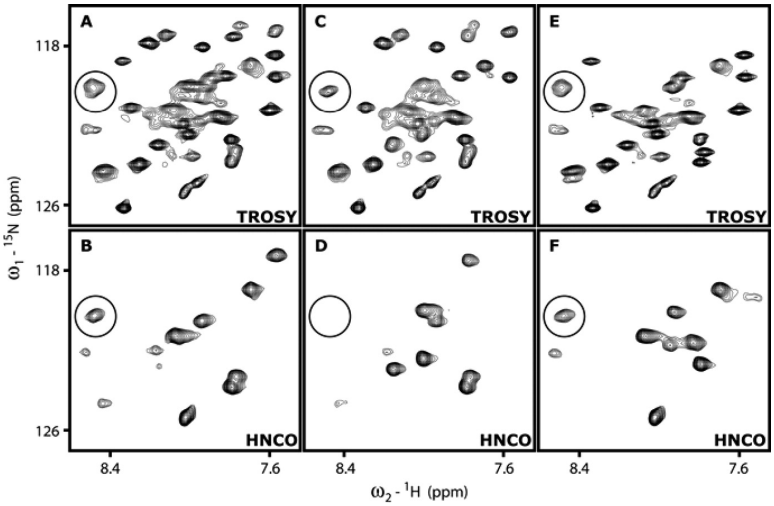
Communication

Efficient Strategy for the Rapid Backbone Assignment of Membrane Proteins

Nikola Trbovic, Christian Klammt, Alexander Koglin, Frank Lhr, Frank Bernhard, and Volker Dtsch

J. Am. Chem. Soc., **2005**, 127 (39), 13504-13505• DOI: 10.1021/ja0540270 • Publication Date (Web): 08 September 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 8 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 09/08/2005

Efficient Strategy for the Rapid Backbone Assignment of Membrane Proteins

Nikola Trbovic, Christian Klammt, Alexander Koglin, Frank Löhr, Frank Bernhard, and Volker Dötsch*

Institute for Biophysical Chemistry, University of Frankfurt and Center for Biomolecular Magnetic Resonance, 60439 Frankfurt, Germany

Received June 17, 2005; E-mail: vdoetsch@em.uni-frankfurt.de

Membrane proteins represent one of the biggest challenges in the area of structural biology based on the fact that they are hard to express, to purify, and difficult to analyze by high-resolution structural methods. Recently, the development of efficient cell-free transcription/translation protocols for the expression of milligram amounts of membrane proteins that cannot be expressed in sufficient quantities in vivo has opened a new avenue toward high-resolution structural investigations by X-ray crystallography and NMR spectroscopy.[1,2] To understand the function of membrane proteins, detailed investigations of their structure and interaction with their binding partners are necessary. NMR spectroscopy is an ideal tool for such investigations since it can provide both the structure and information about binding sites through chemical shift mapping. The basis for such detailed investigations is, however, the assignment of a protein's resonances. Unfortunately, α-helical proteins tend to display narrower chemical shift dispersion[3−5] as compared to that of proteins containing *β*-sheets. Since the majority of integral membrane proteins consists exclusively of α-helices, their NMR spectra tend to show a significant degree of peak overlap. The overlap problem is further aggravated by the often considerable size of the proteins and further enhanced by the fact that the proteins have to be solubilized in detergent micelles, which significantly increases the molecular weight of the protein/micelle particles, resulting in broader line width. Combined, these disadvantages of membrane proteins pose a considerable challenge for the chemical shift assignment, suggesting that new assignment strategies might be necessary in order to make backbone assignment of membrane proteins as routine a task as backbone assignment of soluble proteins. In this communication, we describe our efforts to assign the completely α-helical integral membrane protein TehA based on a combination of standard heteronuclear triple resonance experiments and a combinatorial labeling scheme.

The bacterial protein TehA is a 36 kDa membrane protein that shows limited homology to the family of small multidrug resistance proteins (SMR).[6] Its overexpression in bacteria confers resistance to tellurite compounds as well as to lipophilic cationic dyes. In vivo experiments have demonstrated that a 24 kDa fragment of TehA, which contains seven out of the 10 predicted transmembrane helices, shows the same biological effects as the full length protein. Therefore, we have focused on this 24 kDa fragment. For the expression and labeling with NMR active isotopes, we have employed an in vitro transcription/translation system based on *E. coli* S30 extracts, which yields 3 mg of TehA protein per milliliter of reaction volume.[1] Figure 1 shows the [^{15}N,^{1}H] TROSY spectrum of a ^{2}H/^{13}C/^{15}N triple-labeled sample of the 24 kDa TehA fragment, demonstrating the relatively narrow chemical shift dispersion and peak overlap in the middle of the spectrum.

To assign the backbone of the protein, we have measured HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HNCO, and HN(CA)CO spectra. In addition, we used a ^{2}H/^{15}N-labeled sample to measure a ^{15}N-edited NOESY−TROSY and a [^{15}N,^{1}H] HMQC−NOESY−

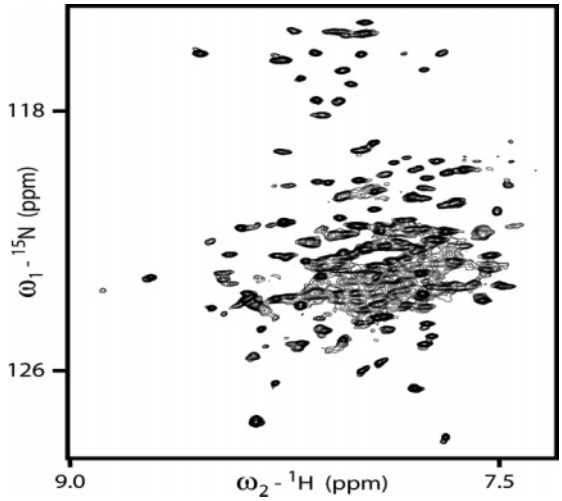


Figure 1. [^{15}N,^{1}H] TROSY spectrum of the 24 kDa fragment of TehA. The concentration of the protein was 0.6 mM, dissolved in 25 mM potassium phosphate buffer (pH 6) containing 5% LMPG (1-myristoyl-2-hydroxy-*sn*-glycero-3-[phosphor-*rac*-(1-glycerol)]). The spectrum was measured at 40 °C on a 800 MHz NMR instrument with 4 scans per FID and 300 increments in the indirect dimension.

TROSY experiment. On the basis of the combination of these experiments, we were able to assign 55% of the protein's backbone unambiguously. However, severe overlap, as well as the absence of some peaks, prevented us from obtaining more assignments. To close these gaps, we have expressed several amino acid type selective-labeled samples. However, despite the labeling of 10 different amino acid types (W, A, V, T, S, R, M, L, I, F), only an additional 10% of unambiguous assignments could be obtained. The main reason for the failure of the selective labeling procedure to result in higher assignment yields was that, in many cases, the N- and C-terminal connectivities to an identified amino acid type were not unambiguous, resulting in more than one possible sequence-specific assignment. To solve this problem, we decided to use a specific labeling procedure based on the simultaneous labeling of certain amino acid types with ^{15}N and other amino acid types with ^{13}C on the backbone carbonyls, which has been used in previous applications for site-specific labeling.[7−9] By measuring two-dimensional versions of an HNCO experiment, it is possible to select only those ^{15}N-labeled amino acids that are N-terminally preceded by a ^{13}C-labeled amino acid type. If that combination occurs only once in the entire protein, that amino acid is site specifically assigned and can be used as an anchor point for further sequence-specific assignments. To optimize this procedure and to minimize the number of samples that have to be produced, we employed a combinatorial approach. As summarized in Table 1, we produced three different samples, each one labeled with two to three different ^{15}N-labeled amino acid types and in addition with two different ^{13}C-carbonyl-labeled amino acid types. By measuring

10.1021/ja0540270 CCC: $30.25

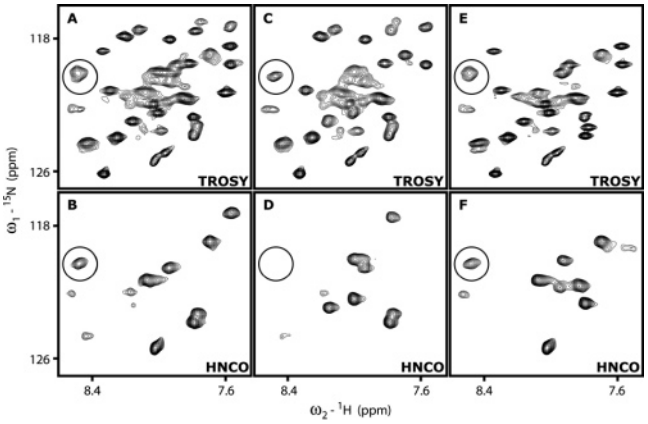


Figure 2. Results of the combinatorial labeling scheme. A, C, and E show TROSY spectra of samples 1−3, and B, D, and F the corresponding HNCO spectra. The circle indicates the resonance position of alanine 48, which in TehA is preceded by leucine 47.

Table 1. Labeling Schemes Used for the Combinatorial Labeling

amino acid type	sample 1	sample 2	sample 3
alanine	^{15}N	^{15}N	^{15}N
phenylalanine	^{15}N	^{15}N	
isoleucine	^{15}N		^{15}N
serine	^{13}C	^{13}C	
leucine	^{13}C		^{13}C
glycine		^{13}C	
valine			^{13}C

a [^{15}N,^{1}H] TROSY spectrum as well as a two-dimensional HNCO spectrum for each of the three samples (Figure 2), the sequence-specific assignment for eight new amino acids could be obtained, which served as specific starting points for more residues, bringing the total backbone assignment to 85% (in addition the assignment for 14 amino acids previously assigned could be confirmed). The amino acid types used for this combinatorial specific labeling approach were selected by an algorithm programmed in Matlab (http://www.biophyschem.uni-frankfurt.de/AK_Doetsch/projects/download/combilabel.m). The input for this algorithm are the amino acid sequence and the unassigned sequence stretches. On the basis of this information, the algorithm calculates the optimal combination of ^{15}N- and ^{13}C-labeled amino acids that will provide the most new specific assignments. Of the remaining 15% of backbone resonances that could not be assigned, 10% are not visible even in a two-dimensional TROSY spectrum while 5% are visible, but do not show sequential connectivities. These 5% can, in principle, be assigned with the specific labeling method. This, however, basically requires one sample per assignment.

Reinvestigation of our data showed that the combination of the nonselective triple resonance experiments with the combinatorial specific labeling strategy would have produced the same level of overall backbone assignment (85%) as the combination of all three assignment strategies (nonselective triple resonance experiments, amino acid type selective labeling, and combinatorial specific labeling). We, therefore, propose as the most straightforward strategy for the backbone assignment of membrane proteins the combination of nonselective triple resonance experiments and a combinatorial specific labeling protocol based on the production of proteins with an in vitro transcription/translation system. Recently, specific labeling in combination with triple resonance experiments has been used to accelerate the assignment process,[8] and a partial assignment procedure based entirely on the use of a combinatorial specific labeling scheme has been proposed for the selective assignment of certain amino acids in soluble proteins.[9] While a pure combinatorial approach is very efficient and useful for applications that only require the assignments of the amide proton and nitrogen frequencies (such as binding assays), full structure determinations are increasingly based on the use of ^{13}C backbone chemical shifts as structural parameters. In particular, for membrane proteins, the use of ^{13}Cα and ^{13}Cβ chemical shifts as indicators of the secondary structure is very important.[3,5] These chemical shifts are automatically provided by the nonselective triple resonance experiments. Furthermore, typically only 40−50% of all amino acids of a protein are part of a unique amino acid pair within the sequence and can, therefore, be unambiguously assigned solely based on combinatorial specific labeling pattern.[9] In the case of TehA, 95 pairs were unique corresponding to 43.4%; 35 pairs occurred twice, 9 pairs three times, and 6 pairs more than three times. By first assigning as many resonances as possible with the nonspecific triple resonance experiments and using an optimization procedure to pick from the remaining sequence stretches those amino acid combinations with the highest number of unique pairs, this problem of multiple possible assignments can be almost completely avoided.

This combinatorial labeling scheme relies on the use of a cell-free transcription/translation system for the production of the protein samples. In principle, amino acid type selective labeling is also possible in auxotrophic bacteria;[10] however, the available strains are only auxotrophic for certain types of amino acids, thus limiting the potential labeling combinations. In contrast, cross labeling in in vitro transcription/translation reactions is negligible.[11] Furthermore, the small amounts of labeled amino acids that are used make selective labeling in this system far less expensive than in cellular systems. Finally, producing NMR samples with an in vitro transcription/translation system is very fast. A typical NMR sample can be produced in less than 24 h since no complicated cell disruption and purification schemes are involved. In fact, since the produced protein is the only labeled macromolecule in the reaction mixture, NMR spectra can, in principle, be measured without any chromatographic purification.[11] We, therefore, believe that the combination of cell-free transcription/translation with standard NMR triple resonance experiments and combinatorial labeling schemes will provide a very efficient avenue toward the backbone assignment of membrane proteins.

Acknowledgment. Financial support was provided by the DFG (SFB 628).

References

(1) Klammt, C.; Löhr, F.; Schäfer, B.; Haase, W.; Dötsch, V.; Rüterjans, H.; Glaubitz, C.; Bernhard, F. *Eur. J. Biochem.* **2004**, *271*, 568−580.
(2) Ishihara, G.; Goto, M.; Saeki, M.; Ito, K.; Hori, T.; Kigawa, T.; Shirouzu, M.; Yokoyama, S. *Protein Expr. Purif.* **2005**, *41*, 27−37.
(3) Oxenoid, K.; Kim, H. J.; Jacob, J.; Sönnichsen, F. D.; Sanders, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 5048−5049.
(4) Tian, C.; Breyer, R. M.; Kim, H. J.; Karra, M. D.; Friedman, D. B.; Karpay, A.; Sanders, C. R. *J. Am. Chem. Soc.* **2005**, *127*, 8010−8011.
(5) Roosild, T. P.; Greenwald, J.; Vega, M.; Castronov, S.; Riek, R.; Choe, S. *Science* **2005**, *307*, 1317−1321.
(6) Turner, R. J.; Taylor, D. E.; Weiner, J. H. *Antimicrob. Agents Chemother.* **1997**, *41*, 440−444.
(7) Weigelt, J.; van Dongen, M.; Uppenberg, J.; Schultz, J.; Wikstrom, M. *J. Am. Chem. Soc.* **2002**, *124*, 2446−2447.
(8) Shi, J.; Pelton, J. G.; Cho, H. S.; Wemmer, D. E. *J. Biomol. NMR* **2004**, *28*, 235−247.
(9) Parker, M. J.; Aulton-Jones, M.; Hounslow, A. M.; Craven, C. J. *J. Am. Chem. Soc.* **2004**, *126*, 5020−5021.
(10) Ou, H. D.; Lai, H. C.; Serber, Z.; Dötsch, V. *J. Biomol. NMR* **2001**, *21*, 269−273.
(11) Ozawa, K.; Headlam, M. J.; Schaeffer, P. M.; Henderson, B. R.; Dixon, N. E.; Otting, G. *Eur. J. Biochem.* **2004**, *271*, 4084−4093.

JA0540270