J Biomol NMR (2010) 46:33–43
DOI 10.1007/s10858-009-9364-5

PERSPECTIVE

Cell-free expression and stable isotope labelling strategies for membrane proteins

Solmaz Sobhanifar · Sina Reckel · Friederike Junge · Daniel Schwarz · Lei Kai · Mikhail Karbyshev · Frank Löhr · Frank Bernhard · Volker Dötsch

Received: 18 May 2009 / Accepted: 16 July 2009 / Published online: 13 August 2009

Abstract Membrane proteins are highly underrepresented in the structural data-base and remain one of the most challenging targets for functional and structural elucidation. Their roles in transport and cellular communication, furthermore, often make over-expression toxic to their host, and their hydrophobicity and structural complexity make isolation and reconstitution a complicated task, especially in cases where proteins are targeted to inclusion bodies. The development of cell-free expression systems provides a very interesting alternative to cell-based systems, since it circumvents many problems such as toxicity or necessity for the transportation of the synthesized protein to the membrane, and constitutes the only system that allows for direct production of membrane proteins in membrane-mimetic environments which may be suitable for liquid state NMR measurements. The unique advantages of the cell-free expression system, including strong expression yields as well as the direct incorporation of almost any combination of amino acids with very little metabolic scrambling, has allowed for the development of a wide-array of isotope labelling techniques which facilitate structural investigations of proteins whose spectral congestion and broad line-widths may have earlier rendered them beyond the scope of NMR. Here we explore various labelling strategies in conjunction with cell-free developments, with a particular focus on α-helical transmembrane proteins which benefit most from such methods.

S. Sobhanifar · S. Reckel · F. Junge · D. Schwarz · L. Kai · M. Karbyshev · F. Löhr · F. Bernhard · V. Dötsch (✉)
Institute of Biophysical Chemistry, Centre for Biomolecular Magnetic Resonance (BMRZ), Goethe University, Frankfurt/Main, Germany
e-mail: vdoetsch@em.uni-frankfurt.de

S. Sobhanifar · S. Reckel · F. Junge · D. Schwarz · L. Kai · M. Karbyshev · F. Löhr · F. Bernhard · V. Dötsch
Cluster of Excellence Frankfurt (Macromolecular Complexes), Goethe University, Frankfurt/Main, Germany

Introduction

Membrane proteins typically adopt one of two types of structural motifs in their membrane bound domains, namely rigid pore β-barrels as is typical of gram-negative bacteria, chloroplasts, and mitochondria, or single or bundled α-helical fragments which are by far more common and constitute roughly one-third of a typical genome. This latter group is of great relevance to the medical field in form of potential therapeutic drug targets, but the progress in structural aspects has been slow, particularly with regard to ones containing multiple-transmembrane helices such as the GPCRs. Lack of progress has been largely due to the fact that such proteins tend to give low expression yields in standard cell-based systems, and in many cases are difficult to isolate and reconstitute in active form. Even when a viable sample is obtained, liquid state NMR investigation, up until recently, has in most cases been unfeasible; typical membrane proteins in a micelle environment tend to be large and thus prone to signal broadening and crowding. Advancements in the field of NMR such as the use of higher-field magnets, cryogenic probes, deuteration, as well as the introduction of transverse relaxation optimized spectroscopy (TROSY) (Pervushin et al. 1997) which significantly increases both spectral resolution and sensitivity, have greatly extended the size-boundary of molecules that can be studied by NMR. In fact, solution structures of several membrane proteins such as the β-barrel PapG (Hwang et al. 2002), OmpX (Fernandez et al. 2004),

OmpA (Arora et al. 2001; Johansson et al. 2007), OmpG (Liang and Tamm 2007), and VDAC-1 (Bayrhuber et al. 2008; Hiller et al. 2008), as well as those of the α-helical M2 proton channel (Schnell and Chou 2008), DsbB (Zhou et al. 2008), DAGK (Van Horn et al. 2009), and membrane-associated Mystic (Roosild et al. 2005) have recently been solved. Nevertheless, the numbers remain low, especially with regard to α-helical proteins. This is because, despite sensitive detection methods, the shear size of the protein-micelle/bicelle complex still pushes the boundaries and causes considerable line-broadening, which is further aggravated by the narrow-range chemical shifts for which α-helical membrane proteins are especially notorious. Consequently, spectra of uniformly labelled samples are often severely overlapped and require selective labelling for improved resolution. Cell-based expression presents limited options for selective labelling due to the unavailability of certain auxotrophic mutants or mutant combinations, as well as significant amino acid scrambling and dilution (McIntosh and Dahlquist 1990); this often confines labelling to very few amino acid types, limiting the amount of information that may be obtained. Cell-free expression has been routinely applied for production of soluble protein NMR samples (Kigawa et al. 1995; Ozawa et al. 2005, 2006; Vinarov et al. 2006) and has become a standard technique in part due to its widespread use in structural genomics consortia. Recently, cell-free systems have in addition been used successfully for the expression of membrane proteins, circumventing critical obstacles inherent to cell-based systems such as toxicity, as well as the sequestering of protein within inclusion bodies (Klammt et al. 2004, 2005, 2007; Schwarz et al. 2007; Keller et al. 2008). The introduction of cell-free expression not only allows for the synthesis of difficult-to-express proteins in sufficient quantities for structural studies, but in addition, allows for selective labelling with almost any amino acid type (with limited scrambling) according to various selective and combinatorial schemes tailored to NMR analysis. The combination of improved expression and great flexibility for various isotopic labelling schemes makes cell-free expression of membrane proteins a very interesting tool for structural studies of this important protein class, particularly with respect to multi-pass α-helical species.

Overview of cell-free protein expression

Due to low cost and high productivity, proteins are routinely produced using bacteria-based expression systems. However, in cases where yields are limiting, as with many membrane proteins, or where specific labelling strategies are necessary, the cell-free expression system holds several benefits. Firstly, the open nature of the system allows for the addition of agents that support expression and stabilize the synthesized protein, such as RNAse and protease inhibitors or ligands, the latter of which is particularly advantageous for proteins sensitive to proteolytic degradation. The system can also be modified to include non-natural amino acids (Noren et al. 1989; Muranaka et al. 2007; Goerke and Swartz 2009) or to exchange metal ions with those normally not incorporated during cellular expression, such as the paramagnetic lanthanides, whose integration can provide important parameters for NMR structural refinement (Lee and Sykes 1980; Veglia and Opella 2000). Several cell-free expression systems are available based on cell extracts of different origin, including *Escherichia coli*, wheat germ, and rabbit reticulocytes. While the latter may be more suitable for proteins with specific post-translational requirements, the two other systems are more commonly used due to their higher efficiencies. For membrane proteins, cell-free expression provides the added benefit of allowing direct addition of certain detergents or other membrane mimetics such as liposomes and bicelles to the reaction mixture. Cell-free systems are consequently the only expression systems capable of direct expression of membrane proteins into micelles and other hydrophobic environments which may be suited to liquid state NMR spectroscopy. High expression yields can also be obtained in the absence of the above mentioned environments. In this case, the expressed membrane protein forms a precipitate which appears in nature very different from the inclusion bodies which often form during expression in bacteria, in that it readily solubilizes in detergents suitable for NMR investigation (Klammt et al. 2004). The three different modes of cell-free membrane protein expression—as precipitate (P-CF), in presence of detergent (D-CF), or in presence of lipids (L-CF)—are depicted in Fig. 1. Of these three modes, expression in P-CF generally provides the highest yields, although expression in D-CF, in a case-dependent manner and under certain detergent conditions, can produce comparable protein levels (Berrier et al. 2004; Elbaz et al. 2004; Ishihara et al. 2005; Klammt et al. 2005). While these two methods may provide reasonable amounts of protein, functional protein with the correct conformation may not always be achievable in detergent. The expression of fully functional protein directly into liposomes has been achieved using the L-CF mode (Kalmbach et al. 2007; Wuu and Swartz 2008; Hovijitra et al. 2009), although thus far, yields have tended to be insufficient for structural analysis in this mode.

A further advantage of cell free systems is that milligram amounts of protein sufficient for structural studies may be obtained within small reaction volumes (1–3 ml), and given that target proteins are the only ones synthesized, isotope-labelled amino acids are used efficiently and

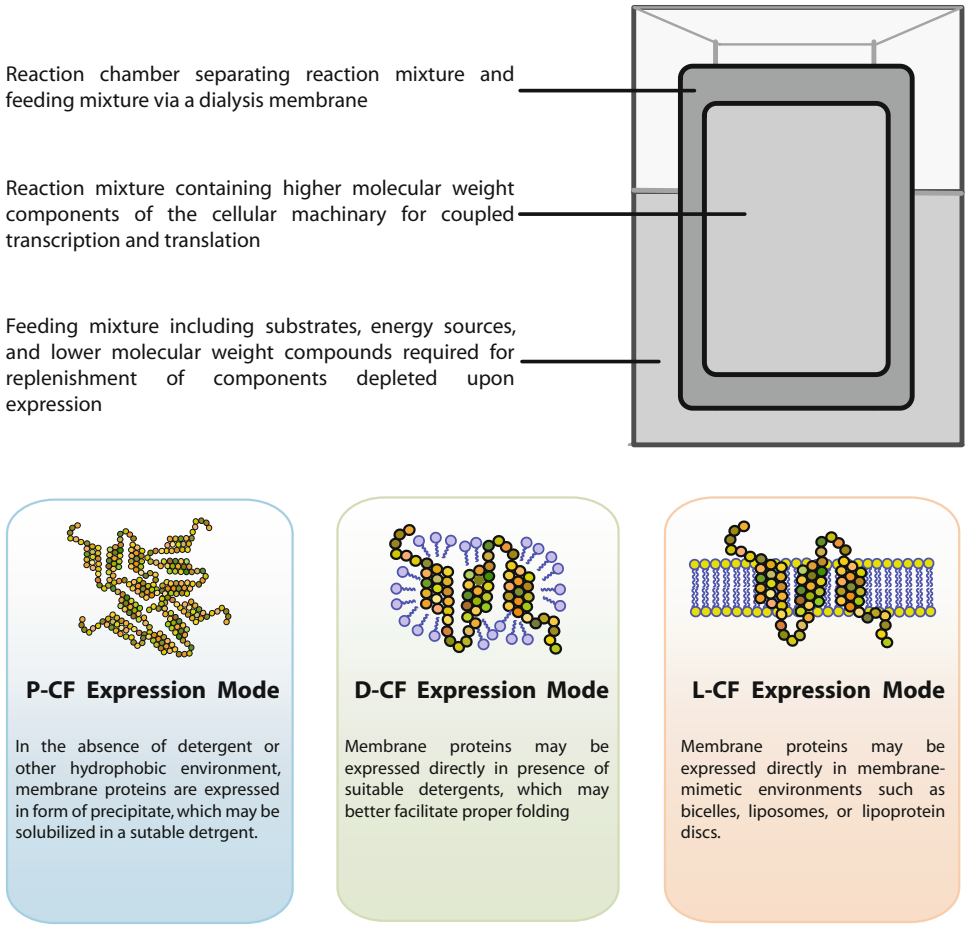


Fig. 1 Schematic of a typical continuous-exchange cell-free expression system. Proteins may be expressed in three possible modes: in absence of detergent in form of precipitate (P-CF); in presence of detergent (D-CF); in presence of bicelles, liposomes, or lipoproteins (L-CF). The reaction mixture, containing the cellular machinery, is separated from the feeding mixture, containing substrates and the energy source, by a semi-permeable membrane (10-14MWCO)

purification may not be necessary since only the expressed protein is detectable in heteronuclear NMR experiments (Ozawa et al. 2006). Most importantly, since metabolic enzymes are not newly expressed throughout the reaction, isotope scrambling is kept at a minimum. The first system configuration to produce enough protein for structural investigations was termed "continuous-flow" (CFCF), where continuous removal of reaction products including the synthesized polypeptides, as well as a continuous supply of energy source and substrates, are relied upon for synthesis (Spirin et al. 1988; Kigawa and Yokoyama 1991). This technique was later improved upon in the form of the "continuous exchange" cell-free (CECF) system, where the high molecular weight components of the cellular machinery are held in a smaller and separate compartment within a dialysis membrane, surrounded by a large feeding mixture containing the energy source and precursors (Kim and Choi 1996; Kigawa et al. 1999). The advantage of such a system is that a relatively large feeding mixture can be used to supply a smaller reaction volume (typically between 10:1 and 30:1 feeding volume to reaction volume ratio), such that less extract and enzymes are required per reaction, and breakdown products such as pyrophosphate diffuse freely through the semi-permeable membrane so that their concentrations are kept low enough not to significantly inhibit the reaction.

One of the most important questions concerning expressed membrane proteins is whether these proteins adopt a functional fold. This problem is not specific to cell-free expression but affects membrane protein preparations regardless of the expression system, since even the extraction of a membrane protein from its natural membrane environment with detergent can influence the conformation of the protein. To answer the question of whether cell-free expressed membrane proteins can at least in principle be active, we have tested several cell-free expressed membrane proteins for their functionality. In the case of cell-free expressed EmrE, we as well as others have demonstrated that this protein, when incorporated into liposomes, is capable of transporting organic compounds such as ethidium bromide in a proton-gradient dependent manner, and that it has the same structure and binds compounds in the same way as protein expressed in *E. coli* cells (Klammt et al. 2004; Chen et al. 2007). We have also shown that the β-barrel protein TSX can form pores in membranes, which can be investigated by measuring the increase in conductivity.

The pore-forming capability of TSX is, however, critically dependent on the type of detergent used for resolubilization: With Triton X-100 a high conductivity could be measured, whereas after resolubilization with LMPG, the conductivity reached only background levels (Klammt et al. 2005). Functionality of cell-free expressed membrane proteins is not restricted to these rather small and simple proteins. Recently, we could show that the transport characteristics of the cell-free expressed cation-exporter rOCT-1 from rat are virtually identical to those determined for the transporter expressed in insect cells (Keller et al. 2008). This protein with its 12 transmembrane helices, a molecular weight of 60 kDa, and assumed dimerization via a disulfide bond in the first loop region, is rather complex. These results demonstrate that cell-free expressed proteins can in principle be functional and active, if the right detergent/lipid environment can be identified. It is therefore important to have a set of criteria that help identify conditions which give rise to functionally folded protein before engaging with often costly and time-consuming structure-targeted experiments. Target-specific activity and interaction assays would of course be ideal for verifying the functional fold of a protein. In some cases however, this may not be possible; for example, certain detergents may interfere with binding by shielding the hydrophobic interaction interfaces of proteins occupying separate micelles, or the protein investigated may be part of a larger complex which as a whole is required for activity. In such cases, one should examine other parameters related to folding and sample quality. These include looking for mono-dispersion via gel filtration, dynamic or static light scattering and single-particle analysis by electron microscopy, analyzing the multi-meric state of the protein via native gel, ultracentrifugation or mass spectrometry (LILBID, Morgner et al. 2008), and investigating topology and tertiary fold via far and near-UV circular dichroism spectroscopy. Once there is evidence of a mono-disperse, well-behaved and well-folded protein specimen, one should proceed with HSQC measurements to examine the dispersion of the resonance signals and to determine whether and which labelling strategies may be beneficial.

Cell-free approaches for stable isotope labelling

Due to its high versatility, the cell-free approach has become popular for the production of proteins which are difficult to express in cell-based systems, or have requirements which can be most readily facilitated by the addition of agents directly to the reaction mixture. Descriptions of cell-free systems and their various applications, as well as published protocols for cell-free protein synthesis are hence widely available (Jermutus et al. 1998; Kim and Swartz 1999, 2000; Guignard et al. 2002; Hirao et al. 2002; Hwang et al. 2002; Jewett and Swartz 2004a, b, c; Swartz et al. 2004; Schwarz et al. 2007; Hino et al. 2008; Kopeina et al. 2008). Here the focus is placed on cell-free expression in the context of labelling applications; several modifications have been made to standard protocols to address labelling efficiency, in way of removal of endogenous amino-acids, elimination of minimal but still present scrambling, and production of fully deuterated protein samples.

When expressing labelled proteins, extra care must be taken in removal of endogenous templates still present within the ribosomes. Most protocols employ a step where additional amino acids are added and translation is allowed to reach completion so that the mRNA may be released. This approach may, however, be problematic for isotope labelling due to the addition of further unlabelled amino acids to those already present. One suggested method is to apply the lysate to a gel-filtration column to remove the low molecular weight amino acids, which is reported to increase labelling efficiency from $\sim$90 to $\sim$96% (Torizawa et al. 2004). An alternative is the addition of NaCl (to 400 mM) to the lysate (rather than amino acid addition), followed by incubation at 42°C for $\sim$45 min, which causes the release of the mRNA, followed by degradation and overnight dialysis to remove the added NaCl and any amino acids still present (Klammt et al. 2004).

In most cell-free expression systems, potassium L-Glu is used as the potassium ion source which is necessary for efficient translation (Kigawa et al. 1995; Jewett and Swartz 2004a, b, c). This system is nevertheless clearly unsuitable for stable isotope labelling where Glu is involved, whether in uniform or selective schemes, due to isotope dilution. An alternative has been the use of potassium D-Glu, whose productivity is reported to equal that of the potassium L-Glu system with no compromise in labelling efficiency (Matsuda et al. 2007). Recently, potassium *N*-acetyl-L-Glu and potassium glutarate buffers have also been shown capable of replacing the potassium L-Glu in cell-free reactions with comparable yields, although the use of the L-Glu buffer is still recommended in cases where glutamate is not labelled, since L-Glu appears to better suppress amino-transferase activity (Jia et al. 2009). It is also common to use the acetate salt as a potassium source for stable isotope labelling (Ozawa et al. 2004), although the productivity of this system is reported to be half that of the L-Glu system.

Escherichia coli extract-based cell-free protein preparations, especially those produced from strains such as the A19, generally suffer little from amino acid scrambling, except in cases of certain amino acids such as Glu and Gln (Ozawa et al. 2004). Although uniform labelled samples are not affected from scrambling, in selective labelling protocols, such amino acid types should generally be

avoided. In the wheat-germ system, interconversion between Glu and Asp, Glu and Gln, and Ala and Gln are particularly problematic, but may be overcome by the addition of transaminase and Glu synthase inhibitors (Morita et al. 2004).

As most membrane proteins are large polytopic structures, it is often difficult to interpret spectra due to congestion and line-broadening resulting from transverse relaxation. In these cases, deuteration is often necessary to temper relaxation in order to obtain useful spectra (Clore and Gronenborn 1998; Gardner and Kay 1998). Simply using perdeuterated amino acids, however, still leads to the introduction of a certain degree of protonation upon back-exchange at H^{α} and, to a lesser extent, H^{β} positions due to exchange with bulk water (Etezady-Esfarjani et al. 2007). The production of proteins with high deuteration levels, therefore, requires expression in D_2O. Given that in a standard CECF expression system, the concentrated cell-free or S-30 extract (Nirenberg 1963; Liu et al. 2005), designated according to the sedimentation characteristics of its components, accounts for approximately 30% of the reaction mixture, it would be useful to work with an extract prepared in D_2O, such as would be the case with the other cell-free reagents. Growth of *E. coli* cells in D_2O is nevertheless a rather expensive option in terms of shear volume; 10 l of bacterial culture typically yield ~60–100 ml of S-30 extract. Etezady-Esfarjani et al. (2007) have proposed a protocol for the production of a deuterated S-30 extract capable of producing protein with up to 95% deuteration in non-labile proton positions, with comparable expression levels to the H_2O-based counterpart. This simple but effective protocol involves buffer exchange of a normal H_2O-based S-30 extract against an equal volume of D_2O S-30 buffer, which is repeated six times using a centrifugal device (10 kDa MWCO). An alternative may be the use of the "Pure Translation System", developed by Shimizu et al. (2001, 2005) which consists entirely of only those components required for protein translation, and does not contain metabolic enzyme activities which are likely responsible for back-protonation (Kainosho, personal communication); nevertheless, as of yet, the high cost and low reaction yield of the system may somewhat reserve its use.

Labelling stratagies for backbone assignment

For the majority of proteins, assignment may be easily obtained by measuring standard 3D spectra with a uniformly $^{15}N/^{13}C$-labelled sample. For α-helical membrane proteins, however, the chemical shift information is often dispersed within little more than 1 ppm on the amide proton scale. A uniformly labelled sample thus can often be prone to severe overlap even in the context of a 3D experiment and in certain cases may not be sufficient for high degrees of assignment. The large overlap in resonance has two origins, firstly the broadening resulting from the shear size of the protein-micelle/bicelle complex, and secondly, the similar chemical environment of limited types of hydrophobic residues which are abundant and tend to cluster in transmembrane α-helical regions (Fig. 2). The

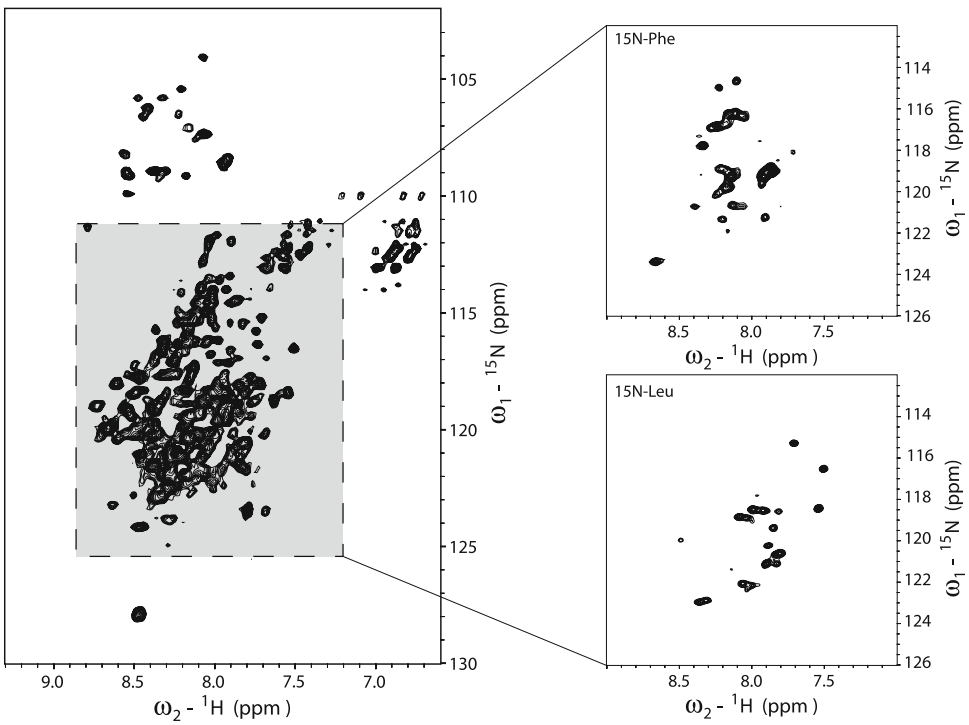


Fig. 2 [^{15}N,^{1}H]-TROSY-HSQC spectra of uniformly (*left*) and selectively (*right*) labelled FLAP (5-lipoxygenase activating protein). The narrow range of the chemical shifts and the spectral congestion are typical of α-helical membrane proteins above 20 kD, necessitating the use of selective labels to identify individual residues

first problem may be to some extent alleviated by adjusting pH and temperature, as well as by the already mentioned high-level deuteration methods, whereas the second may be exploited using selective labelling techniques, such as transmembrane segment enhanced (TMS) labelling (Reckel et al. 2008): Transmembrane helices are predominantly composed of six amino acid types, Ala, Phe, Gly, Ile, Leu, and Val, all of which together account for some 60% of the amino acids in the transmembrane sections of a typical α-helical membrane protein (Senes et al. 2000). By the selective $^{13}C/^{15}N$-labelling of only these six amino acid types, one can therefore significantly reduce spectral overlap, while retaining long consecutive stretches of these labelled amino acids to still enable the use of standard 3D NMR measurements such as the HNCA and HN(CO)CA for obtaining the backbone assignment (Fig. 3). Although there is remaining overlap between the C^{α} chemical shift

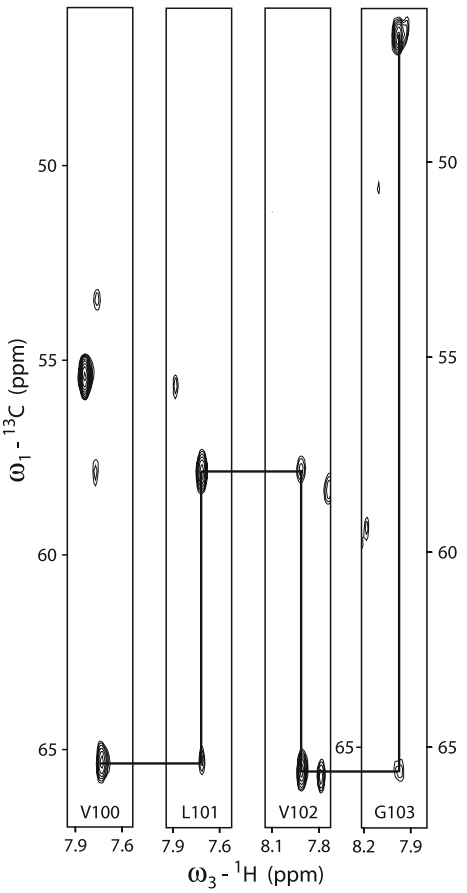


Fig. 3 Sequential connectivities identified in the TMS-labelled C-terminal fragment of Presenilin. Strips from selected planes of a 3D HNCA spectrum were chosen to demonstrate the sequential connectivity of the residues. Although only six out of the twenty amino acids were selectively labelled (alanine, phenylalanine, glycine, isoleucine, leucine, valine), it is possible to obtain many connectivities due to the predominance of the given amino acid types within α-helical transmembrane regions

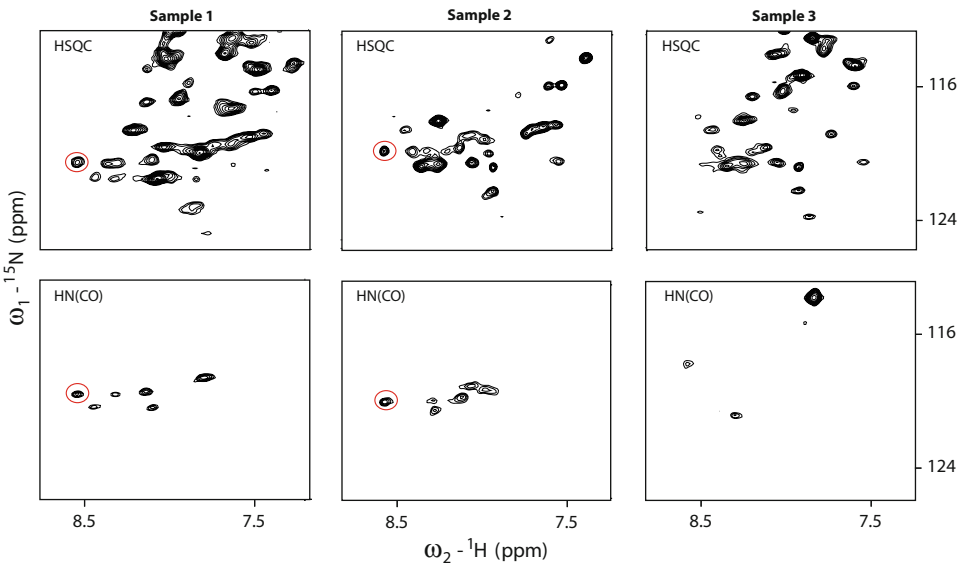
values of Ile and Val, and among Ala, Phe and Leu, these may be resolved with the HNCACB, or in the case of larger proteins, with the use of one or two amino acid type-selective labelled samples. The TMS labelling, nevertheless, may be modified to include or substitute other abundant residues within α-helical regions of a desired protein, or in a reverse labelling approach to predominantly select for the (hydrophilic) residues in the loop regions of the membrane protein. Such an approach, however, is only then economically possible when a significant clustering of only a few residue types occurs, a condition more often met in the hydrophobic transmembrane regions than in the hydrophilic loops. After assignment of the backbone resonances of TMS-labelled stretches, conventional 3D spectra, such as HNCA, and HN(CO)CA, can be measured on uniformly $^{15}N/^{13}C$-labeled samples to assign the connecting regions.

In our experience, the significant chemical shift overlap of α-helical membrane proteins can often prevent the completion of backbone assignment based only on uniform- and TMS-labelled samples, which may be case-dependent [α-helical proteins such as DsbB (Zhou et al. 2008) and DAGK (Van Horn et al. 2009) have been assigned without the aid of selective labelling]. In such instances, unassigned stretches usually remain, since very often the high chemical shift degeneracy of the residues makes identification of the correct sequential connectivity impossible. Although with the cell-free system it is possible to separately label all 19 non-proline residues, one would benefit more from a combinatorial selective labelling scheme, which would require a smaller subset of samples without compromising the amount of information obtained. In this regard, Wu et al. (2006) have proposed a protocol using a minimum of five samples as opposed to 19, where combinations of amino acids are distributed within different samples based on their relative frequencies; those most abundant are included in only one sample, whereas those least abundant are included in up to three. [^{15}N,^{1}H] HSQC spectra are then measured for all samples, and the pattern of presence and absence of a particular peak can be used to determine its identity. Furthermore, individual samples would contain roughly one-third of the peaks present in the spectra of the corresponding uniformly labelled sample, which would effectively decrease the level of overlap.

Although this technique is very useful in providing amino acid-type specific assignments, α-helical proteins tend still to show congestion in 3D spectral planes, such that sequential assignment may remain challenging even when individual amino acid types may be identified. Accordingly, a complimentary technique that may be used is “dual combinatorial selective labelling”, which provides sequence specific assignment by identifying amino acid

pairs with a single occurrence within the protein sequence. These can be uniquely assigned and may act as anchor points for further assignments. The strategy is based on a double labelling technique (Kainosho and Tsuji 1982; Yabuki et al. 1998; Weigelt et al. 2002; Shi et al. 2004), where selected amino acids are ^{15}N labelled and others ^{13}CO labelled. HSQC and proton/nitrogen projected 2D HN(CO) spectra are then measured to obtain the chemical shifts of all amino acids of a certain type from the HSQC, as well as to identify specific amino acids in the HN(CO) spectrum. The simplest implementation of this technique requires a separate sample for each prospective amino acid pair. Others take advantage of a combinatorial method, where fewer samples are required to obtain comparable information. In principle such methods can be used to obtain a large percentage of the backbone assignment—all unique amino acid combinations which typically constitute ~50% of a protein—even without classical 3D NMR experiments. The method proposed by Parker et al. (2004) involves five samples, the first of which contains sixteen ^{13}C/^{15}N doubly labelled amino acids, and the rest a combination of the sixteen, where some are 50% ^{15}N labelled, diluted with their non-labelled counterparts, and others 100% ^{13}C/^{15}N labelled. HSQC and 2D HNCO spectra are subsequently measured and analyzed to obtain sequence specific assignments based on the presence, absence, and intensity of the peaks in the spectra. While this technique provides many advantages for soluble proteins, the significantly higher signal overlap and broader line width require modifications for use with membrane proteins. A simpler method developed by Trbovic et al. (2005) uses the same combinatorial concept but requires labelling of fewer amino acid types. In addition, these amino acids need only be 100% ^{15}N or ^{13}C single-labelled. The combination of signals appearing in the HSQC or HNCO spectra, rather than differences in intensities, is used to obtain amino acid-type specific assignments (Fig. 4). Although this method inherently works with less information, it has some obvious benefits for membrane proteins, such as the fewer number of peaks in the HSQC. The cost is furthermore reduced since fewer labelled amino acids are required. In addition, relying only on the presence or absence of peaks rather than intensity makes this method robust which is particularly important for membrane proteins with large line widths. Although this technique does not allow significant backbone determination on its own, combined with standard 3D experiments such as the HNCA, it provides important sequence specific anchor points which greatly facilitate assignment.

Based on our experience, a basic backbone assignment strategy may be to use first TMS-labelling to assign stretches in the transmembrane regions followed by analysis of



Amino acid type	Sample 1	Sample 2	Sample 3
Phenylalanine	^{15}N	^{15}N	^{15}N
Threonine	^{15}N		^{15}N
Valine	^{15}N	^{15}N	
Glycine	^{13}C	^{13}C	
Leucine	^{13}C		^{13}C
Isoleucine		^{13}C	^{13}C

Fig. 4 2D [^{15}N,^{1}H]-TROSY-HSQC and 2D [^{15}N,^{1}H]-TROSY-HN(CO) spectra of the selective combinatorial labelled bacterial cysteine exporter YfiK. Selected amino acids were ^{15}N or ^{13}C labelled and distributed over three different samples in such a way that each produced pattern would be indicative of a unique amino acid pair combination. TROSY-HSQC and 2D versions of HNCO experiments were then measured and analyzed. In the above example, the encircled peak appears in the HSQC and HNCO spectra of samples 1 and 2, which, based on the labelling pattern, is designated as a valine preceded by a glycine

spectra obtained with uniformly double-labelled samples to connect the TMS-assigned stretches. Remaining gaps may be targeted for selective combinatorial labelling for the identification of unique amino acid pairs to be used as anchor points for further assignments.

Cell-free protein synthesis also provides the basis for an elegant alternative to this double labelling strategy. The method pioneered by the laboratory of Peter G. Schultz uses reprogrammed suppressor codons that allow for the site specific labelling of exactly one amino acid (Noren et al. 1989; Bain et al. 1991; Ellman et al. 1992; Sonar et al. 1994; Liu et al. 1995). Such approaches require the insertion of an (amber) stop codon in place of the to-be-labelled amino acid codon, along with the aminoacetylation of the corresponding modified tRNA with the labelled amino acid. This innovative approach utilizes the cell-free expression system to produce site-specifically stable isotope labelled protein, resulting in NMR samples that show exactly one peak in their HSQC spectra. While this strategy offers many opportunities, it has so far not found wide spread use in the NMR community (Yabuki et al. 1998).

The above mentioned methods address the assignment of backbone resonances. To obtain high resolution structures, however, one in addition requires assignment of side-chains, which in many membrane proteins display even greater degeneracy. As mentioned before, deuteration is often necessary to reduce line-broadening, but this also removes all carbon bound protons and accompanying side-chain information. A very promising alternative may be the use of stereo-array isotope labelling (SAIL) (Kainosho et al. 2006), where certain proton positions are strategically deuterated in a stereospecific manner in such a way to reduce the number of peaks present, while retaining critical side-chain information. Although this method has not yet been applied to membrane proteins, the above authors have solved the structure of the relatively large (41 kDa) maltodextrin-binding protein (displaying severe overlap in side-chain resonances) using SAIL. As such, this technique may be potentially useful in the case of membrane proteins to obtain long-distance NOE restraints that depend heavily on side-chain information, which currently are very difficult to come by.

Outlook

Cell-free protein expression presents a powerful method for high-yield production of near 100% isotope labelled proteins in virtually any combination for NMR spectroscopy. Given its open nature, the system may be elegantly modified to accommodate the demands of NMR, in way of minimization of isotope scrambling which is instrumental in non-uniform labelling, as well as high-level deuteration necessary for larger species, in particular membrane proteins. This versatility has allowed for the development of several labelling strategies which may be used in combination, directed at facilitating sequence assignment in the context of severe overlap and line-broadening. Furthermore, innovative approaches such as SAIL may in the future lead the way to high resolution structures for membrane proteins, utilizing side-chain information as is routine in the case of most moderately-sized soluble proteins. Given that structural information is most valuable in the context of observed functionality, cell-free expressed proteins may be incorporated in functional form into more native hydrophobic environments such as proteoliposomes (Nozawa et al. 2007; Goren and Fox 2008; Gourdon et al. 2008; Kamonchanok et al. 2008; Keller et al. 2008), which may then be analyzed with solid state NMR, or alternatively, into promising native-like systems such as bicelles (Sanders and Schwonek 1992; Vold et al. 1997; Prosser et al. 2006, 2007) and nano-discs (Leitz et al. 2006; Whorton et al. 2007; Bayrhuber et al. 2008; Cappuccio et al. 2008), which may in the near future present lucrative opportunities for the study of functional membrane proteins using liquid state NMR spectroscopy.

References

Arora A, Abildgaard F, Bushweller JH, Tamm LK (2001) Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. Nat Struct Biol 8:334–338

Bain JD, Diala ES, Glabe CG, Wacker DA, Lyttle MH, Dix TA, Chamberlin AR (1991) Site-specific incorporation of nonnatural residues during in vitro protein biosynthesis with semisynthetic aminoacyl-tRNAs. Biochemistry 30:5411–5421

Bayrhuber M, Meins T, Habeck M, Becker S, Giller K, Villinger S, Vonrhein C, Griesinger C, Zweckstetter M, Zeth K (2008) Structure of the human voltage-dependent anion channel. Proc Natl Acad Sci USA 105:15370–15375

Berrier C, Park KH, Abes S, Bibonne A, Betton JM, Ghazi A (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

Cappuccio JA, Blanchette CD, Sulchek TA, Arroyo ES, Kralj JM, Hinz AK, Kuhn EA, Chromy BA, Segelke BW, Rothschild KJ, Fletcher JE, Katzen F, Peterson TC, Kudlicki WA, Bench G, Hoeprich PD, Coleman MA (2008) Cell-free co-expression of functional membrane proteins and apolipoprotein, forming soluble nanolipoprotein particles. Mol Cell Proteomics 7:2246–2253

Chen YJ, Pornillos O, Lieu S, Ma C, Chen AP, Chang G (2007) X-ray structure of EmrE supports dual topology model. Proc Natl Acad Sci USA 104:18999–19004

Clore GM, Gronenborn AM (1998) NMR structure determination of proteins and protein complexes larger than 20 kDa. Curr Opin Chem Biol 2:564–570

Elbaz Y, Steiner-Mordoch S, Danieli T, Schuldiner S (2004) In vitro synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. Proc Natl Acad Sci USA 101:1519–1524

Ellman JA, Volkman BF, Mendel D, Schulz PG, Wemmer DE (1992) Site-specific isotopic labeling of proteins for NMR studies. J Am Chem Soc 114:7959–7961

Etezady-Esfarjani T, Hiller S, Villalba C, Wuthrich K (2007) Cell-free protein synthesis of perdeuterated proteins for NMR studies. J Biomol NMR 39:229–238

Fernandez C, Hilty C, Wider G, Guntert P, Wuthrich K (2004) NMR structure of the integral membrane protein OmpX. J Mol Biol 336:1211–1221

Gardner KH, Kay LE (1998) The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. Annu Rev Biophys Biomol Struct 27:357–406

Goerke AR, Swartz JR (2009) High-level cell-free synthesis yields of proteins containing site-specific non-natural amino acids. Biotechnol Bioeng 102:400–416

Goren MA, Fox BG (2008) Wheat germ cell-free translation, purification, and assembly of a functional human stearoyl-CoA desaturase complex. Protein Expr Purif 62:171–178

Gourdon P, Alfredsson A, Pedersen A, Malmerberg E, Nyblom M, Widell M, Berntsson R, Pinhassi J, Braiman M, Hansson O, Bonander N, Karlsson G, Neutze R (2008) Optimized in vitro and in vivo expression of proteorhodopsin: a seven-transmembrane proton pump. Protein Expr Purif 58:103–113

Guignard L, Ozawa K, Pursglove SE, Otting G, Dixon NE (2002) NMR analysis of in vitro-synthesized proteins without purification: a high-throughput approach. FEBS Lett 524:159–162

Hiller S, Garces RG, Malia TJ, Orekhov VY, Colombini M, Wagner G (2008) Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. Science 321:1206–1210

Hino M, Kataoka M, Kajimoto K, Yamamoto T, Kido J, Shinohara Y, Baba Y (2008) Efficiency of cell-free protein synthesis based on a crude cell extract from *Escherichia coli*, wheat germ, and rabbit reticulocytes. J Biotechnol 133:183–189

Hirao I, Ohtsuki T, Fujiwara T, Mitsui T, Yokogawa T, Okuni T, Nakayama H, Takio K, Yabuki T, Kigawa T, Kodama K, Yokogawa T, Nishikawa K, Yokoyama S (2002) An unnatural base pair for incorporating amino acid analogs into proteins. Nat Biotechnol 20:177–182

Hovijitra NT, Wuu JJ, Peaker B, Swartz JR (2009) Cell-free synthesis of functional aquaporin Z in synthetic liposomes. Biotechnol Bioeng 104:40–49

Hwang PM, Choy WY, Lo EI, Chen L, Forman-Kay JD, Raetz CR, Prive GG, Bishop RE, Kay LE (2002) Solution structure and dynamics of the outer membrane enzyme PagP by NMR. Proc Natl Acad Sci USA 99:13560–13565

Ishihara G, Goto M, Saeki M, Ito K, Hori T, Kigawa T, Shirouzu M, Yokoyama S (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. Protein Expr Purif 41:27–37

Jermutus L, Ryabova LA, Pluckthun A (1998) Recent advances in producing and selecting functional proteins by using cell-free translation. Curr Opin Biotechnol 9:534–548

Jewett MC, Swartz JR (2004a) Rapid expression and purification of 100 nmol quantities of active protein using cell-free protein synthesis. Biotechnol Prog 20:102–109

Jewett MC, Swartz JR (2004b) Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. Biotechnol Bioeng 87:465–472

Jewett MC, Swartz JR (2004c) Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. Biotechnol Bioeng 86:19–26

Jia X, Ozawa K, Loscha K, Otting G (2009) Glutarate and N-acetyl-L-glutamate buffers for cell-free synthesis of selectively (15)N-labelled proteins. J Biomol NMR 44:59–67

Johansson MU, Alioth S, Hu K, Walser R, Koebnik R, Pervushin K (2007) A minimal transmembrane beta-barrel platform protein studied by nuclear magnetic resonance. Biochemistry 46: 1128–1140

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

Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Mei Ono A, Guntert P (2006) Optimal isotope labelling for NMR protein structure determinations. Nature 440:52–57

Kalmbach R, Chizhov I, Schumacher MC, Friedrich T, Bamberg E, Engelhard M (2007) Functional cell-free synthesis of a seven helix membrane protein: in situ insertion of bacteriorhodopsin into liposomes. J Mol Biol 371:639–648

Kamonchanok S, Balog CI, van der Does AM, Booth R, de Grip WJ, Deelder AM, Bakker RA, Leurs R, Hensbergen PJ (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

Keller T, Schwarz D, Bernhard F, Dotsch V, Hunte C, Gorboulev V, Koepsell H (2008) Cell free expression and functional reconstitution of eukaryotic drug transporters. Biochemistry 47:4552–4564

Kigawa T, Yokoyama S (1991) A continuous cell-free protein synthesis system for coupled transcription-translation. J Biochem 110:166–168

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

Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T, Yokoyama S (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. FEBS Lett 442: 15–19

Kim DM, Choi CY (1996) A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. Biotechnol Prog 12:645–649

Kim DM, Swartz JR (1999) Prolonging cell-free protein synthesis with a novel ATP regeneration system. Biotechnol Bioeng 66:180–188

Kim DM, Swartz JR (2000) Prolonging cell-free protein synthesis by selective reagent additions. Biotechnol Prog 16:385–390

Klammt C, Lohr F, Schafer B, Haase W, Dotsch V, Ruterjans H, Glaubitz C, Bernhard F (2004) High level cell-free expression and specific labeling of integral membrane proteins. Eur J Biochem 271:568–580

Klammt C, Schwarz D, Fendler K, Haase W, Dotsch V, Bernhard F (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

Klammt C, Schwarz D, Eifler N, Engel A, Piehler J, Haase W, Hahn S, Dotsch V, Bernhard F (2007) Cell-free production of G protein-coupled receptors for functional and structural studies. J Struct Biol 158:482–493

Kopeina GS, Afonina ZA, Gromova KV, Shirokov VA, Vasiliev VD, Spirin AS (2008) Step-wise formation of eukaryotic double-row polyribosomes and circular translation of polysomal mRNA. Nucleic Acids Res 36:2476–2488

Lee L, Sykes BD (1980) Strategies for the uses of lanthanide NMR shift probes in the determination of protein structure in solutio. Application to the EF calcium binding site of carp parvalbumin. Biophys J 32:193–210

Leitz AJ, Bayburt TH, Barnakov AN, Springer BA, Sligar SG (2006) Functional reconstitution of Beta2-adrenergic receptors utilizing self-assembling Nanodisc technology. Biotechniques 40:601–602, 604, 606, passim

Liang B, Tamm LK (2007) Structure of outer membrane protein G by solution NMR spectroscopy. Proc Natl Acad Sci USA 104: 16140–16145

Liu XM, Sonar S, Lee CP, Coleman M, RajBhandary UL, Rothschild KJ (1995) Site-directed isotope labeling and FTIR spectroscopy: assignment of tyrosine bands in the bR→M difference spectrum of bacteriorhodopsin. Biophys Chem 56:63–70

Liu DV, Zawada JF, Swartz JR (2005) Streamlining *Escherichia coli* S30 extract preparation for economical cell-free protein synthesis. Biotechnol Prog 21:460–465

Matsuda T, Koshiba S, Tochio N, Seki E, Iwasaki N, Yabuki T, Inoue M, Yokoyama S, Kigawa T (2007) Improving cell-free protein synthesis for stable-isotope labeling. J Biomol NMR 37:225–229

McIntosh LP, Dahlquist FW (1990) Biosynthetic incorporation of 15N and 13C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. Q Rev Biophys 23:1–38

Morgner N, Barth HD, Brutschy B, Scheffer U, Breitung S, Gobel M (2008) Binding sites of the viral RNA element TAR and of TAR mutants for various peptide ligands, probed with LILBID: a new laser mass spectrometry. J Am Soc Mass Spectrom 19:1600–1611

Morita EH, Shimizu M, Ogasawara T, Endo Y, Tanaka R, Kohno T (2004) A novel way of amino acid-specific assignment in (1)H-(15)N HSQC spectra with a wheat germ cell-free protein synthesis system. J Biomol NMR 30:37–45

Muranaka N, Miura M, Taira H, Hohsaka T (2007) Incorporation of unnatural non-alpha-amino acids into the N terminus of proteins in a cell-free translation system. Chembiochem 8:1650–1653

Nirenberg MW (1963) Cell-free protein synthesis directed by messenger RNA. Methods Enzymol 6:17–23

Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG (1989) A general method for site-specific incorporation of unnatural amino acids into proteins. Science 244:182–188

Nozawa A, Nanamiya H, Miyata T, Linka N, Endo Y, Weber AP, Tozawa Y (2007) A cell-free translation and proteoliposome reconstitution system for functional analysis of plant solute transporters. Plant Cell Physiol 48:1815–1820

Ozawa K, Headlam MJ, Schaeffer PM, Henderson BR, Dixon NE, Otting G (2004) Optimization of an *Escherichia coli* system for cell-free synthesis of selectively N-labelled proteins for rapid analysis by NMR spectroscopy. Eur J Biochem 271:4084–4093

Ozawa K, Dixon NE, Otting G (2005) Cell-free synthesis of 15N-labeled proteins for NMR studies. IUBMB Life 57:615–622

Ozawa K, Wu PS, Dixon NE, Otting G (2006) N-Labelled proteins by cell-free protein synthesis. Strategies for high-throughput NMR studies of proteins and protein-ligand complexes. FEBS J 273:4154–4159

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

Pervushin K, Riek R, Wider G, Wuthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci USA 94:12366–12371

Poget SF, Cahill SM, Girvin ME (2007) Isotropic bicelles stabilize the functional form of a small multidrug-resistance pump for NMR structural studies. J Am Chem Soc 129:2432–2433

Prosser RS, Evanics F, Kitevski JL, Al-Abdul-Wahid MS (2006) Current applications of bicelles in NMR studies of membrane-associated amphiphiles and proteins. Biochemistry 45:8453–8465

Reckel S, Sobhanifar S, Schneider B, Junge F, Schwarz D, Durst F, Lohr F, Guntert P, Bernhard F, Dotsch V (2008) Transmembrane segment enhanced labeling as a tool for the backbone assignment of alpha-helical membrane proteins. Proc Natl Acad Sci USA 105:8262–8267

Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R, Choe S (2005) NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. Science 307:1317–1321

Sanders CRII, Schwonek JP (1992) Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR. Biochemistry 31:8898–8905

Schnell JR, Chou JJ (2008) Structure and mechanism of the M2 proton channel of influenza A virus. Nature 451:591–595

Schwarz D, Junge F, Durst F, Frolich N, Schneider B, Reckel S, Sobhanifar S, Dotsch V, Bernhard F (2007) Preparative scale expression of membrane proteins in *Escherichia coli*-based continuous exchange cell-free systems. Nat Protoc 2: 2945–2957

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

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

Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol 19:751–755

Shimizu Y, Kanamori T, Ueda T (2005) Protein synthesis by pure translation systems. Methods 36:299–304

Sonar S, Lee CP, Coleman M, Patel N, Liu X, Marti T, Khorana HG, RajBhandary UL, Rothschild KJ (1994) Site-directed isotope labelling and FTIR spectroscopy of bacteriorhodopsin. Nat Struct Biol 1:512–517

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

Swartz JR, Jewett MC, Woodrow KA (2004) Cell-free protein synthesis with prokaryotic combined transcription-translation. Methods Mol Biol 267:169–182

Torizawa T, Shimizu M, Taoka M, Miyano H, Kainosho M (2004) Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. J Biomol NMR 30:311–325

Trbovic N, Klammt C, Koglin A, Lohr F, Bernhard F, Dotsch V (2005) Efficient strategy for the rapid backbone assignment of membrane proteins. J Am Chem Soc 127:13504–13505

Van Horn WD, Kim HJ, Ellis CD, Hadziselimovic A, Sulistijo ES, Karra MD, Tian C, Sonnichsen FD, Sanders CR (2009) Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. Science 324:1726–1729

Veglia G, Opella SJ (2000) Lanthanide ion binding to adventitious sites aligns membrane proteins in micelles for solution NMR spectroscopy. J Am Chem Soc 122:11733–11734

Vinarov DA, Loushin Newman CL, Markley JL (2006) Wheat germ cell-free platform for eukaryotic protein production. FEBS J 273:4160–4169

Vold RR, Prosser RS, Deese AJ (1997) Isotropic solutions of phospholipid bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides. J Biomol NMR 9:329–335

Weigelt J, van Dongen M, Uppenberg J, Schultz J, Wikstrom M (2002) Site-selective screening by NMR spectroscopy with labeled amino acid pairs. J Am Chem Soc 124:2446–2447

Whorton MR, Bokoch MP, Rasmussen SG, Huang B, Zare RN, Kobilka B, Sunahara RK (2007) A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. Proc Natl Acad Sci USA 104:7682–7687

Wu PS, Ozawa K, Jergic S, Su XC, Dixon NE, Otting G (2006) Amino-acid type identification in 15N-HSQC spectra by combinatorial selective 15N-labelling. J Biomol NMR 34:13–21

Wuu JJ, Swartz JR (2008) High yield cell-free production of integral membrane proteins without refolding or detergents. Biochim Biophys Acta 1778:1237–1250

Yabuki T, Kigawa T, Dohmae N, Takio K, Terada T, Ito Y, Laue ED, Cooper JA, Kainosho M, Yokoyama S (1998) Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis. J Biomol NMR 11:295–306

Zhou Y, Cierpicki T, Jimenez RH, Lukasik SM, Ellena JF, Cafiso DS, Kadokura H, Beckwith J, Bushweller JH (2008) NMR solution structure of the integral membrane enzyme DsbB: functional insights into DsbB-catalyzed disulfide bond formation. Mol Cell 31:896–908