequential connectivities could be identified in the HNCA spectrum of the TMS-labeled sample, but, because of observed structural heterogeneity, some of these will represent the same amino acid in different conformational states. With this labeling strategy 38% of the amino acids in the transmembrane helices of FLAP, corresponding to 24% of the total protein, and 59% of the transmembrane part of presenilin1-CTF (23% of the total protein) could thus be assigned. In addition, this labeling scheme yielded assignments in the loop regions, covering 10% of the total loop region (4% of the entire protein sequence) for FLAP and 17% (11% for the entire protein sequence) for presenilin1-CTF.
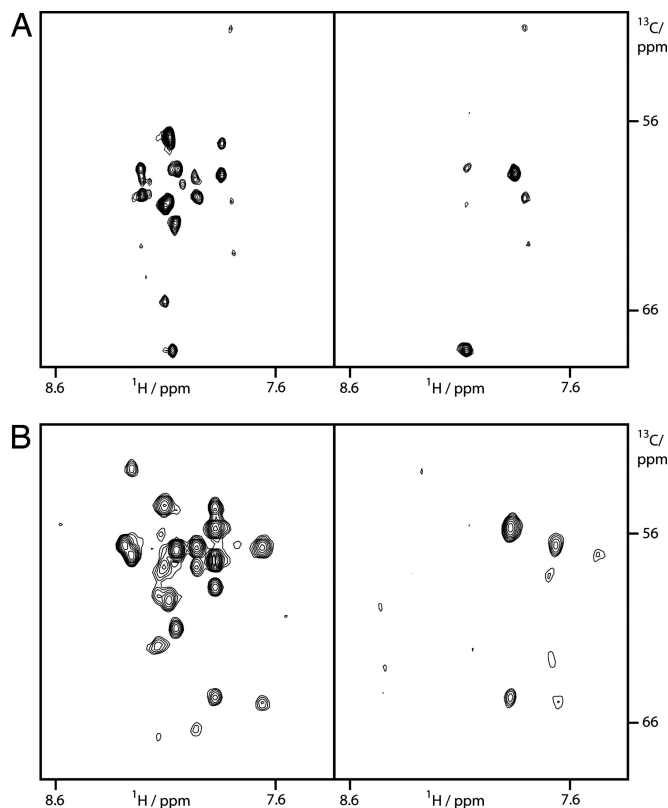
Fig. 4. Sections taken from [^{13}C, ^{1}H]-planes of [^{15}N, ^{1}H]-TROSY HNCA spectra measured with uniformly ^{15}N/^{13}C-labeled (*Left*) and TMS-labeled (*Right*) samples of FLAP (*A*) and presenilin1-CTF (*B*).

Fig. 5. Uniqueness of the sequence mapping of segments of a given length or longer that are composed exclusively of the six amino acid types AFGILV, computed for the five helical membrane proteins (red), the three β-sheet membrane proteins (blue), and the five soluble proteins (black) of Fig. 3. The solid lines were computed under the assumption that the six amino acid types AFGILV can be distinguished unambiguously from each other. The dashed curve shows the uniqueness of the sequence mappings for the helical membrane proteins assuming that only the three subsets of amino acid types, G, IV, and AFL, can be distinguished unambiguously from each other.

Discussion

NMR-based structure determination of small- and medium-sized soluble proteins has become a standard tool thanks to the development of many sophisticated NMR techniques over the past 20 years. Nevertheless, most of these powerful techniques are not applicable to membrane proteins because of their large size (including the micelle) and limited chemical shift dispersion (20). Because of the large effective size of the micelle-solubilized membrane proteins only the most basic and robust NMR pulse sequences with the shortest coherence transfer delays can be used. This limitation often prevents the use of pulse sequences such as HNCACB, CBCACONH, or CBCANH that transfer the magnetization further along the side chain of the amino acids to decrease the resonance overlap and to identify amino acid types. To obtain almost complete backbone assignments of membrane proteins despite the significant resonance overlap we have started to develop and to employ amino acid type selective labeling schemes. One such scheme, the TMS-labeling described above, leads to a rough partitioning of the resonances of a membrane protein in the transmembrane segments and the solvent-exposed loops. Because the transmembrane segments predominantly consist of only six different amino acid types, they can be efficiently labeled selectively. The reduced resonance overlap allows us to obtain the assignments of a large percentage of these transmembrane helices by applying only the most robust and most sensitive pulse sequences, HNCA and HNCOCA.

Identifying sequential amino acid stretches is only the first step of the assignment process. In the second step these stretches have to be mapped onto the amino acid sequence of the protein. In principle, the use of only a small number of different amino acids makes this mapping process more difficult because fewer unique sequences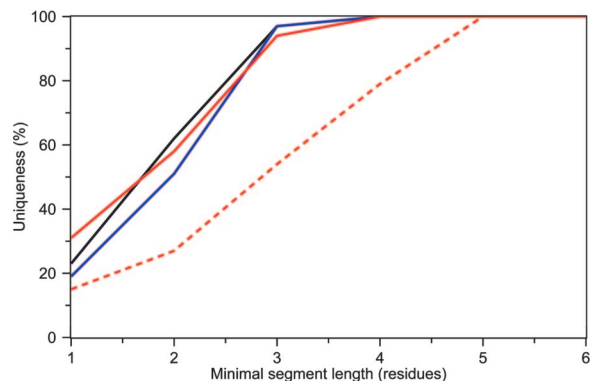 are available (21). Fig. 5 shows a plot of the unique-ness of TMS-labeled sequence stretches of a given minimal length within the protein sequence for α-helical membrane proteins, β-barrel membrane proteins and soluble proteins. This figure demonstrates that in all three classes the sequence mapping of TMS-labeled stretches of three or more sequentially following amino acids is virtually unique. This result is valid if all six different amino acid types AFGILV can be distinguished. In practice, however, the backbone chemical shifts allow differentiation of only a classification into three groups of amino acids, G, IV, AFL. Glycine has unique chemical shifts both in the ^{15}N and in the ^{13}C$^\alpha$ dimensions. The ^{13}C$^\alpha$ chemical shifts of valine and isoleucine are sufficiently separated from those of the other 4 aa to be able to assign them as either of these 2 aa. In the third group comprising alanine, leucine, and phenylalanine, overlap occurs, in particular, between leucine and phenylalanine. Calculating uniqueness by using only these three distinguishable groups shows that the stretch length required for unique sequence mapping increases to five (Fig. 5). In practice this means that for smaller stretches additional information is necessary to achieve unambiguous mapping of the fragments onto the protein sequence. This additional information can either be obtained from uniformly labeled samples or from further amino acid type selective labeling. With the development of the TMS-labeling scheme we have now modified and further optimized our originally suggested protocol for the backbone assignment of membrane proteins. In our current protocol, we produce both TMS- and uniformly labeled samples to obtain as many assignments as possible from HNCA and HNCOCA experiments. To obtain assignments for both completely unassigned parts of the sequence, as well as for stretches that could not yet be mapped unambiguously onto the sequence, we use the combinatorial labeling approach described in ref. 13. Combining these strategies enabled us to obtain 90% of the backbone resonance assignments of the C-terminal fragment of presenilin1-CTF.

In principle, selective labeling is possible for the loop regions as well. However, in contrast to the transmembrane segments the amino acid composition of the loops is more diverse and labeling with >6 aa types will, in most cases, be necessary to obtain a reasonable number of sequential connectivities. Although labeling with more than six amino acid types is technically simple, the high costs of double-labeled amino acids make it uneconomical. In some proteins, however, clustering of certain amino acid types in the loops might occur that can be exploited for selective

labeling of these regions. Similarly, in some membrane proteins the amino acid composition of the transmembrane segments could differ from the average, making necessary the adjustment of both the number and types of amino acids for TMS-labeling. In general, however, TMS-labeling will be most useful based on the six amino acid types mentioned above.

One prerequisite for the use of sophisticated labeling schemes is the availability of a cell-free expression system. Cell-based expression systems suffer not only from (often) low expression yields for membrane proteins, but also from metabolic scrambling, thus making clean amino acid type specific labeling difficult. In principle, auxotrophic bacterial strains are available for all 20 aa (22), but their expression yield is often reduced relative to wild-type bacteria, aggravating the problem of low expression yield of membrane proteins in such systems. Nonetheless, a reverse labeling scheme based on cellular expression systems as suggested by Baldus and coworkers can simplify spectra to a certain degree. By using sensory rhodopsin II as an example, it was shown that addition of unlabeled amino acids to an otherwise uniformly labeled medium leads to significant peak reduction. In contrast to cell-based expression systems metabolic scrambling is virtually absent in cell-free expression systems, enabling selective labeling in almost any combination. Only in the case of glutamate and aspartate, scrambling might cause problems (depending on the source of extract), which can, however, be overcome by using transaminase inhibitors.

The resonance overlap of membrane proteins can also be reduced by methods other than selective labeling schemes. Bax and coworkers have demonstrated with the tetrameric potassium ion transporter KcsA in SDS micelles that loop regions and transmembrane regions can be distinguished by their different amide proton exchange rates (23). Dissolving the protein in D_2O resulted in amide proton signals exclusively from the transmembrane regions whereas the loop regions became invisible. Expression of the protein in deuterated media and solvent exchange into H_2O led to the selective detection of the loop regions. Similar experiments in our laboratory with the three membrane proteins TehA, FLAP, and presenilin1-CTF, however, revealed that 80–90% of their resonances (including the transmembrane segments) disappeared within 16 h. In the case of FLAP all remaining resonances belonged to amino acids located in the first transmembrane helix and in presenilin1-CTF to residues in the second transmembrane helix. The reduced overlap of these spectra allowed us to confirm our assignments and to obtain additional information about the structure of the proteins. In the case of CTF the first transmembrane helix harbors an aspartic acid that is part of the active site. The third helix is kinked by a proline. Both the aspartic acid and the proline in the middle of the transmembrane segments reduce the stability of those helices, resulting in relatively fast proton exchange. These results show that differences in exchange rates of amide protons can be used as an alternative assignment method for very stable proteins (such as KcsA or β-barrel proteins) or for obtaining further structural information in other proteins. For even larger membrane proteins TMS-labeling and D_2O-exchange experiments might be combined to further reduce spectral overlap. Our exchange experiments have demonstrated that entire helices are either stable or exchange fast. Because these stable helices will contain stretches of TMS-labeled amino acids, large stretches of stable transmembrane helices can thus be assigned in TMS-labeled and D_2O-exchanged samples.

Another approach that reduces the spectral overlap and can provide complementary information is the titration with paramagnetic spin labels as suggested by Wüthrich and coworkers (24). Although a hydrophilic probe will affect residues in solvent-exposed regions, a hydrophobic paramagnetic compound causes relaxation of residues within the micellar environment. Conducting different sets of experiments by using spin labels of

different polarity will thus result in less crowded spectra with complementary information. In addition to reducing the spectral overlap, thus helping with the assignment, these experiments also reveal details about the protein–lipid interaction. We have titrated both FLAP and CTF with different spin labels and use this information to obtain additional information about the length of the helices.

The unique advantage of our TMS-labeling scheme, however, is that it provides a "clean" and stable sample environment that does not depend on differences in exchange rates or differences in the distribution of spin labels between the aqueous solution and the micelles which can result in some peaks just being broadened but not completely absent from the spectra. Another advantage of the TMS-labeling strategy is that potential assignments of peaks are confined to six rather than all 20 amino acid types.

In conclusion, TMS-labeling provides advantages that can be combined with additional, complementary approaches to further reduce spectral overlap for large membrane proteins or to provide structural information.

Methods

Protein Expression. Presenilin1-CTF (residues 291–467) and FLAP were expressed in a S30-based continuous exchange cell-free system (1). Presenilin1-CTF was cloned into a modified pet21a vector having a C-terminal His-tag with 10 histidines, FLAP was cloned into the pBH4 vector. Expression of all proteins followed published procedures and yielded an average of 1.5 mg of protein per ml of reaction mixture for both proteins. Before producing NMR samples the expression yield of each protein was optimized by screening for the optimal potassium and magnesium concentrations on an analytical scale of a 50-μl reaction volume. For the production of the TMS-labeled samples, $^{15}N/^{13}C$ double-labeled amino acids (A, F, G, I, L, V) were purchased from CIL and used at the same final concentration of 0.5 mM as the unlabeled amino acids.

FLAP was produced solubly in the presence of 0.1% (wt/vol) Brij35 in the reaction mixture. For purification and detergent exchange into LPPG (1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-RAC-(1-glycerol)]) the reaction mixture was diluted 1:10 in a phosphate buffer (20 mM Na_2HPO_4, 400 mM NaCl, 0.1% Brij 35, pH 7.5) before loading onto a Ni-chelate column. After washing the column with 5 volumes of 50 mM imidazole in Brij35 containing buffer, the detergent exchange was achieved by using 20 column volumes of buffer with 0.05% LPPG. The protein now in LPPG micelles was eluted from the column in 350 mM imidazole. Subsequent buffer exchange into a suitable NMR buffer (20 mM Mes, 25 mM Bis-Tris, 1 mM DTT, pH 6.0) was achieved by overnight dialysis at 4°C. The protein solution was then concentrated to ≈300 μl by using a stirred cell (Amicon) with a YM10 membrane (MWCO 10 kDa) and the final LPPG concentration was adjusted to 2% and protease inhibitors were added to prevent proteolysis.

Presenilin1-CTF was expressed as a precipitate and solubilized by adding 150 mM SDS in buffer containing 20 mM BisTrisPropane (pH 6.8) and 20 mM KCl. Because of sufficient purity of the precipitate form of the protein, further purification steps were not required. Homogeneity of the sample was also tested by gel filtration showing a single peak in the elution profile.

Rotational correlations times for both proteins were measured by cross-correlated relaxation according to Wüthrich and coworkers (25). FLAP in LPPG micelles has a rotational correlation time of 16 ns and presenilin1-CTF in SDS micelles 8 ns. Because all amide protons of a protein are used to calculate the rotational correlation time these data represent mere lower limits. In particular, the short correlation time of Presenilin1-CTF can be explained by this effect because approximately half of the protein consists of unfolded loops. Flap was further characterized by pulsed field gradient-based diffusion measurements resulting in a diffusion coefficient of $9.2 \times 10^{-11} m^2/s$.

NMR Spectroscopy. Spectra were measured on Bruker Avance spectrometers operating at proton resonance frequencies of either 700 MHz, 800 MHz, or 900 MHz. All spectrometers were equipped with cryogenic probes. Experiments with FLAP were performed at 309 K with protein concentrations varying between 200 and 250 μM. TROSY-HSQC spectra of the uniformly labeled sample were recorded with eight scans per increment and spectral widths of 12 ppm (^1H) and 50 ppm (^{15}N). For the TMS-labeled sample the TROSY-HSQC was recorded with 24 scans per increment and 30 ppm spectral width in the nitrogen dimension. The [^{15}N,^1H]-TROSY HNCA spectra (26) were recorded

with 16 scans per increment for both samples and acquisition times of 50, 6.4, and 42 ms for ^{1}H, ^{13}C, and ^{15}N, respectively. The [^{15}N,^{1}H]-TROSY HNCOCA experiments (27) were run with 24 scans per increment for the uniformly labeled sample and 64 scans for the TMS-labeled sample.

NMR samples of presenilin1-CTF contained 0.3–0.5 mM ^{15}N/^{13}C TMS- or uniformly labeled protein in conditions described. All experiments were performed at 313 K. Backbone assignments were obtained by using 3D TROSY-based HNCA and HN(CO)CA experiments with respective acquisition times of 54, 8.6, and 34 ms, and 16 scans per increment. All NMR spectra were processed by using Topspin (Bruker) and analyzed by using CARA (www.nmr.ch) and Sparky software (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

1. Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YB (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* 242:1162–1164.
2. Kigawa T, Muto Y, Yokoyama S (1995) Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J Biomol NMR* 6:129–134.
3. Ozawa K, Dixon NE, Otting G (2005) Cell-free synthesis of 15N-labeled proteins for NMR studies. *IUBMB Life* 57:615–622.
4. Klammt C, et al. (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem* 271:568–580.
5. Klammt C, et al. (2005) Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *FEBS J* 272:6024–6038.
6. Klammt C, et al. (2006) Cell-free expression as an emerging technique for the large scale production of integral membrane protein. *FEBS J* 273:4141–4153.
7. Ishihara G, et al. (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expression Purif* 41:27–37.
8. Berrier C, et al. (2004) Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry* 43:12585–12591.
9. Chen YJ, et al. (2007) X-ray structure of EmrE supports dual topology model. *Proc Natl Acad Sci USA* 104:18999–19004.
10. Sansuk K, et al. (2008) GPCR Proteomics: Mass spectrometric and functional analysis of histamine H1 receptor after baculovirus-driven and in vitro cell free expression. *J Proteome Res* 7:621–629.
11. Keller T, et al. (2008) Cell free-expression and functional reconstitution of eucaryotic drug transporters. *Biochemistry*, 47:4552–4564.
12. Hiller S, Fiorito F, Wuthrich K, Wider G (2005) Automated projection spectroscopy (APSY). *Proc Natl Acad Sci USA* 102:10876–10881.
13. Trbovic N, et al. (2005) Efficient strategy for the rapid backbone assignment of membrane proteins. *J Am Chem Soc* 127:13504–13505.
14. Koglin A, et al. (2006) Combination of cell-free expression and NMR spectroscopy as a new approach for structural investigation of membrane proteins. *Magn Reson Chem* 44(Spec No):S17–S23.
15. Shi J, Pelton JG, Cho HS, Wemmer DE (2004) Protein signal assignments using specific labeling and cell-free synthesis. *J Biomol NMR* 28:235–247.
16. Parker MJ, Aulton-Jones M, Hounslow AM, Craven CJ (2004) A combinatorial selective labeling method for the assignment of backbone amide NMR resonances. *J Am Chem Soc* 126:5020–5021.
17. Kainosho M, Tsuji T (1982) Assignment of the three methionyl carbonyl carbon resonances in Streptomyces subtilisin inhibitor by a carbon-13 and nitrogen-15 double-labeling technique. A new strategy for structural studies of proteins in solution. *Biochemistry* 21:6273–6279.
18. Senes A, Gerstein M, Engelman DM (2000) Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J Mol Biol* 296:921–936.
19. Ferguson AD, McKeever BM, Xu S, Wisniewski D, Miller DK, et al. (2007) Crystal structure of inhibitor-bound human 5-lipoxygenase-activating protein. *Science* 317:510–512.
20. Krueger-Koplin RD, et al. (2004) An evaluation of detergents for NMR structural studies of membrane proteins. *J Biomol NMR* 28:43–57.
21. Wüthrich K (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
22. Waugh DS (1996) Genetic tools for selective labeling of proteins with a-^{15}N-aminoacids. *J Biomol NMR* 8:184–192.
23. Chill JH, Louis JM, Miller C, Bax A (2006) NMR study of the tetrameric KcsA potassium channel in detergent micelles. *Protein Sci* 15:684–698.
24. Hilty C, Wider G, Fernandez C, Wüthrich K (2004) Membrane protein-lipid interactions in mixed micelles studied by NMR spectroscopy with the use of paramagnetic reagents. *ChemBioChem* 5:467–473.
25. Lee D, Hilty C, Wider G, Wüthrich K (2006) Effective rotational correlation times of proteins from NMR relaxation interference. *J Magn Reson* 178:72–76.
26. Salzmann M, Pervushin K, Wider G, Senn H, Wüthrich K (1998) TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. *Proc Natl Acad Sci USA* 95:13585–13590.
27. Salzmann M, Wider G, Pervushin K, Senn H, Wüthrich K (1999) TROSY-type triple resonance experiments for sequential NMR assignments of large proteins. *J Am Chem Soc* 121:844–848.

BIOPHYSICS