

Strategies for dealing with conformational sampling in structural calculations of flexible or kinked transmembrane peptides¹

Jan K. Rainey, Larry Fliegel, and Brian D. Sykes

Abstract: Peptides corresponding to transmembrane (TM) segments from membrane proteins provide a potential route for the determination of membrane protein structure. We have determined that 2 functionally critical TM segments from the mammalian Na⁺/H⁺ exchanger display well converged structure in regions separated by break points. The flexibility of these break points results in conformational sampling in solution. A brief review of available NMR structures of helical membrane proteins demonstrates that there are a number of published structures showing similar properties. Such flexibility is likely indicative of kinks in the full-length protein. This minireview focuses on methods and protocols for NMR structure calculation and analysis of peptide structures under conditions of conformational sampling. The methods outlined allow the identification and analysis of structured peptides containing break points owing to conformational sampling and the differentiation between oligomerization and ensemble-averaged observation of multiple peptide conformations.

Key words: NMR spectroscopy, membrane proteins, peptide structure, conformational sampling, oligomerization.

Résumé : Les peptides correspondant aux segments transmembranaires (TM) de protéines membranaires sont des outils potentiels à la détermination de la structure de telles protéines. Nous avons déterminé que deux segments TM fonctionnellement importants de l'échangeur Na⁺/H⁺ de mammifère possèdent des structures convergentes dans ces régions, séparées par des points de rupture. La flexibilité autour de ces points de rupture résulte en un échantillonnage conformationnel en solution. Une brève revue des structures disponibles déterminées par RMN de protéines à hélice membranaire démontre qu'il existe un certain nombre de structures publiées possédant des propriétés similaires. Une telle flexibilité est probablement indicative de la présence de coudes (*kinks*) dans la protéine de pleine longueur. Cette mini-revue se concentre sur les méthodes et les protocoles de calcul de la structure en RMN et d'analyse de la structure de peptides sous des conditions d'échantillonnage conformationnel. Les méthodes esquissées permettent l'identification et l'analyse de peptides structurés contenant des points de rupture dus à l'échantillonnage conformationnel, et la différenciation entre l'oligomérisation et l'observation pondérée des conformations de multiples peptides.

Mots clés : spectroscopie par RMN, protéines membranaires, structure peptidique, échantillonnage conformationnel, oligomérisation.

[Traduit par la Rédaction]

Introduction

The following minireview summarizes methods that we have developed to facilitate structural studies on the mammalian Na⁺/H⁺ exchanger isoform 1 (NHE1). The N-terminal cation exchange region of this protein, composed of

~500 out of 815 residues, is believed to be arranged in 12 transmembrane (TM) segments (Wakabayashi et al. 2000). Since full-length NHE1 cannot yet be expressed and purified in sufficient quantity for structural studies, we have been carrying out structural studies of isolated TM segments from functionally critical regions of NHE1. These peptides

Received 24 July 2006. Revision received 20 September 2006. Accepted 21 September 2006. Published on the NRC Research Press Web site at <http://bcf.nrc.ca> on 19 December 2006.

J.K. Rainey^{2,3} and B.D. Sykes⁴ Protein Engineering Network of Centres of Excellence and Department of Biochemistry, University of Alberta, Edmonton, AB T6G 2H7, Canada.

L. Fliegel. Department of Biochemistry, University of Alberta, Edmonton, AB T6G 2H7, Canada.

¹This paper is one of a selection of papers published in this Special Issue, entitled CSBMCB — Membrane Proteins in Health and Disease.

²Corresponding author (e-mail: jan.rainey@dal.ca).

³Present address: Department of Biochemistry & Molecular Biology, Dalhousie University, Halifax, NS B3H 1X5, Canada.

⁴Corresponding author (e-mail: brian.sykes@ualberta.ca).

Table 1. Summary of deposited PDB (Berman et al. 2000) entries of α -helical membrane proteins sorted by class of structure available in PDB.

Single TM segment, only 1 conformer provided	1CEK ¹ , 1EMZ ² , 1FDF ³ , 1HO2/1HO7 ⁴ , 1MP6 ⁵ , 1PJD ⁶ , 2C0X ⁷
TM helix showing variability at N and (or) C termini	1A11 ¹ , 1AMB/1AMC1 ⁸ , 1B9U ⁹ , 1BTQ/1BTR/1BTS/1BTT ¹⁰ , 1CKW/1CKX ¹¹ , 1DLZ ¹² , 1DTC ¹³ , 1DXZ ¹⁴ , 1GQ0 ¹⁵ , 1IH9 ¹⁶ , 1IIJ ¹⁷ , 1JDM ¹⁸ , 1QG9 ¹⁹ , 1R7C/1R7D/1R7E/1R7F/1R7G ²⁰ , 1SKH ²¹ , 1SPF ²² , 1Z65 ²³ , 2AJJ/2AJM/2AJN/2AJO ²⁴ , 2ARI ²⁵ , 2BP4 ²⁶ , 2NR1 ¹ , 3MRA ²⁷
Kinked TM segment with conformational sampling	1ATY ²⁸ , 1BA4 ²⁹ /1BA6 ³⁰ , 1BJB/1BJC ³¹ , 1BZK ³² , 1FJK/1FJP ³³ , 1Y4E ³⁴ , 2HTG ³⁵
Single TM helix with extramembrane portion	1BNX ³⁶ , 1FDM ³⁷ , 1IYT ³⁸ , 1JO5 ³⁹ , 1MOT ⁴⁰ , 1N7L ⁴¹ , 1WRG ⁴² , 1XRD ⁴² , 1Z9I ⁴³ , 1ZZA ⁴⁴ , 2CPB/2CPS ⁴⁵
Oligomer (no. of monomers)	1AFO ⁴⁶ (2), 1BHA/1BHB ⁴⁷ (2), 1EQ8 ¹ (5), 1NYJ ⁴⁸ (4), 1PI8/1PJE ⁴⁹ (4), 1ZLL ⁵⁰ (5), 2A9H ⁵¹ (4x2TM)
Multiple TM segments (no. of TM segments) ^d	1A91 ⁵² (2), 1BCT ⁵³ (2), 1C0V(2)/1C99(2)/1C17(12) ⁵⁴ , 1IJP ⁵⁵ (2), 1JFP ⁵⁶ (7), 1L0M ⁵⁷ (7), 1L6T ⁵⁸ (2), 1LN6 ⁵⁹ (7), 1VRY ⁶⁰ (2), 1WAZ ⁶¹ (2), 1WU0 ⁶² (2), 1YGM ⁶³ (4)

Note: The list has been annotated manually using results from both the Protein Data Bank and the Membrane Protein Data Bank (Raman et al. 2006) search engines. Superscripted numbers refer to the primary reference for each structure as follows: ¹, Opella et al. 1999; ², Op De Beeck et al. 2000; ³, Yeagle et al. 2000; ⁴, Ohlenschlager et al. 2002; ⁵, Wang et al. 2001; ⁶, Valentine et al. 2001; ⁷, Marvin et al. 2006; ⁸, Talafous et al. 1994; ⁹, Dmitriev et al. 1999; ¹⁰, Gargaro et al. 1994; ¹¹, Massiah et al. 1999; ¹², Balashova et al. 2000; ¹³, Bladon et al. 1992; ¹⁴, Pashkov et al. 1999; ¹⁵, Galbraith et al. 2003; ¹⁶, Shenkarev et al. 2002; ¹⁷, Goetz et al. 2001; ¹⁸, Mascioni et al. 2002; ¹⁹, Doak et al. 1996; ²⁰, Penin et al. 2004; ²¹, Biverstahl et al. 2004; ²², Johansson et al. 1994; ²³, Papadopoulos et al. 2006; ²⁴, Sapay et al. 2006; ²⁵, Jaroniec et al. 2005; ²⁶, Zirah et al. 2006; ²⁷, Lugovskoy et al. 1998; ²⁸, Girvin and Fillingame 1995; ²⁹, Coles et al. 1998; ³⁰, Watson et al. 1998; ³¹, Poulsen et al. 2000; ³², Chambers et al. 1999; ³³, Lamberth et al. 2000; ³⁴, Slepov et al. 2005; ³⁵, Ding et al. 2006; ³⁶, Chambers et al. 1998; ³⁷, Almeida and Opella 1997; ³⁸, Crescenzi et al. 2002; ³⁹, Sorgen et al. 2002; ⁴⁰, Yushmanov et al. 2003; ⁴¹, Zamoon et al. 2003; ⁴², Wang et al. 2005; ⁴³, Choowongkorn et al. 2005; ⁴⁴, Buck-Koehntop et al. 2005; ⁴⁵, Papavoine et al. 1998; ⁴⁶, MacKenzie et al. 1997; ⁴⁷, Pervushin et al. 1994; ⁴⁸, Nishimura et al. 2002; ⁴⁹, Park et al. 2003; ⁵⁰, Oxenoid and Chou 2005; ⁵¹, Yu et al. 2005; ⁵², Girvin et al. 1998; ⁵³, Barsukov et al. 1992; ⁵⁴, Rastogi and Girvin 1999; ⁵⁵, Dmitriev and Fillingame 2001; ⁵⁶, Yeagle et al. 2001; ⁵⁷, Katragadda et al. 2001a; ⁵⁸, Dmitriev et al. 2002; ⁵⁹, Choi et al. 2002; ⁶⁰, Ma et al. 2005; ⁶¹, Howell et al. 2005; ⁶², Nakano et al. 2006; ⁶³, Roosild et al. 2005;.

^dNote that some structures were solved by piecing together multiple segments and that some are isolated portions of a larger protein.

may be produced either by expression as a fusion protein (Lindhout et al. 2003; Slepov et al. 2005) or by chemical synthesis (Ding et al. 2006; Naider et al. 2005). Literature precedents show that isolated TM segments from the cystic fibrosis transmembrane conductance regulator (Oblatt-Montal et al. 1994; Wigley et al. 1998), bacteriorhodopsin (Hunt et al. 1997; Katragadda et al. 2001a), rhodopsin (Choi et al. 2002; Katragadda et al. 2001b; Yeagle et al. 2001), and the fungal G-protein-coupled receptor Ste2p (Naider et al. 2005) are all functional and have structures under membrane mimetic conditions that agree with the segment in the context of the protein as a whole, where this structure is known. It should be noted that some instances of variation in structural characteristics under different membrane mimetic conditions have been observed, such as bacteriorhodopsin (Hunt et al. 1997; Katragadda et al. 2001a), phospholamban (Lamberth et al. 2000, Oxenoid and Chou 2005, Zamoon et al. 2003), and TM IV of NHE1 (Slepov et al. 2005). Our philosophy is that the structure should be examined in multiple conditions where possible, although we have come across a case now (TM VII of NHE1), in which the peptide only stayed in solution under 1 condition out of ~ 10 that were tried. As abilities to predict membrane protein structure and packing improve, NMR structures of individual TM segments may work very nicely into full-length protein structure predictions, since non-canonical structures can be reliably determined and used as a basis for prediction.

Table 1 gives an overview by structural character of Protein Data Bank (PDB) (Berman et al. 2000) entries of transmembrane segments or membrane proteins with α -helical character solved by NMR methods to date. This list is an at-

tempt to be comprehensive based upon searches using both the PDB and Membrane Protein Data Bank (Raman et al. 2006) search engines. Division into the classes given in Table 1 is somewhat arbitrary, since some structures may display properties of more than 1 class in the table. In our experience, although isolated TM segment peptides are generally easier to obtain and handle than full-length membrane proteins, production and subsequent purification are still not trivial. This minireview covers the strategies that we have developed, which will be useful once this hurdle is overcome, to deal with the sampling of multiple peptide conformations that we have observed in our high-resolution solution-state NMR studies of TMs IV (Slepov et al. 2005) and VII (Ding et al. 2006) of NHE1.

Determining peptide structures

Compared with proteins that have a defined fold, peptides tend to be flexible and able to sample numerous conformations (comprehensively reviewed in Williamson and Waltho 1992). Solution-state methods to study proteins and peptides probe the entire ensemble of conformers present at any given time, and the type of structural information available is highly dependent upon the technique employed (Cai and Dass 2003). In many cases, it is impractical or impossible to deconvolute the contributions of different conformations to an ensemble-averaged experiment. In solution-state NMR spectroscopy, atomic-level resolution and differentiation of signals is possible. The typical, classical protocol (Wüthrich 1986) for solving a protein or peptide structure by NMR spectroscopy involves the following steps: first, determine chemical shift values characteristic of as many atoms as

Fig. 1. Conformational sampling about break points in peptide structures of isolated TM segments from α -helical membrane proteins seen in NMR structures deposited in the PDB (Table 1; other examples exist in the literature, but either are not deposited in the PDB or do not have multiple conformers available). Colours are blue-red over the spectrum from the N to C termini of the peptide in question. Arrows indicate the break point(s) in structure for each case. 1BZK, 1Y4E, and 2HTG were superposed over residues 15–28, 165–168, and 255–261, respectively. Trisolvent refers to methanol:chloroform:water (4:4:1, by volume).

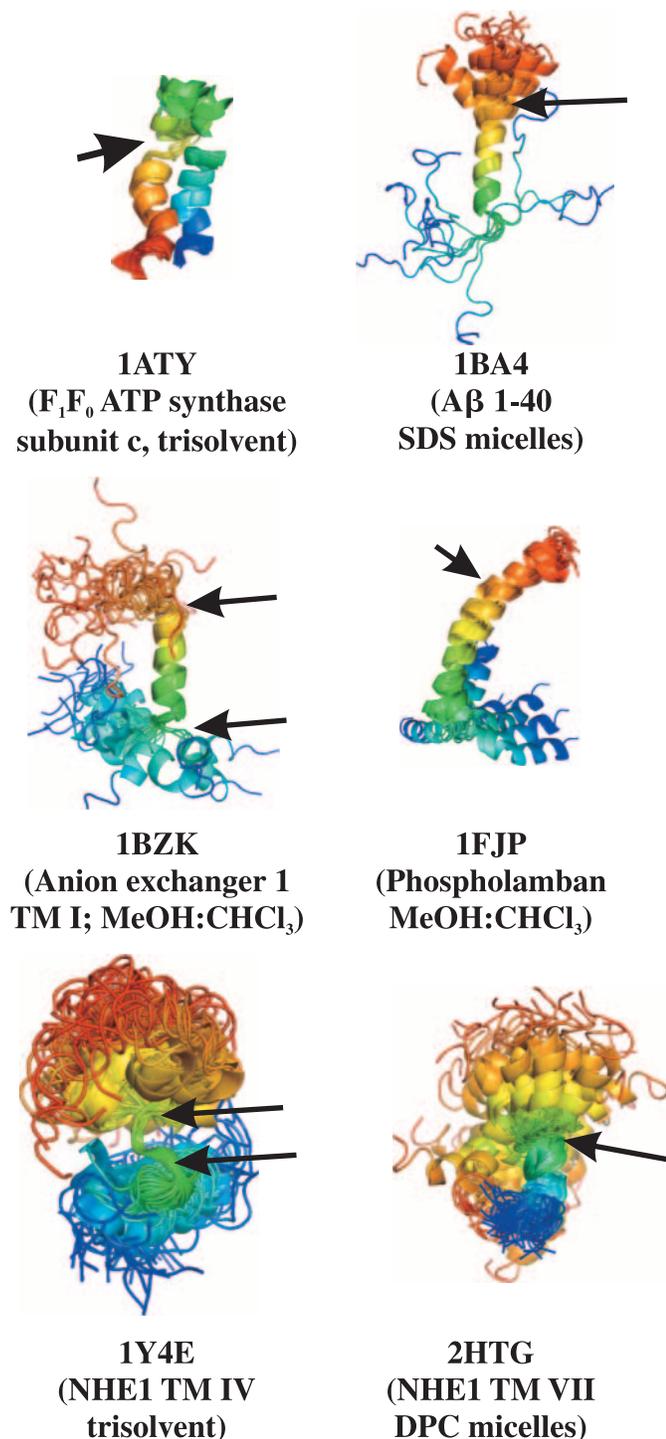
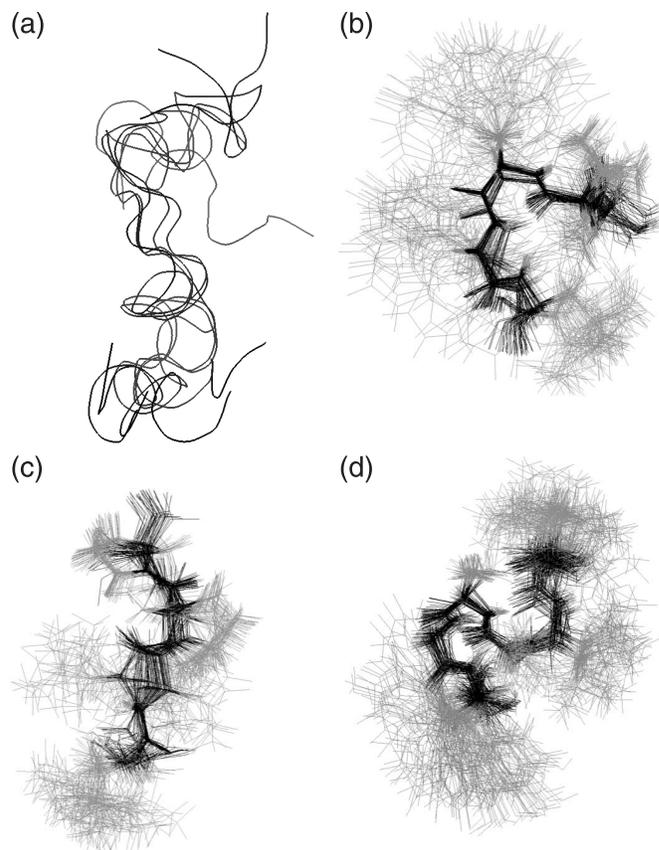
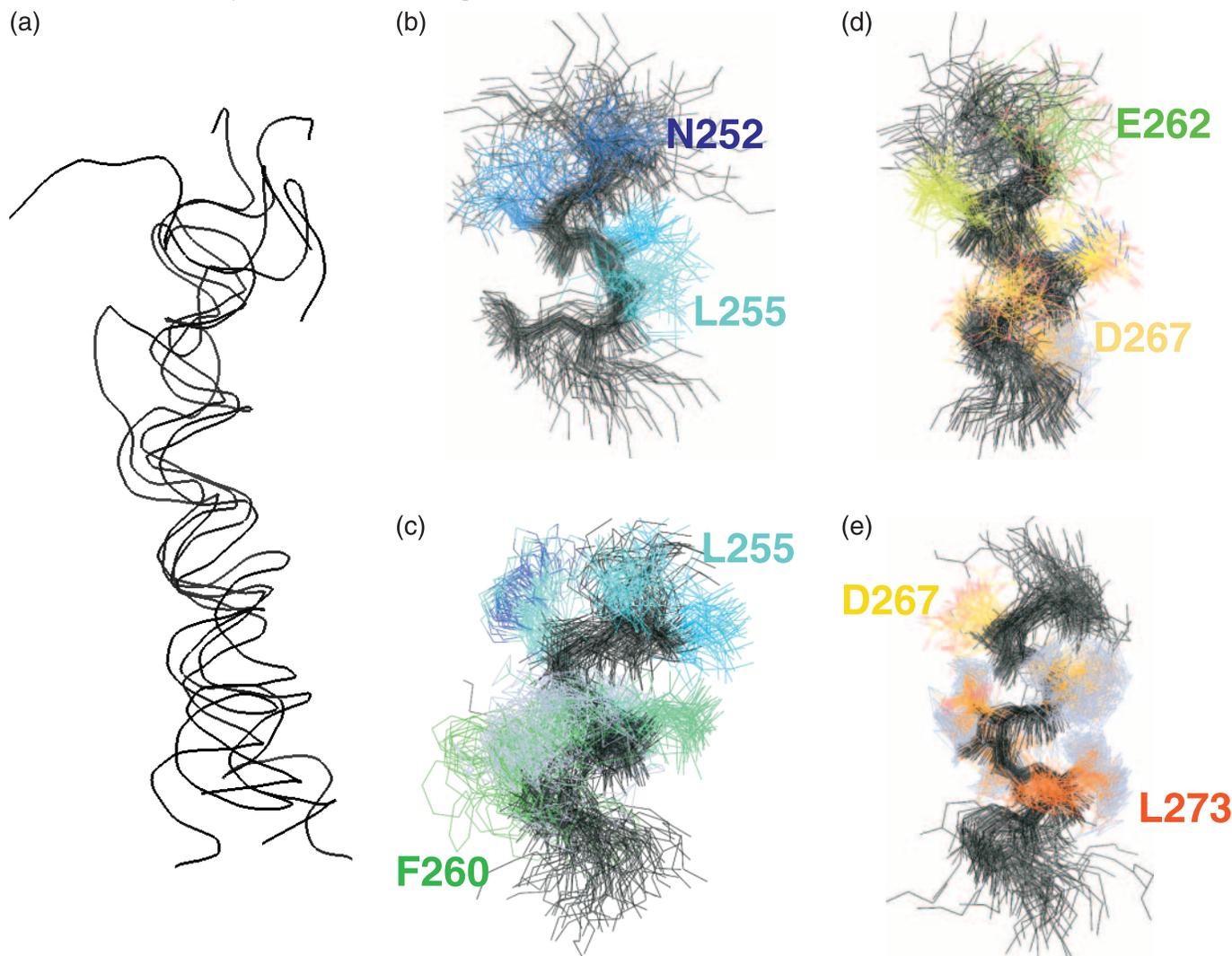


Fig. 2. Superpositions of the NMR structure of TM IV of NHE1 in mimetic solvent (methanol:chloroform:water; 4:4:1 by volume). (a) Superposition over backbone atoms residues S158–P178 (5 lowest energy structures shown out of 600) demonstrates poor convergence over the length of the peptide; N to C termini from top to bottom. (b–d) Superpositions over C α atoms of b, D159–L163; c, L165–P168; and d, I169–F176, with the backbone shown in black and side chains in grey for the 60 lowest energy structures out of 600 showing good backbone and reasonable side-chain convergence over each of the 3 separate structurally convergent regions. Break points in structural convergence of the peptide were determined to be at F164 and P168–I169.



possible within the sample; second, analyze ^1H - ^1H distances using the nuclear Overhauser effect (NOE); third, characterize deviations of chemical shift from random-coil values and through-bond J-couplings to assist in determining secondary structure; and, fourth, use NOE derived ^1H - ^1H distance constraints and chemical shift and (or) J-coupling-derived dihedral angle constraints to carry out simulated annealing calculations to produce polypeptide structures in agreement with the experimental data. Note that this is frequently an iterative process. Additional restraints, such as residual dipolar couplings, T_1 - T_2 relaxation data, or knowledge of molecular shape (e.g., radius of gyration) are often introduced during simulated annealing (comprehensively reviewed in Schwieters et al. 2006). However, because peptides are most typically produced through chemical synthesis, the former 2 types of restraints requiring isotopic labeling are often not financially feasible; in the latter case, the conformational flexibility we are trying to handle makes a description of molecular shape difficult, if at all meaningful.

Fig. 3. Superpositions of the NMR structure of TM VII of NHE1 in DPC micelles. (a) Superposition over backbone atoms residues L254–V271 (5 lowest energy structures shown out of 66) demonstrates poor convergence over length of peptide; N to C termini from top to bottom. (b–d) Superpositions over backbone atoms of b, I251–L255; c, L255–F260; d, E262–V269; and e, D267–Y274, with the backbone shown in black and side chains in grey for all 66 retained ensemble members showing good backbone and reasonable side-chain convergence over each of the 4 separate structurally convergent regions. A major break point in converged structure was determined to be at G261; backbone variability was observed about the points L255 and T270–V271.



Although many peptides exhibit flexibility in solution, sample a variety of conformers, and display only nascent structure indicated by low numbers of ^1H – ^1H NOE contacts, this has not been our experience with either the TM IV or TM VII segments of NHE1. This is not a surprising observation if the physiologically relevant structure of the TM segment is preferred in membrane mimetic environment, as has been reported for the membrane proteins listed in the Introduction, since in this case, structural convergence would be expected. Furthermore, numerous PDB structures of transmembrane peptides show structures that have been well resolved by NMR (Table 1).

Both of the TM segments of NHE1 that we have studied have had break points in converged structure, which we believe to be indicative of physiologically relevant kinks in the TM segment structure. Several examples of structures show-

ing this characteristic are now deposited in the PDB (Table 1; Fig. 1). Figures 2 and 3 show examples, using the solved structures of the TM IV and TM VII segments of NHE1. In each case, attempts to superpose the backbone over the majority of the peptide length do not provide good superposition. For example, even with the 5 lowest energy structures out of the entire ensemble of 600 (TM IV) or 66 (TM VII) structures, a great deal of variation in backbone structure can be seen with an attempt to superpose over a major portion of the peptide (Figs. 2a and 3a). However, it is possible to achieve excellent superpositions over 4–9 residue stretches of each peptide (Figs. 2b–2d and Figs. 3b–3e) after taking into account the identified break points in structural convergence. As a whole, these break points lead to conformational sampling of a restricted nature. In Figs. 2a and 3a, this is reflected in the fact that the 5 lowest energy

structures out of the ensemble of 600 retained NMR structures are all extended to approximately the same degree, even though the backbones themselves do not superpose well. This restricted sampling has allowed us to develop methods to examine these peptide structures in detail. These methods are applicable to any situation in which regions of a peptide or protein have well defined structures but display a variety of positions and orientations relative to other structurally defined regions.

Summary of structure calculation and analysis strategies

For each of the 2 TM segments of NHE1 that we have studied so far, a different calculation protocol was used. In the first case, the TM IV segment studied in membrane mimetic solvents (Slepko et al. 2005), the experimental data observed was well represented through the calculation of a single conformer. However, flexibility at distinct points in the peptide led to conformational sampling by TM IV. In structure calculations, we used large ensembles of structures (~ 1000) to allow extensive sampling about these break points. Analysis of the resulting ensembles (typically 600 of 1000 structures were retained) required some specialized strategies. In the second case, the TM VII segment was studied in micelles (Ding et al. 2006). We could not satisfy $\sim 1/3$ of the experimental NOE restraints using a single conformer calculation. However, by introducing a second conformer and allowing each NOE restraint to be satisfied in 1, both, or between conformers, we were able to satisfy all restraints. The resulting ensembles of structures also showed break points in converged structure similar to those observed in TM IV. Therefore, a similar analysis of the ensemble was carried out to determine features of the structure. The following discussion covers 3 major points: using structure calculations to differentiate between data indicative of single-conformer, multiple-conformer, or oligomeric peptides; analysis of ensembles of peptide conformers undergoing conformational sampling about break points in converged structure; and validation and analysis of such a peptide structure.

Single-conformer, oligomeric, or multiple-conformer structure calculations?

In the case of wide variability in relative positions of structurally converged peptide segments about a break point, conformational sampling may result in a relative lack of NOE constraints between atoms on either side of the break point. With a situation of this sort, structure calculation using a single conformer is likely to satisfy the vast majority of observed NOE constraints. The most important considera-

tion for structure calculation with this type of sample is the production of sufficient ensemble members to ensure that the calculated structures provide a representation of the actual range of sample conformations, rather than an artificially limited subset of these conformations. In the case of TM IV of NHE1, we used very large ensembles of ~ 1000 structures to ensure that this was the case. Two points should be considered here: the experimental data should be well matched by the ensemble of NMR structures, although higher numbers of NOE violations should be expected and allowed for such ensembles of structures than for a folded protein; and superposition over the entire structure will not generally be possible, making analysis much more cumbersome, as discussed below. Typically, this should be the first type of structure calculation attempted.

If a single-conformer calculation cannot consistently provide structures that fit a large proportion of the experimental data, then the unsatisfied NOE restraints may be the result of preferential sampling of certain conformations and (or) oligomerization of the peptide. Because either of these situations is readily possible, we developed a dual-conformer calculation protocol that does not preferentially select for one case or the other. In this case, we carried out structure calculations, starting with 2 extended polypeptide chains (we will refer to these as chains A and B). For any given NOE (say between residue x – atom 1 and residue y – atom 2), we set up an ambiguous restraint which may be satisfied by one or more of the following:

chain A – residue x – atom 1 \rightarrow chain A – residue y – atom 2
 chain B – residue x – atom 1 \rightarrow chain B – residue y – atom 2
 chain A – residue x – atom 1 \rightarrow chain B – residue y – atom 2
 chain B – residue x – atom 1 \rightarrow chain A – residue y – atom 2
 The distances between each of these pairs of atoms, d_i , d_{ii} , d_{iii} , and d_{iv} are then averaged in the following manner

$$d_{\text{avg}} = (d_i^{-6} + d_{ii}^{-6} + d_{iii}^{-6} + d_{iv}^{-6})^{-1/6}$$

during structure calculation. Although unintuitive, this style of d^{-6} summed averaging means that if any 1 of the 4 distances is much shorter than the rest, the value of d_{avg} would represent this fact. For comparison, a standard d^{-6} average would be given by

$$d_{\text{avg}} = \left(\frac{d_i^{-6} + d_{ii}^{-6} + d_{iii}^{-6} + d_{iv}^{-6}}{4} \right)^{-1/6}$$

Using XPLOR-NIH (Schwieters et al. 2003), this is implemented using the SUM averaging option. The actual syntax to specify a given ambiguous restraint in XPLOR-NIH is as follows:

((residue x and atom 1 and segid A) OR (residue x and atom 1 and segid B))

((residue y and atom 2 and segid A) OR (residue y and atom 2 and segid B))

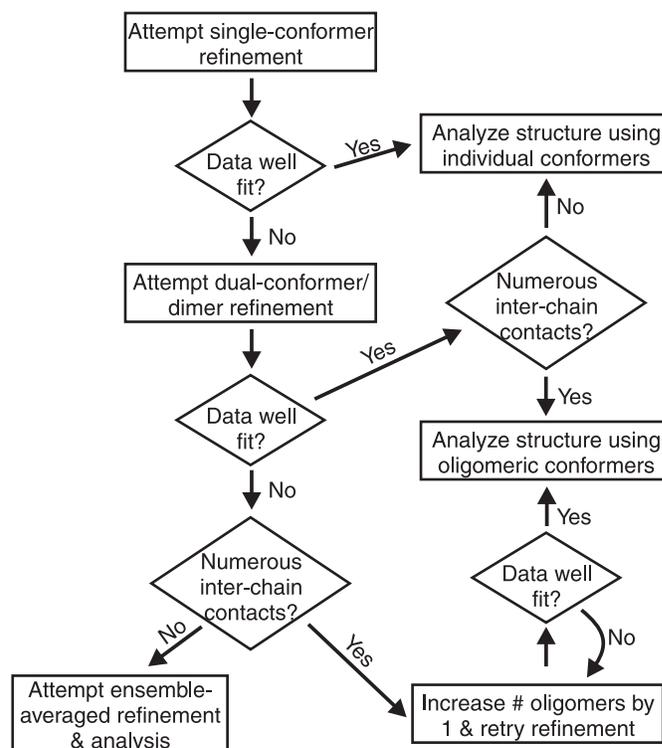
$d_{\text{minus}} \ d_{\text{plus}}$

where d , d_{minus} , and d_{plus} will be interpreted in the manner specified by the NOE restraining function selected using the potential statement. Note that further ambiguity can easily be nested within a given selection statement of the restraint, whether through nonstereospecific pseudoatom use or by additional OR statements. Production of these ambiguous restraints from a starting single-conformer restraint file or from a NOE spectrum peak data file is readily automated, and analysis of the resulting ensemble of structures in terms of degree of constraint agreement and oligomerization vs multiple conformers is best done using custom scripts (Tcl/Tk scripts for both constraint file production and ensemble analysis are freely available upon request).

If this dual-conformer–dimer approach still is unable to provide an ensemble of structures in agreement with the NMR data, 2 possible routes to a structure are feasible. If many inter-chain contacts are being observed, it is likely that oligomer formation is represented by the data. In this case, the dimer–dual-conformer method outlined above can be iteratively expanded with increasing numbers of polypeptide chains until the experimental data are satisfied by the ensemble of calculated structures. Alternately, if dimer formation is not apparent in the resulting ensemble of structures, methods that implement ensemble-averaged NOE restraints are likely required. Essentially, for an NOE to be observed in a flexible peptide, meaning that 2 protons are in close proximity (i.e., $< \sim 6 \text{ \AA}$; $1 \text{ \AA} = 0.1 \text{ nm}$), 2 cases are possible: either a significant population of the peptides in solution are in a conformation where the protons are in close proximity (i.e., the average of the ensemble of structures gives an observable ^1H – ^1H contact) or the ensemble of peptides in solution is undergoing conformational fluctuation, in which these protons come into proximity for some fraction of the time (i.e., the average over time of the ^1H – ^1H distance obeys the restraint). A number of methods to carry out ensemble refinement have been published (Bonvin and Brunger 1995; Bonvin et al. 1994; Bruschiweiler et al. 1991; Wang et al. 1995), and XPLOR-NIH now has this capability as a standard (Schwieters et al. 2003, 2006). Therefore, we will not discuss these methods in detail in this minireview. The results of an ensemble-averaged refinement carried out under any of these protocols may be readily analyzed using the methods we present.

Based upon our experiences with TMs IV and VII of NHE1, the protocol shown in the flowchart in Fig. 4 is recommended. Note that, in these methods, it is entirely up to the individual to determine what constitutes a reasonable or significant number of NOE violations. For example, in the case of our published TM IV peptide structure, single-conformer methods provided an ensemble of structures with few NOE violations (12 violations of magnitude $>0.2 \text{ \AA}$ and $<0.3 \text{ \AA}$ for 600 structures) and strong structural convergence in 3 regions separated by 2 break points. This convergence and degree of NOE agreement was observed after several rounds of iteratively lengthening NOE restraints. If we introduce a second conformer, NOE violations become nonexistent with very few modifications to the lengths of restraints, but convergence of the polypeptide backbone (or of all heavy atoms) in the 3 structured regions is decreased compared with the single conformer calculation (results not shown). Therefore, in this case, the single-conformer method

Fig. 4. Flowchart demonstrating a suggested NMR structure calculation strategy for peptides undergoing conformational sampling and (or) oligomerization. Note that each “data well fit?” test may require significant refinement of restraint data prior to making a yes or no decision.



provides a superior structure. Alternatively, our structure of TM VII of NHE1 shows much better structural convergence in the double-conformer case than the single-conformer case. (In the single-conformer case, reasonable convergence and violation statistics required discarding of $\sim 1/3$ of the NOE restraints.) Each of these cases demonstrates a clearly better ability to satisfy the experimental data in structure calculation using one method over the other. Furthermore, the dual-conformer (easily extended to multiconformer) calculation method we have developed readily allows the differentiation between oligomer formation and conformational sampling. In the case that a dimer or higher oligomer is observed, the degree to which the structure is affected by the oligomerization should be examined. If the structure is indeed strongly affected, the question of whether this is an artifact owing to the solution conditions chosen or whether this is actually a feature that would stabilize the peptide within a physiological context must be addressed. This would likely require solution of the structure under a further set of conditions, if possible.

Determination of break points in structure

A typical peptide having break points in structure will not appear to be well converged, since attempts to superpose the backbone over the majority of its length will lead to poor superpositions shown by high root mean square deviations (RMSDs). This can be seen in Figs. 2a and 3a, in which even the 5 lowest-energy structures of TMs IV (Fig. 2a)

and VII (Fig. 3a) do not superpose well over the length of the peptide. We find a very useful starting point to be the NMRCORE program, which is designed to search an ensemble of protein NMR structures for “core domains” of atoms that always occupy the same relative positions between ensemble members (Kelley et al. 1997). In our experience, NMRCORE is also well suited to determining regions converged structures within peptides. Our typical analysis consists of the following: (i) NMRCORE is used to suggest regions of structural convergence, typically using C α , as opposed to backbone, provides a better overall estimate; (ii) the ensemble of structures starting from suggested core regions of NMRCORE is superposed, and the RMSD of superposed ensemble is determined; (iii) the region used for superposition is iteratively varied, moving both the N and C termini from suggested NMRCORE core regions until the optimal RMSD is found.

To carry out superpositions, we have had great success with the LSQKAB implementation (Collaborative Computational Project, Number 4 1994) of the method of Kabsch (Kabsch 1976). This software is readily run from scripts, allowing numerous superpositions to be carried out over user specified ranges, with minimal changes required to the script. Examples of the results of this process are shown in Figs. 5a–5b, using the RMSDs observed over structurally convergent regions of the TM IV (Fig. 2) and VII (Fig. 3) segments of NHE1 for, respectively, the 600 and 66 member ensembles of structures upon optimal superposition. In each case, the boxed residues show a distinctly lower RMSD than those that are immediately N and C terminal, which fall outside of the converged region. These jumps in RMSD serve as a strong indicator of the best regions for superposition. Break points in structure occur at F164 and at P168–I169 for TM IV, and at L255, G261, and T270–V271 for TM VII.

Although we find NMRCORE to be useful for finding general bounds of converged regions, another strong diagnostic of break points in peptide structure is unusual clustering or larger than normal spread of (ϕ, ψ) dihedral angles. PROCHECK–NMR (Laskowski et al. 1993, 1996) is a useful tool for producing Ramachandran plots on a per-residue basis. Unfortunately, because many of the peptide ensembles we use have more than 50 members, this software may not be able to provide full plots of all ensemble members. Therefore, we have also developed our own scripts (freely available upon request) that calculate dihedral angles and allow us to look for these diagnostic dihedral angle clusters or spreads in an ensemble with any number of members. A further useful diagnostic is the dihedral angle order parameter (S) (Hyberts et al. 1992). Calculation of S for a given dihedral angle, α , of a given residue in an ensemble of structures employs the two-dimensional vector average

$$S = \frac{1}{N} \left| \sum_{i=1}^N [(\cos\alpha_i)\hat{x} + (\sin\alpha_i)\hat{y}] \right|$$

where N is the number of structures in the ensemble, α_i is the dihedral angle of the residue in question for structure i , and \hat{x} and \hat{y} are orthogonal unit vectors. The value of S (which varies from 0–1) is 1 for the case in which all dihedrals are equal, and small for widely spread angles. In the case of TM VII from NHE1, for example, there is a major

break point in structure at G261 (seen in Fig. 1) and minor breaks that are N terminal to L255 and V271. These break points are reflected by higher than usual average deviations in ϕ and ψ for G261, in ϕ for L255 and V271 (Fig. 5d), and drops in S for ϕ and ψ (Fig. 5f). The dihedral angle average deviations (Fig. 5c) and S values (Fig. 5e) are less optimal for TM IV, despite good structural convergence (Fig. 5a). In this case, however, these values are still helpful in initial assignment of the well converged regions of the structure (particularly from residues 171–176) and in the identification of major break points in structure. An examination of these kinds of trends, as well as actual Ramachandran plots, assisted us in selecting the optimal regions for superposition, which were slightly different from the core regions determined by NMRCORE.

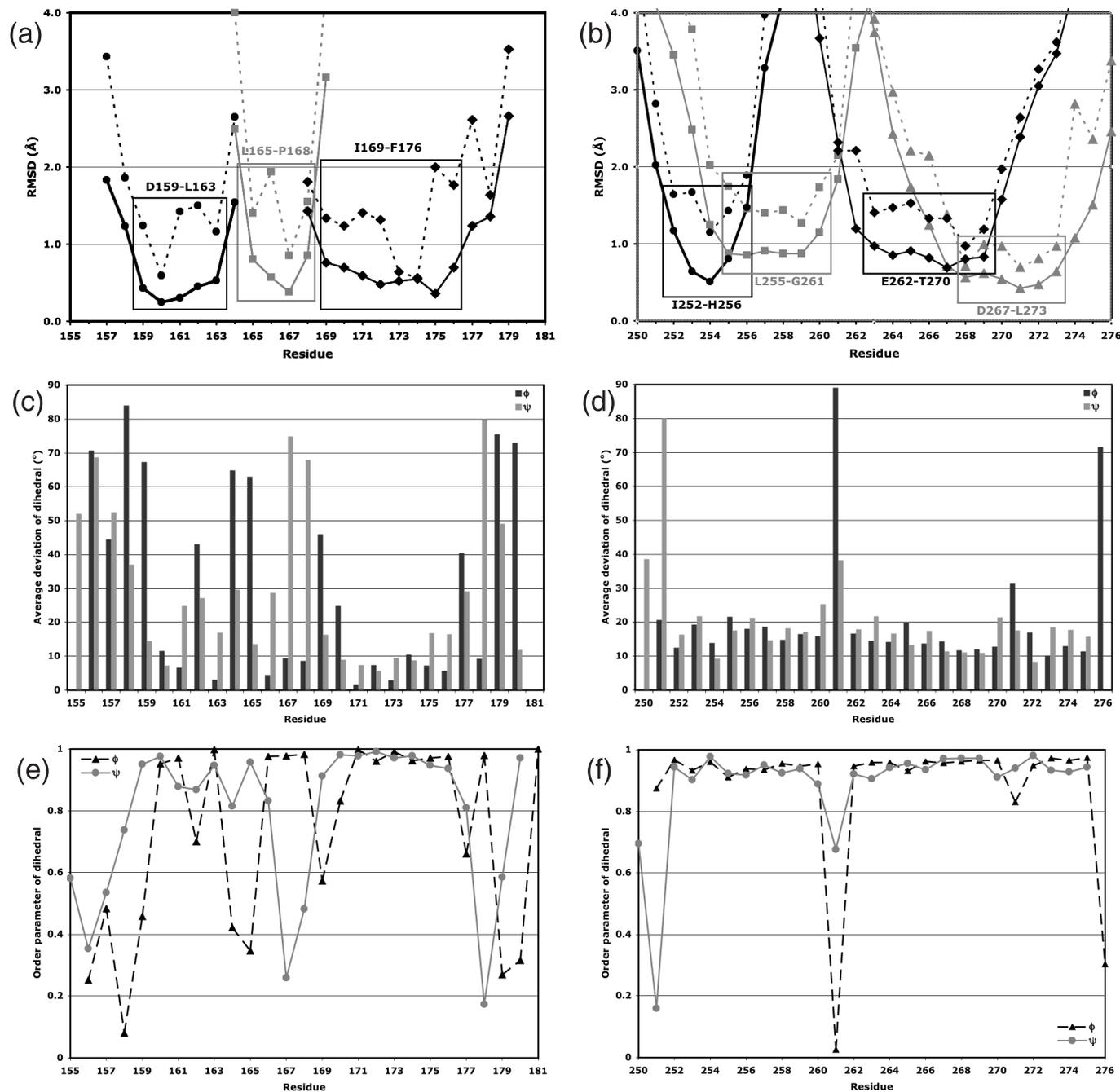
The use of hydrogen–deuterium exchange rates is also potentially very valuable for the identification of break points in a peptide structure. Examples of these data in the literature show that slower-exchanging regions are typically identified as those that are helical and have low RMSDs (Chambers et al. 1999; Coles et al. 1998; Girvin and Fillingame 1995; Lamberth et al. 2000; Poulsen et al. 2000; Watson et al. 1998), while regions that we would label structural break points are faster exchanging. As yet, there has not been an isotopically labeled peptide showing conformational sampling in the TM region (Table 1) characterized using dynamics measurement. However, it is likely that order parameters obtained from ^{15}N and (or) ^{13}C relaxation measurements would also be a useful indicator for break-points vs structurally convergent stretches, as was seen at the break between the AP and TM helices in the phospholamban pentamer structure of Oxenoid and Chou (2005).

Analysis and validation of peptide structures

To analyze NOE violations, dihedral angle spread and violations, determine RMSDs for structurally converged regions of a peptide, and examine the propensity for dimer formation vs observation of more than one conformation in the double-conformer calculation case, we tend to use custom scripts (freely available upon request). To provide standard Ramachandran plot statistics, we typically make use of PROCHECK–NMR by subdividing an ensemble of structures into smaller, equal-sized ensembles and averaging the PROCHECK derived statistics.

The software PROMOTIF–NMR (Hutchinson and Thornton 1996) is extremely useful for analyzing structural characteristics of an ensemble of structures. Because peptide structures have some degree of inherent variability, unless strong H-bond-type constraints are introduced, attempts to classify secondary structure are sometimes challenging. For example, a given secondary structure (i.e., an α helix) may have slight variability in its N- and C-terminal residues over the ensemble of structures. Rather than relying entirely upon secondary structure assignments for individual peptide structures, we have found that the most useful way to analyze structural features is to look for consensus in the PROMOTIF or Kabsch–Sander (Kabsch and Sander 1983) secondary structure assignments over the ensemble of structures. In regions with well defined RMSDs upon superposition and consistent dihedral angle observations, such consensus as-

Fig. 5. RMSD (on a per-residue basis) for (a) all 600 members of the NMR ensemble for 3 superpositions of the TM IV segment of NHE1, and (b) all 66 members of the NMR ensemble for TM VII. The residue ranges shown correspond to the superposition range used (using C α (a) or backbone atoms (b) to superpose) for each structurally converged segment. Boxes surround the well converged regions, as assigned by low RMSDs. Dashed lines are all-atom RMSDs, while solid lines are C α (a) or backbone (b) RMSDs. Average deviation ((c) TM IV; (d) TM VII) and order parameter (*S*; (e) TM IV; (f) TM VII) of ϕ and ψ dihedral angles for each entire retained ensemble of NMR structures.



signments are indicative of the structure that would be expected for the peptide, even if not all members of the ensemble do not display the same degree of structuring.

Summary

We have found that 2 functionally important transmembrane segments from mammalian NHE1 display conformational sampling when studied as isolated peptides in a

membrane mimetic environment. It makes sense that a TM segment which undergoes conformational changes as a part of regular protein function will display some inherent structural flexibility. In the case of a kinked α -helical TM segment, the break in regular H bonding at the kink may also be expected to allow flexibility, depending upon the structure at the kink, upon isolation from the protein. Therefore, the types of conformational sampling we and other groups have observed are likely to be an ongoing issue with the

study of isolated TM segments. Although these studies are not trivial, we hope that the different strategies that we have outlined here to calculate these classes of structures, to analyze structurally converged regions separated by flexible break points, and to validate an ensemble of structures relative to the available NMR data will provide useful starting points and methods for studies of isolated TM segments or of any peptide undergoing conformational sampling. The method that we present for multiple-conformer calculation also allows easy testing of oligomer formation, and differentiation between oligomerization and conformational sampling.

Acknowledgements

We thank Ryan Hoffman and Olivier Julien (B.D.S. laboratory) for helpful conversations regarding structural calculations and PyMol, and Charles Schwieters from the National Institutes of Health (Bethesda, Md.) for helpful discussions regarding XPLOR-NIH. J.K.R. is grateful for Postdoctoral Fellowships from the Natural Sciences and Engineering Research Council of Canada, the Alberta Heritage Foundation for Medical Research (AHFMR), and the Canadian Institutes of Health Research (CIHR) Strategic Training Program in Membrane Proteins and Cardiovascular Disease. L.F. is supported by a grant from CIHR and a Scientist award from AHFMR. B.D.S. receives support as a Canada Research Chair in Structural Biology. Structural calculation method development was funded by grants awarded to B.D.S. by the Canadian Protein Engineering Network of Centres of Excellence.

References

- Almeida, F.C., and Opella, S.J. 1997. fd coat protein structure in membrane environments: structural dynamics of the loop between the hydrophobic trans-membrane helix and the amphipathic in-plane helix. *J. Mol. Biol.* **270**: 481–495. doi:10.1006/jmbi.1997.1114. PMID:9237913.
- Balashova, T.A., Shenkarev, Z.O., Tagaev, A.A., Ovchinnikova, T.V., Raap, J., and Arseniev, A.S. 2000. NMR structure of the channel-former zervamicin IIB in isotropic solvents. *FEBS Lett.* **466**: 333–336. doi:10.1016/S0014-5793(99)01707-X. PMID:10682854.
- Barsukov, I.L., Nolde, D.E., Lomize, A.L., and Arseniev, A.S. 1992. Three-dimensional structure of proteolytic fragment 163–231 of bacterioopsin determined from nuclear magnetic resonance data in solution. *Eur. J. Biochem.* **206**: 665–672. doi:10.1111/j.1432-1033.1992.tb16972.x. PMID:1606953.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. 2000. The protein data bank. *Nucleic Acids Res.* **28**: 235–242. doi:10.1093/nar/28.1.235. PMID:10592235.
- Biverstahl, H., Andersson, A., Graslund, A., and Maler, L. 2004. NMR solution structure and membrane interaction of the N-terminal sequence (1–30) of the bovine prion protein. *Biochemistry*, **43**: 14940–14947. doi:10.1021/bi0485070. PMID:15554701.
- Bladon, C.M., Bladon, P., and Parkinson, J.A. 1992. Delta-toxin and analogues as peptide models for protein ion channels. *Biochem. Soc. Trans.* **20**: 862–864. PMID:1283135.
- Bonvin, A.M., and Brunger, A.T. 1995. Conformational variability of solution nuclear magnetic resonance structures. *J. Mol. Biol.* **250**: 80–93. doi:10.1006/jmbi.1995.0360. PMID:7602599.
- Bonvin, A.M.J.J., Boelens, R., and Kaptein, R. 1994. Time-and-ensemble-averaged direct NOE restraints. *J. Biomol. NMR*, **4**: 143–149. doi:10.1007/BF00178343.
- Bruschweiler, R., Blackledge, M., and Ernst, R.R. 1991. Multi-conformational peptide dynamics derived from NMR data: a new search algorithm and its application to antamanide. *J. Biomol. NMR*, **1**: 3–11. PMID:1841688.
- Buck-Koehntop, B.A., Mascioni, A., Buffy, J.J., and Veglia, G. 2005. Structure, dynamics, and membrane topology of stannin: a mediator of neuronal cell apoptosis induced by trimethyltin chloride. *J. Mol. Biol.* **354**: 652–665. PMID:16246365.
- Cai, X.M., and Dass, C. 2003. Conformational analysis of proteins and peptides. *Curr. Org. Chem.* **7**: 1841–1854. doi:10.2174/1385272033486161.
- Chambers, E.J., Bloomberg, G.B., Ring, S.M., and Tanner, M.J. 1999. Structural studies on the effects of the deletion in the red cell anion exchanger (band 3, AE1) associated with South East Asian ovalocytosis. *J. Mol. Biol.* **285**: 1289–1307. doi:10.1006/jmbi.1998.2392. PMID:9887277.
- Chambers, E.J., Askin, D., Bloomberg, G.B., Ring, S.M., and Tanner, M.J. 1998. Studies on the structure of a transmembrane region and a cytoplasmic loop of the human red cell anion exchanger (band 3, AE1). *Biochem. Soc. Trans.* **26**: 516–520. PMID:9765907.
- Choi, G., Landin, J., Galan, J.F., Birge, R.R., Albert, A.D., and Yeagle, P.L. 2002. Structural studies of metarhodopsin II, the activated form of the G-protein coupled receptor, rhodopsin. *Biochemistry*, **41**: 7318–7324. doi:10.1021/bi025507w. PMID:12044163.
- Choowongkorn, K., Carlin, C.R., and Sonnichsen, F.D. 2005. A structural model for the membrane-bound form of the juxta-membrane domain of the epidermal growth factor receptor. *J. Biol. Chem.* **280**: 24043–24052. doi:10.1074/jbc.M502698200. PMID:15840573.
- Coles, M., Bicknell, W., Watson, A.A., Fairlie, D.P., and Craik, D.J. 1998. Solution structure of amyloid beta-peptide(1–40) in a water-micelle environment. Is the membrane-spanning domain where we think it is? *Biochemistry*, **37**: 11064–11077. doi:10.1021/bi972979f. PMID:9693002.
- Collaborative Computational Project, Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**: 760–763. doi:10.1107/S0907444994003112. PMID:15299374.
- Crescenzi, O., Tomaselli, S., Guerrini, R., Salvadori, S., D’Ursi, A.M., Temussi, P.A., and Picone, D. 2002. Solution structure of the Alzheimer amyloid beta-peptide (1–42) in an apolar micro-environment. Similarity with a virus fusion domain. *Eur. J. Biochem.* **269**: 5642–5648. doi:10.1046/j.1432-1033.2002.03271.x. PMID:12423364.
- Ding, J., Rainey, J.K., Xu, C., Sykes, B.D., and Fliegel, L. 2006. Structural and functional characterization of TM VII of the NHE1 isoform of the Na⁺/H⁺ exchanger. *J. Biol. Chem.* **281**: 29817–29829. doi:10.1074/jbc.M606152200. PMID:16861220.
- Dmitriev, O.Y., and Fillingame, R.H. 2001. Structure of Ala(20) → Pro/Pro(64) → Ala substituted subunit c of *Escherichia coli* ATP synthase in which the essential proline is switched between transmembrane helices. *J. Biol. Chem.* **276**: 27449–27454. doi:10.1074/jbc.M100762200. PMID:11331283.
- Dmitriev, O.Y., Jones, P.C., and Fillingame, R.H. 1999. Structure of the subunit c oligomer in the F1Fo ATP synthase: model derived from solution structure of the monomer and cross-linking in the native enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **96**: 7785–7790. doi:10.1073/pnas.96.14.7785. PMID:10393899.
- Dmitriev, O.Y., Abildgaard, F., Markley, J.L., and Fillingame, R.H. 2002. Structure of Ala24/Asp61 → Asp24/Asn61 substituted subunit c of *Escherichia coli* ATP synthase: implications for the

- mechanism of proton transport and rotary movement in the F₀ complex. *Biochemistry*, **41**: 5537–5547. doi:10.1021/bi012198l. PMID:11969414.
- Doak, D.G., Mulvey, D., Kawaguchi, K., Villalain, J., and Campbell, I.D. 1996. Structural studies of synthetic peptides dissected from the voltage-gated sodium channel. *J. Mol. Biol.* **258**: 672–687. doi:10.1006/jmbi.1996.0278. PMID:8637001.
- Galbraith, T.P., Harris, R., Driscoll, P.C., and Wallace, B.A. 2003. Solution NMR studies of antiameobin, a membrane channel-forming polypeptide. *Biophys. J.* **84**: 185–194. PMID:12524274.
- Gargaro, A.R., Bloomberg, G.B., Dempsey, C.E., Murray, M., and Tanner, M.J. 1994. The solution structures of the first and second transmembrane-spanning segments of band 3. *Eur. J. Biochem.* **221**: 445–454. doi:10.1111/j.1432-1033.1994.tb18757.x. PMID:8168533.
- Girvin, M.E., and Fillingame, R.H. 1995. Determination of local protein structure by spin label difference 2D NMR: the region neighboring Asp61 of subunit c of the F₁F₀ ATP synthase. *Biochemistry*, **34**: 1635–1645. doi:10.1021/bi00005a020. PMID:7849023.
- Girvin, M.E., Rastogi, V.K., Abildgaard, F., Markley, J.L., and Fillingame, R.H. 1998. Solution structure of the transmembrane H⁺-transporting subunit c of the F₁F₀ ATP synthase. *Biochemistry*, **37**: 8817–8824. doi:10.1021/bi980511m. PMID:9636021.
- Goetz, M., Carlotti, C., Bontems, F., and Dufourc, E.J. 2001. Evidence for an alpha-helix → pi-bulge helicity modulation for the neu/erbB-2 membrane-spanning segment. A ¹H NMR and circular dichroism study. *Biochemistry*, **40**: 6534–6540. doi:10.1021/bi0027938. PMID:11371217.
- Howell, S.C., Mesleh, M.F., and Opella, S.J. 2005. NMR structure determination of a membrane protein with two transmembrane helices in micelles: MerF of the bacterial mercury detoxification system. *Biochemistry*, **44**: 5196–5206. doi:10.1021/bi048095v. PMID:15794657.
- Hunt, J.F., Earnest, T.N., Bousche, O., Kalghatgi, K., Reilly, K., Horvath, C., Rothschild, K.J., and Engelman, D.M. 1997. A biophysical study of integral membrane protein folding. *Biochemistry*, **36**: 15156–15176. doi:10.1021/bi970146j. PMID:9398244.
- Hutchinson, E.G., and Thornton, J.M. 1996. PROMOTIF—a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**: 212–220. PMID:8745398.
- Hyberts, S.G., Goldberg, M.S., Havel, T.F., and Wagner, G. 1992. The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures. *Protein Sci.* **1**: 736–751. PMID:1304915.
- Jaroniec, C.P., Kaufman, J.D., Stahl, S.J., Viard, M., Blumenthal, R., Wingfield, P.T., and Bax, A. 2005. Structure and dynamics of micelle-associated human immunodeficiency virus gp41 fusion domain. *Biochemistry*, **44**: 16167–16180. doi:10.1021/bi051672a. PMID:16331977.
- Johansson, J., Szyperski, T., Curstedt, T., and Wuthrich, K. 1994. The NMR structure of the pulmonary surfactant-associated polypeptide SP-C in an apolar solvent contains a valyl-rich alpha-helix. *Biochemistry*, **33**: 6015–6023. doi:10.1021/bi00185a042. PMID:8180229.
- Kabsch, W. 1976. Solution for best rotation to relate 2 sets of vectors. *Acta Crystallogr. A*, **32**: 922–923. doi:10.1107/S0567739476001873.
- Kabsch, W., and Sander, C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**: 2577–2637. doi:10.1002/bip.360221211. PMID:6667333.
- Katragadda, M., Alderfer, J.L., and Yeagle, P.L. 2001a. Assembly of a polytopic membrane protein structure from the solution structures of overlapping peptide fragments of bacteriorhodopsin. *Biophys. J.* **81**: 1029–1036. PMID:11463644.
- Katragadda, M., Chopra, A., Bennett, M., Alderfer, J.L., Yeagle, P.L., and Albert, A.D. 2001b. Structures of the transmembrane helices of the G-protein coupled receptor, rhodopsin. *J. Pept. Res.* **58**: 79–89. doi:10.1034/j.1399-3011.2001.00904.x. PMID:11454172.
- Kelley, L.A., Gardner, S.P., and Sutcliffe, M.J. 1997. An automated approach for defining core atoms and domains in an ensemble of NMR-derived protein structures. *Protein Eng.* **10**: 737–741. PMID:9278289.
- Lamberth, S., Griesinger, C., Schmid, H., Carafoli, E., Muenchbach, M., Vorherr, T., and Krebs, J. 2000. NMR solution structure of phospholamban. *Helv. Chim. Acta*, **83**: 2141–2152. doi:10.1002/1522-2675(20000906)83:9<2141::AID-HLCA2141>3.0.CO;2-W.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. Procheck—a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**: 283–291. doi:10.1107/S0021889892009944.
- Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**: 477–486. PMID:9008363.
- Lindhout, D.A., Thiessen, A., Schieve, D., and Sykes, B.D. 2003. High-yield expression of isotopically labeled peptides for use in NMR studies. *Protein Sci.* **12**: 1786–1791. doi:10.1110/ps.0376003. PMID:12876327.
- Lugovskoy, A.A., Maslennikov, I.V., Utkin, Y.N., Tsetlin, V.I., Cohen, J.B., and Arseniev, A.S. 1998. Spatial structure of the M3 transmembrane segment of the nicotinic acetylcholine receptor alpha subunit. *Eur. J. Biochem.* **255**: 455–461. doi:10.1046/j.1432-1327.1998.2550455.x. PMID:9716388.
- Ma, D., Liu, Z., Li, L., Tang, P., and Xu, Y. 2005. Structure and dynamics of the second and third transmembrane domains of human glycine receptor. *Biochemistry*, **44**: 8790–8800. doi:10.1021/bi050256n. PMID:15952785.
- MacKenzie, K.R., Prestegard, J.H., and Engelman, D.M. 1997. A transmembrane helix dimer: structure and implications. *Science (Washington, D.C.)*, **276**: 131–133. doi:10.1126/science.276.5309.131. PMID:9082985.
- Marvin, D.A., Welsh, L.C., Symmons, M.F., Scott, W.R., and Straus, S.K. 2006. Molecular structure of fd (f1, M13) filamentous bacteriophage refined with respect to X-ray fibre diffraction and solid-state NMR data supports specific models of phage assembly at the bacterial membrane. *J. Mol. Biol.* **355**: 294–309. doi:10.1016/j.jmb.2005.10.048. PMID:16300790.
- Mascioni, A., Karim, C., Barany, G., Thomas, D.D., and Veglia, G. 2002. Structure and orientation of sarcolipin in lipid environments. *Biochemistry*, **41**: 475–482. doi:10.1021/bi011243m. PMID:11781085.
- Massiah, M.A., Ko, Y.H., Pedersen, P.L., and Mildvan, A.S. 1999. Cystic fibrosis transmembrane conductance regulator: solution structures of peptides based on the Phe508 region, the most common site of disease-causing DeltaF508 mutation. *Biochemistry*, **38**: 7453–7461. doi:10.1021/bi9903603. PMID:10360942.
- Naider, F., Khare, S., Arshava, B., Severino, B., Russo, J., and Becker, J.M. 2005. Synthetic peptides as probes for conformational preferences of domains of membrane receptors. *Biopolymers*, **80**: 199–213. doi:10.1002/bip.20183. PMID:15622547.
- Nakano, T., Ikegami, T., Suzuki, T., Yoshida, M., and Akutsu, H. 2006. A new solution structure of ATP synthase subunit c from thermophilic *Bacillus* PS3, suggesting a local conformational

- change for H⁺-translocation. *J. Mol. Biol.* **358**: 132–144. doi:10.1016/j.jmb.2006.01.011. PMID:16497328.
- Nishimura, K., Kim, S., Zhang, L., and Cross, T.A. 2002. The closed state of a H⁺ channel helical bundle combining precise orientational and distance restraints from solid state NMR. *Biochemistry*, **41**: 13170–13177. doi:10.1021/bi0262799. PMID:12403618.
- Oblatt-Montal, M., Reddy, G.L., Iwamoto, T., Tomich, J.M., and Montal, M. 1994. Identification of an ion channel-forming motif in the primary structure of CFTR, the cystic fibrosis chloride channel. *Proc. Natl. Acad. Sci. U.S.A.* **91**: 1495–1499. doi:10.1073/pnas.91.4.1495. PMID:7509074.
- Ohlenschläger, O., Hojo, H., Ramachandran, R., Gorkach, M., and Haris, P.I. 2002. Three-dimensional structure of the S4–S5 segment of the Shaker potassium channel. *Biophys. J.* **82**: 2995–3002. PMID:12023222.
- Op De Beeck, A., Montserret, R., Duvet, S., Cocquerel, L., Cacan, R., Barberot, B., et al. 2000. The transmembrane domains of hepatitis C virus envelope glycoproteins E1 and E2 play a major role in heterodimerization. *J. Biol. Chem.* **275**: 31428–31437. doi:10.1074/jbc.M003003200. PMID:10807921.
- Opella, S.J., Marassi, F.M., Gesell, J.J., Valente, A.P., Kim, Y., Oblatt-Montal, M., and Montal, M. 1999. Structures of the M2 channel-lining segments from nicotinic acetylcholine and NMDA receptors by NMR spectroscopy. *Nat. Struct. Biol.* **6**: 374–379. doi:10.1038/7610. PMID:10201407.
- Oxenoid, K., and Chou, J.J. 2005. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc. Natl. Acad. Sci. U.S.A.* **102**: 10870–10875. doi:10.1073/pnas.0504920102. PMID:16043693.
- Papadopoulos, E., Oglecka, K., Maler, L., Jarvet, J., Wright, P.E., Dyson, H.J., and Graslund, A. 2006. NMR solution structure of the peptide fragment 1–30, derived from unprocessed mouse Doppel protein, in DHPC micelles. *Biochemistry*, **45**: 159–166. doi:10.1021/bi051313f. PMID:16388591.
- Papavoine, C.H., Christiaans, B.E., Folmer, R.H., Konings, R.N., and Hilbers, C.W. 1998. Solution structure of the M13 major coat protein in detergent micelles: a basis for a model of phage assembly involving specific residues. *J. Mol. Biol.* **282**: 401–419. doi:10.1006/jmbi.1998.1860. PMID:9735296.
- Park, S.H., Mrse, A.A., Nevzorov, A.A., Mesleh, M.F., Oblatt-Montal, M., Montal, M., and Opella, S.J. 2003. Three-dimensional structure of the channel-forming trans-membrane domain of virus protein “u” (Vpu) from HIV-1. *J. Mol. Biol.* **333**: 409–424. doi:10.1016/j.jmb.2003.08.048. PMID:14529626.
- Pashkov, V.S., Maslennikov, I.V., Tchikin, L.D., Efremov, R.G., Ivanov, V.T., and Arseniev, A.S. 1999. Spatial structure of the M2 transmembrane segment of the nicotinic acetylcholine receptor alpha-subunit. *FEBS Lett.* **457**: 117–121. doi:10.1016/S0014-5793(99)01023-6. PMID:10486576.
- Penin, F., Brass, V., Appel, N., Ramboarina, S., Montserret, R., Ficheux, D., et al. 2004. Structure and function of the membrane anchor domain of hepatitis C virus nonstructural protein 5A. *J. Biol. Chem.* **279**: 40835–40843. doi:10.1074/jbc.M404761200. PMID:15247283.
- Pervushin, K.V., Orekhov, V., Popov, A.I., Musina, L., and Arseniev, A.S. 1994. Three-dimensional structure of (1–71)bacterioopsin solubilized in methanol/chloroform and SDS micelles determined by 15N–1H heteronuclear NMR spectroscopy. *Eur. J. Biochem.* **219**: 571–583. doi:10.1111/j.1432-1033.1994.tb19973.x. PMID:8307023.
- Poulsen, S.A., Watson, A.A., Fairlie, D.P., and Craik, D.J. 2000. Solution structures in aqueous SDS micelles of two amyloid beta peptides of A beta(1–28) mutated at the alpha-secretase cleavage site (K16E, K16F). *J. Struct. Biol.* **130**: 142–152. doi:10.1006/jsbi.2000.4267. PMID:10940222.
- Raman, P., Cherezov, V., and Caffrey, M. 2006. The membrane protein data bank. *Cell. Mol. Life Sci.* **63**: 36–51. doi:10.1007/s00018-005-5350-6. PMID:16314922.
- Rastogi, V.K., and Girvin, M.E. 1999. Structural changes linked to proton translocation by subunit c of the ATP synthase. *Nature (London)*, **402**: 263–268. PMID:10580496.
- Roosild, T.P., Greenwald, J., Vega, M., Castronovo, S., Riek, R., and Choe, S. 2005. NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science (Washington, D.C.)*, **307**: 1317–1321. doi:10.1126/science.1106392. PMID:15731457.
- Sapay, N., Montserret, R., Chipot, C., Brass, V., Moradpour, D., Deleage, G., and Penin, F. 2006. NMR structure and molecular dynamics of the in-plane membrane anchor of nonstructural protein 5A from bovine viral diarrhea virus. *Biochemistry*, **45**: 2221–2233. doi:10.1021/bi0517685. PMID:16475810.
- Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. 2003. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**: 65–73. doi:10.1016/S1090-7807(02)00014-9. PMID:12565051.
- Schwieters, C.D., Kuszewski, J.J., and Clore, G.M. 2006. Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**: 47–62. doi:10.1016/j.pnmrs.2005.10.001..
- Shenkarev, Z.O., Balashova, T.A., Efremov, R.G., Yakimenko, Z.A., Ovchinnikova, T.V., Raap, J., and Arseniev, A.S. 2002. Spatial structure of zervamicin IIB bound to DPC micelles: implications for voltage-gating. *Biophys. J.* **82**: 762–771. PMID:11806918.
- Slepko, E.R., Rainey, J.K., Li, X., Liu, Y., Cheng, F.J., Lindhout, D.A., Sykes, B.D., and Fliegel, L. 2005. Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na⁺/H⁺ exchanger. *J. Biol. Chem.* **280**: 17863–17872. PMID:15677483.
- Sorgen, P.L., Cahill, S.M., Krueger-Koplin, R.D., Krueger-Koplin, S.T., Schenck, C.C., and Girvin, M.E. 2002. Structure of the *Rhodobacter sphaeroides* light-harvesting 1 beta subunit in detergent micelles. *Biochemistry*, **41**: 31–41. doi:10.1021/bi011576j. PMID:11772000.
- Talafous, J., Marcinowski, K.J., Klopman, G., and Zagorski, M.G. 1994. Solution structure of residues 1–28 of the amyloid beta-peptide. *Biochemistry*, **33**: 7788–7796. doi:10.1021/bi00191a006. PMID:7516706.
- Valentine, K.G., Liu, S.F., Marassi, F.M., Veglia, G., Opella, S.J., Ding, F.X., et al. 2001. Structure and topology of a peptide segment of the 6th transmembrane domain of the *Saccharomyces cerevisiae* alpha-factor receptor in phospholipid bilayers. *Biopolymers*, **59**: 243–256. doi:10.1002/1097-0282(20011005)59:4<243::AID-BIP1021>3.0.CO;2-H. PMID:11473349.
- Wakabayashi, S., Pang, T., Su, X., and Shigekawa, M. 2000. A novel topology model of the human Na⁽⁺⁾/H⁽⁺⁾ exchanger isoform 1. *J. Biol. Chem.* **275**: 7942–7949. doi:10.1074/jbc.275.11.7942. PMID:10713111.
- Wang, J., Kim, S., Kovacs, F., and Cross, T.A. 2001. Structure of the transmembrane region of the M2 protein H⁽⁺⁾ channel. *Protein Sci.* **10**: 2241–2250. doi:10.1110/ps.17901. PMID:11604531.
- Wang, J.J., Hodges, R.S., and Sykes, B.D. 1995. Generating multiple conformations of flexible peptides in solution based on NMR nuclear Overhauser effect data—application to desmopressin. *J. Am. Chem. Soc.* **117**: 8627–8634. doi:10.1021/ja00138a019.
- Wang, Z.Y., Gokan, K., Kobayashi, M., and Nozawa, T. 2005. So-

- lution structures of the core light-harvesting alpha and beta polypeptides from *Rhodospirillum rubrum*: implications for the pigment-protein and protein-protein interactions. *J. Mol. Biol.* **347**: 465–477. doi:10.1016/j.jmb.2005.01.017. PMID:15740753.
- Watson, A.A., Fairlie, D.P., and Craik, D.J. 1998. Solution structure of methionine-oxidized amyloid beta-peptide (1–40). Does oxidation affect conformational switching? *Biochemistry*, **37**: 12700–12706. doi:10.1021/bi9810757. PMID:9737846.
- Wigley, W.C., Vijayakumar, S., Jones, J.D., Slaughter, C., and Thomas, P.J. 1998. Transmembrane domain of cystic fibrosis transmembrane conductance regulator: design, characterization, and secondary structure of synthetic peptides m1-m6. *Biochemistry*, **37**: 844–853. doi:10.1021/bi972293n. PMID:9454574.
- Williamson, M.P., and Waltho, J.P. 1992. Peptide structure from NMR. *Chem. Soc. Rev.* **21**: 227–236. doi:10.1039/cs9922100227.
- Wüthrich, K. 1986. *NMR of proteins and nucleic acids*. Wiley, New York.
- Yeagle, P.L., Choi, G., and Albert, A.D. 2001. Studies on the structure of the G-protein-coupled receptor rhodopsin including the putative G-protein binding site in unactivated and activated forms. *Biochemistry*, **40**: 11932–11937. doi:10.1021/bi015543f. PMID:11570894.
- Yeagle, P.L., Danis, C., Choi, G., Alderfer, J.L., and Albert, A.D. 2000. Three dimensional structure of the seventh transmembrane helical domain of the G-protein receptor, rhodopsin. *Mol. Vis.* **6**: 125–131. PMID:10930473.
- Yu, L., Sun, C., Song, D., Shen, J., Xu, N., Gunasekera, A., Hajduk, P.J., and Olejniczak, E.T. 2005. Nuclear magnetic resonance structural studies of a potassium channel-charybdotoxin complex. *Biochemistry*, **44**: 15834–15841. doi:10.1021/bi051656d. PMID:16313186.
- Yushmanov, V.E., Mandal, P.K., Liu, Z., Tang, P., and Xu, Y. 2003. NMR structure and backbone dynamics of the extended second transmembrane domain of the human neuronal glycine receptor alpha1 subunit. *Biochemistry*, **42**: 3989–3995. doi:10.1021/bi026767g. PMID:12667090.
- Zamoon, J., Mascioni, A., Thomas, D.D., and Veglia, G. 2003. NMR solution structure and topological orientation of monomeric phospholamban in dodecylphosphocholine micelles. *Biophys. J.* **85**: 2589–2598. PMID:14507721.
- Zirah, S., Kozin, S.A., Mazur, A.K., Blond, A., Cheminant, M., Segalas-Milazzo, I., Debey, P., and Rebuffat, S. 2006. Structural changes of region 1–16 of the Alzheimer disease amyloid beta-peptide upon zinc binding and in vitro aging. *J. Biol. Chem.* **281**: 2151–2161. PMID:16301322.