Opinion TRENDS in Biochemical Sciences Vol.31 No.4 April 2006

Has the code for protein translocation been broken?

Dalit Shental-Bechor*, Sarel J. Fleishman* and Nir Ben-Tal

Department of Biochemistry, George S. Wise Faculty of Life Sciences, Tel-Aviv University, 69978 Ramat Aviv, Israel

Polypeptides chains are segregated by the translocon channel into secreted or membrane-inserted proteins. Recent reports claim that an *in vivo* system has been used to break the 'amino acid code' used by translocons to make the determination of protein type (i.e. secreted or membrane-inserted). However, the experimental setup used in these studies could have confused the derivation of this code, in particular for polar amino acids. These residues are likely to undergo stabilizing interactions with other protein components in the experiment, shielding them from direct contact with the inhospitable membrane. Hence, it is our view that the 'code' for protein translocation has not yet been deciphered and that further experiments are required for teasing apart the various energetic factors contributing to protein translocation.

Introduction

Co-translational translocation is the process by which ribosomes that are attached to the endoplasmic reticulum (ER) extrude proteins through the translocon channel, giving rise to two different classes of proteins: those that are secreted or inserted into cellular membranes [1]. This crucial classification process is conducted on the basis of the sequence of the translated protein, which has led to the expectation that a 'sequence code' exists. If identified, this 'code' could be used to explain and predict which proteins would eventually reside within the membrane and which would be secreted into the ER lumen (and subsequently transported to various cellular compartments or expelled from the cell) [2]. This assumption was the foundation for several hydrophobicity scales (a ranking of the 20 amino acids according to their polarity), which were computed either from physical principles or from experiments that quantitatively compare the equilibrium distribution of amino acid residues in hydrophobic and hydrophilic media [3–8]. Such scales have been extremely useful, and have remained the principal means for identifying transmembrane (TM) segments in protein sequences for more than two decades [9].

An *in vivo* system for probing the energetics of translocation

Recently, Hessa *et al.* [10] carried out a series of experiments designed to decipher, for the first time, the translocation sequence code using an *in vivo* system containing ER membranes, ribosomes and the translocon channel – an approach that is far more realistic than the simple model systems previously employed. The experimental procedure is based on the use of an artificially designed variant of the leader peptidase (Lep) protein from *Escherichia coli*. This protein includes two endogenous TM helices (TM1 and TM2) and a soluble domain (P2) (Figure 1a). Using Lep as a host, an additional sequence segment (termed H) was engineered as a probe downstream of TM2 so that the equilibrium concentrations of the inserted versus the translocated H could be measured *in vitro* [10]. The procedure is attractive because of the clarity of the experimental readout, in addition to its simplicity, even though it addresses a highly complicated physiological system.

Hessa *et al.* [10] proceeded to read the 'sequence code' by translating the equilibrium concentrations of secreted and inserted Lep that contained various H probe segments into free-energy differences between the two states, inserted and translocated (Figure 1a). However, this treatment implicitly postulates two crucial, albeit unproven, thermodynamic assumptions: (i) that the H segment forms the same secondary structure, presumably an α-helix, in both the translocated and inserted states; and (ii) that the H segment is isolated from other protein components and contacts only the lipid molecules. Deviations from helicity or association with other protein components would mean that more thermodynamic states would need to be considered and that energetic contributions other than direct peptide–membrane interactions were involved, thus confounding the derivation of a hydrophobicity scale (Box 1).

In our opinion, the experimental setup used by Hessa *et al.* [10] cannot discriminate between the effects of the interactions of H with the membrane (hydrophobicity) and with TM1 and TM2, which are specific for the Lep host. Rather than the two thermodynamic states suggested by Hessa *et al.* [10] (Figure 1a), we believe that the H segment resides in at least five different states (Figure 1b), three corresponding to the classification of membrane-inserted [Figure 1b(i–iii)] and two to that of translocated [Figure 1b(iv–v)]. The inserted states differ from one another in the extent to which they expose the sidechains or backbone of their H segments to the lipid milieu. In the conformation in which the H segment is separate from the remainder of the TM domain [Figure 1b(i)], all of its sidechains are exposed to the lipid and its backbone is

Corresponding author: Ben-Tal, N. (nirb@tauex.tau.ac.il).
* Authors contributed equally.
Available online 10 March 2006

 0968-0004/$ - see front matter doi:10.1016/j.tibs.2006.02.002

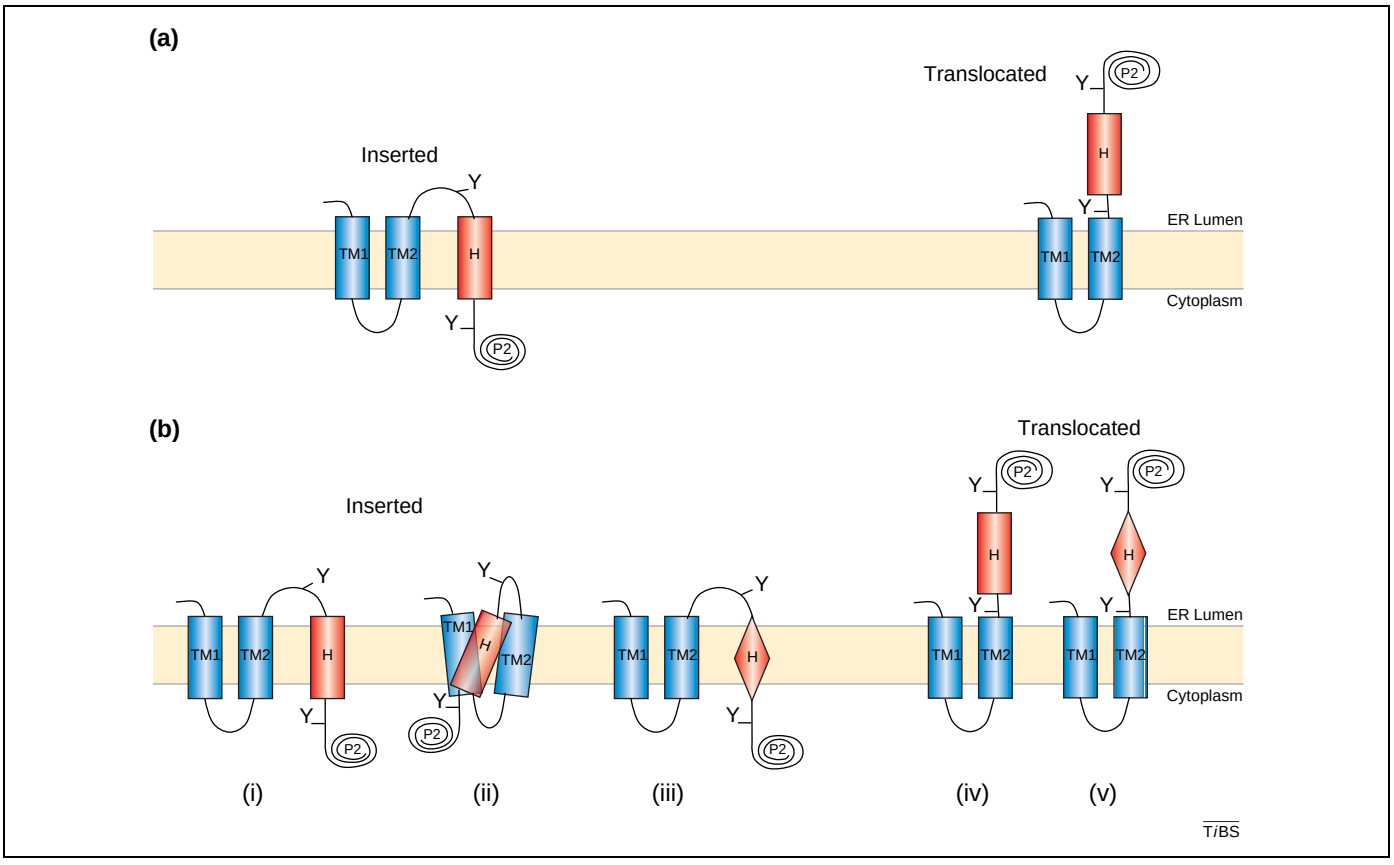

Figure 1. Schematic representation of the thermodynamic states associated with the *in vivo* system for probing the energetics of transloca tion. Hessa *et al.* [10] modified the leader peptidase (Lep) protein to include a probe H segment. Thus, the protein included two endogenous TM domains (TM1 and TM2), an extramembrane domain (P2) and the engineered H segment with two glycosylation sites on each end (Y). Glycosylation takes place only in the luminal side of the membrane, such that the inserted and translocated states can be differentiated from each other by the number of glycosylations that took place. **(a)** Following the suggestion of Hessa et al. [10], the translocation process can be described as a chemical equilibrium between two states – inserted and translocated. **(b)** Inserted Lep might assume a conformation in which the H segment is isolated from the remainder of the TM domain (i). However, other conformations, in which H is packed against TM1 and TM2 and interacts specifically with their sidechains and backbones (ii) or one in which H deviates from α-helicity [red rectangle in (i) versus diamond in (iii)] are also feasible. In the studies by Hessa *et al.* [10,16], it is impossible to distinguish between these three different thermodynamic states because all of them would be denoted as 'inserted Lep'. Similar confusion will arise between the two translocated states [(iv) and (v)], which differ from one another in the conformation of H.

maintained in an α-helical conformation. This is the only inserted state postulated by Hessa *et al.* [10]. But conformations in which H interacts with TM1 and TM2 shielding some of the sidechains of H from lipid [Figure 1b(ii)] or in which H deviates from α-helicity [Figure 1b(iii)], are also possible. Similarly, Hessa *et al.* [10] acknowledge the conformation in which an α-helical H segment is translocated [Figure 1b(iv)], but they ignore the fact that a non-helical conformation of the translocated H is also likely [Figure 1b(v)].

We perceive that the main flaw in the interpretation of experimental results by Hessa *et al.* [10] is that it is unlikely to apply to other TM proteins because the conformation in which H interacts directly with TM1 and TM2 [Figure 1b(ii)], which was not considered by Hessa *et al.*, is specific for Lep; therefore, the generality of the scale derived by Hessa *et al.* [10] requires substantiation. For example, among the inserted states, the conformation for membrane-inserted H considered by Hessa *et al.* [Figure 1b(i)] is likely to dominate in Lep variants with hydrophobic H segments because the interactions with the hydrophobic lipids are probably at least as favourable as the interactions with protein for such segments. However, as the polarity of H increases,

Box 1. The thermodynamic interpretation of the experimental results of Hessa *et al.*

Hessa *et al.* [10] engineered a probe H segment flanked by two *N*-linked glycosylation sites (see Figure 1 in the main text). Glycosylation took place only on the luminal side of the microsomes that were generated in the experiment, so the extent of Lep glycosylation is indicative of the state of the H segment. That is, inserted and translocated H are associated with singly and doubly glycosylated Lep, respectively. The proportion of singly (f_{1g}) and doubly (f_{2g}) glycosylated Lep were measured *in vitro* using SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) gels, and an 'apparent equilibrium constant' was assigned to their ratio according to equation 1.

$$K_{app} = \frac{f_{1g}}{f_{2g}} \qquad \text{(Eqn 1)}$$

The results were represented by converting K_{app} into apparent free energy according to the conventional thermodynamic definition: $\Delta G_{app} = -RT\ln K_{app}$, where R is the gas constant, T is the absolute temperature and ln is the natural logarithm. We contend that, in these experiments, H can be in an ensemble of at least five rather than two states (see Figure 1b in the main text) and that, therefore, it is overly simplistic to describe the equilibrium between inserted and translocated H using the ΔG_{app} formula used by Hessa *et al.* [10].

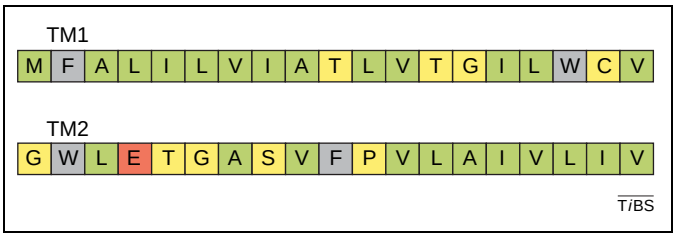

Figure 2. Amino acid sequences of the two transmembrane spans (TM1 and TM2) of Lep (SWISSPROT entry P00803). These segments contain several polar residues, which could form contacts with other polar residues on the probe H segment, shielding them from the hydrophobic membrane environment. Because the interactions of H with TM1 and TM2 are likely to be specific for Lep, they would alter its insertion propensity in a way that does not solely reveal the interactions of H with membrane and would, thus, limit the generality of the hydrophobicity scale derived by Hessa *et al.* [10]. Grey, aromatic residues; green, hydrophobic residues; yellow, small polar residues; red, negatively charged glutamic acid.

the conformation in which H interacts with TM1 and TM2 [Figure 1b(ii)] is likely to become the more populated state because of the known tendency of highly polar or charged residues (e.g. asparagine and glutamic acid) to drive the packing of their host helix against the polar backbone and sidechains of other TM helices [5,11–15]. The TM1 and TM2 segments of Lep contain one titratable (glutamic acid) and several residues that are small and polar (e.g. glycine and serine) (Figure 2), which would enable the other polar residues on the H segment to be shielded from the inhospitable membrane environment. If this conformation dominates [Figure 1b(ii)], then changing the position of the polar residues in the H segment would alter the stability of the protein in the membrane as a result of the interactions between H and the endogenous TM1 and TM2 of Lep. Indeed, the results reported by Hessa *et al.* [10,16] revealed that such positional dependence is observed for H segments that contain highly polar or charged residues. Moreover, the simple additivity of contributions to stability observed for apolar and mildly polar residues (e.g. leucine and serine, respectively) breaks down with the introduction of highly polar residues. One way to explain positional dependence and deviations from additivity, which is not refuted by Hessa *et al.*, is that the polar helices introduced into Lep as H segments formed stabilizing interactions with the endogenous TM helices of Lep. The expectation that the conformation in which H interacts with TM1 and TM2 [Figure 1b(ii)] dominates in H segments that contain charged residues might also explain the low values of ΔG_{app} penalties obtained by Hessa *et al.* [10] (Box 1) for the transfer of such residues from translocated to inserted states relative to other hydrophobicity scales [5,6,8]. Furthermore, although the original Lep protein is monomeric in membranes [17], it is not clear whether the Lep variants containing the more polar H segments oligomerize. If so, the equilibrium between inserted and secreted Lep would comprise many more states than the five suggested here (Figure 1b).

Membrane insertion of the S4 segment of voltage-gated K^+ channels

The limitations described here apply equally to the use of engineered Lep for studying the partitioning of natural TM segments. For instance, the same experimental

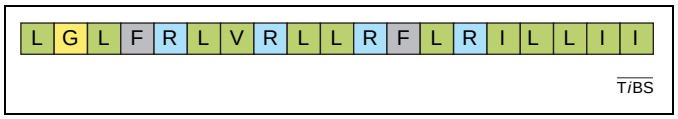

Figure 3. Amino acid sequence of the S4 segment of the voltage sensor of the KvAP channel. The sequence of S4 is mostly hydrophobic but four arginines (blue), at least three of which are charged, occupy conserved positions in the sequence. These charges could interact favourably with polar groups on TM1 and TM2 in a conformation schematically represented in Figure 1b(ii). Such interactions would stabilize the inserted conformation of S4, which would therefore not be exposed completely to membrane. Grey, aromatic residues; green, hydrophobic residues; yellow, small polar residues; blue, positively charged arginine residues.

framework has been used to study the free energy of transfer between translocated and inserted states of the S4 segment [16], which constitutes the core voltage-sensing element in the voltage-gated K^+ channel from *Aeropyrum pernix* (KvAP) [18]. This 19-residue segment is a hydrophobic cation consisting mostly of highly hydrophobic residues interspersed with four arginines at conserved positions (Figure 3), of which at least three are charged [19]. Until recently, it was anticipated – mainly on energetic grounds – that the S4 segment is packed against the other TM helices of the K^+ channel because of its high polarity [20]. By contrast, the first structure of the KvAP channel, which was crystallized in the absence of lipid, showed S4 to be exposed to the membrane [21]. This finding elicited considerable controversy. Hessa *et al.* [10] found that the net polarity of S4 was at the threshold that would enable its efficient insertion into the membrane, supporting the view that it is exposed to lipid.

To provide a thermodynamic explanation to the conclusion drawn by Hessa *et al.* [10] that isolated S4 could insert efficiently into membranes, Freites *et al.* [22] conducted molecular-dynamics simulations of an inserted α-helical segment bearing the S4 sequence, which was placed in a TM orientation and surrounded by lipid on all sides in isolation from other protein components. On the basis of these simulations, they suggested that the S4 segment is stabilized in a TM orientation despite its high polarity owing to contacts formed between the arginine sidechains on S4 and the phosphate headgroups and water molecules. These contacts can form according to the simulations because, in the immediate vicinity of S4, the thickness of the hydrocarbon core of the membrane shrinks from a steady-state width of ~30 Å [3] (Figure 4) to a mere 10 Å, which is considerably thinner than a lipid monolayer (Figure 4). In particular, Freites *et al.* [22] noted that one of the lipids assumes a conformation that spans the entire membrane in the region of S4, further demonstrating the enormous distortion of the membrane in these simulations. We note that a conservative estimate of the energetic penalty of such a large contraction of the membrane lipids that considers only the effects of dihedral-angle strains would be 12 kcal mol^{-1} [5,23]. Furthermore, a snapshot provided by Freites *et al.* [22] reveals that, for some of the lipids, the distortion is so large that their aliphatic chains form contacts with water molecules. Taken together, the strains to the aliphatic chains and the solvation penalty on direct contacts between polar and aliphatic groups that were observed

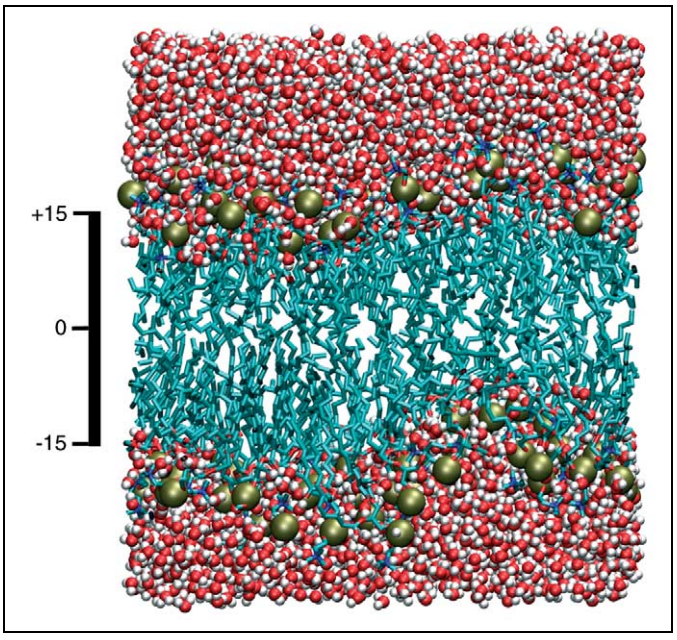

Figure 4. An equilibrated membrane bilayer composed of dimyristoylphosphocholine molecules embedded in water. Cyan, aliphatic chains; gold, phosphates in the headgroups; red, oxygen atoms in water. The bar shows the approximate span of the aliphatic chains in Ångstroms, with an approximation of the membrane midplane marked by 0. The hydrophobic core of each leaflet of the bilayer spans 15 Å. The membrane hydrophobic core fluctuates by a few Ångstroms around an equilibrium width of 30 Å.

in these simulations make membrane contraction an unlikely explanation for efficient insertion of S4.

The suggestion that a TM orientation for S4 is stabilized by such large-scale membrane distortion is made even more implausible in view of biophysical studies of model TM segments [5,24]. The experimental data collected on many different peptides show that the membrane width might decrease by several Ångstroms to better match the hydrophobic length of a peptide, but that peptides with hydrophobic lengths that are considerably shorter than the width of the hydrocarbon region of the bilayer (as in S4) do not partition into the membrane but, instead, reside on the membrane surface. Although S4 was not directly targeted by these experiments, the data provide an indication that the membrane would not undergo contraction of 20 Å (as suggested by Freites *et al.* [22]) to stabilize the inserted conformation of S4. Similar to our criticism of the results on the membrane insertion of polar residues [10], a more likely explanation for the observed tendency of S4 towards membrane insertion in the experimental setup of Hessa *et al.* is that S4 forms stabilizing contacts with the endogenous TM1 and TM2 of Lep, and that lipid contraction has a much smaller role than suggested by the simulations of Freites *et al.* [22]. In this respect, it is notable that a more recent structure of the voltage-gated K^+ channel, which was crystallized in the presence of lipids and is therefore considered a more faithful representation of the physiological structure than the previous crystal structure, showed that two of the four arginines in S4 are buried at a helix–helix interface where they are partially shielded from the inhospitable lipid environment [25].

Concluding remarks

We have raised several points that question the validity of deriving thermodynamic quantities for the interactions between amino acids and peptides with membranes using the *in vivo* system of Hessa *et al.* [10]. The key problems are the lack of experimental controls for the α-helicity of the H segment and, most importantly, whether H associates directly with membranes without forming stabilizing contacts with other protein components in the system. To address these issues, it still needs to be shown that: (i) physical contacts are not formed between H and the two endogenous TM segments of Lep and that engineered Lep does not oligomerize (this could be achieved using fluorescence-labelling techniques [26], for example); and (ii) that the H segment retains its secondary structure (presumably an α-helix) in both the inserted and the translocated states for substitutions of all 20 amino acids. The generality of the scale derived from the results of the study [10] could also be validated by using hosts other than the Lep protein to observe that the same thermodynamic quantities are obtained. The partitioning of isolated S4 segment from KvAP could also be monitored using solid-state NMR [24].

It should be appreciated that Hessa *et al.* [10] have focused on a highly complicated physiological system comprising a plethora of different proteins, which is subject to a complex environment that includes water, protein and membrane. The experimental setup devised by these authors certainly represents a major step towards probing the energetics of protein translocation within a physiologically relevant framework, and could prove useful in future studies of protein–protein interactions within the membrane. However, the readout from this system probably reflects a mixture of hydrophobicity and various contributions stemming from interactions with the host protein Lep, especially for the more polar segments tested. Building on this setup, more experiments will be needed before it can be safely concluded that 'the fundamental code used by the translocon to select polypeptide segments for insertion as TM helices has been broken' [27].

Acknowledgements

We thank I.T. Arkin for kindly providing Figure 4. We also thank I.T. Arkin, J.U. Bowie, J.A. Hirsch, Y. Ofran and O. Yifrach for critical reading of this article. This study was supported by grant 222/04 from the Israel Science Foundation. S.J.F. was supported by a doctoral fellowship from the Clore Israel Foundation.

References

1 Osborne, A.R. *et al.* (2005) protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell Dev. Biol.* 21, 529–550
2 Bowie, J.U. (2005) Cell biology: border crossing. *Nature* 433, 367–369
3 White, S.H. and Wimley, W.C. (1999) Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* 28, 319–365
4 Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157, 105–132
5 Kessel, A. and Ben-Tal, N. (2002) Free energy determinants of peptide association with lipid bilayers. In *Current topics in membranes* (Vol. 52) (Simon, S. and McIntosh, T., eds), pp. 205–253, Academic Press

6 Engelman, D.M. *et al*. (1986) Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* 15, 321–353
7 Eisenberg, D. *et al*. (1986) Hydrophobicity and amphiphilicity in protein structure. *J. Cell. Biochem.* 31, 11–17
8 von Heijne, G. (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* 225, 487–494
9 Chen, C.P. *et al*. (2002) Transmembrane helix predictions revisited. *Protein Sci.* 11, 2774–2791
10 Hessa, T. *et al*. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* 433, 377–381
11 Stevens, T.J. and Arkin, I.T. (1999) Are membrane proteins 'inside-out' proteins? *Proteins* 36, 135–143
12 Fleishman, S.J. and Ben-Tal, N. (2002) A novel scoring function for predicting the conformations of tightly packed pairs of transmembrane α-helices. *J. Mol. Biol.* 321, 363–378
13 Faham, S. *et al*. (2004) Side-chain contributions to membrane protein structure and stability. *J. Mol. Biol.* 335, 297–305
14 Curran, A.R. and Engelman, D.M. (2003) Sequence motifs, polar interactions and conformational changes in helical membrane proteins. *Curr. Opin. Struct. Biol.* 13, 412–417
15 Sternberg, M.J. and Gullick, W.J. (1989) Neu receptor dimerization. *Nature* 339, 587
16 Hessa, T. *et al*. (2005) Membrane insertion of a potassium-channel voltage sensor. *Science* 307, 1427
17 Zwizinski, C. *et al*. (1981) Leader peptidase is found in both the inner and outer membranes of *Escherichia coli*. *J. Biol. Chem.* 256, 3593–3597
18 Ruta, V. *et al*. (2003) Functional analysis of an archaebacterial voltage-dependent K^+ channel. *Nature* 422, 180–185
19 Schoppa, N.E. *et al*. (1992) The size of gating charge in wild-type and mutant Shaker potassium channels. *Science* 255, 1712–1715
20 Miller, C. (2003) A charged view of voltage-gated ion channels. *Nat. Struct. Biol.* 10, 422–424
21 Jiang, Y. *et al*. (2003) X-ray structure of a voltage-dependent K^+ channel. *Nature* 423, 33–41
22 Freites, J.A. *et al*. (2005) Interface connections of a transmembrane voltage sensor. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15059–15064
23 Fattal, D.R. and Ben-Shaul, A. (1993) A molecular model for lipid-protein interaction in membranes: the role of hydrophobic mismatch. *Biophys. J.* 65, 1795–1809
24 Bechinger, B. (2000) Understanding peptide interactions with the lipid bilayer: a guide to membrane protein engineering. *Curr. Opin. Chem. Biol.* 4, 639–644
25 Long, S.B. *et al*. (2005) Crystal structure of a mammalian voltage-dependent Shaker family K+ channel. *Science* 309, 897–903
26 Wallrabe, H. and Periasamy, A. (2005) Imaging protein molecules using FRET and FLIM microscopy. *Curr. Opin. Biotechnol.* 16, 19–27
27 White, S.H. and von Heijne, G. (2005) Transmembrane helices before, during, and after insertion. *Curr. Opin. Struct. Biol.* 15, 378–386