Biochemistry

 Volume 29, Number 17 May 1, 1990

Perspectives in Biochemistry

Membrane Protein Folding and Oligomerization: The Two-Stage Model

J.-L. Popot*,‡ and D. M. Engelman*,§

Institut de Biologie Physico-Chimique and Collège de France, 13 rue Pierre et Marie Curie, F-75005 Paris, France, and Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, New Haven, Connecticut 06511

Received October 30, 1989

ABSTRACT: We discuss the view that the folding of many, perhaps most, integral membrane proteins can be considered as a two-stage process. In stage I, hydrophobic α-helices are established across the lipid bilayer. In stage II, they interact to form functional transmembrane structures. This model is suggested by the nature of transmembrane segments in known structures, refolding experiments, the assembly of integral membrane protein from fragments, and the existence of very small integral membrane protein subunits. It may extend to proteins with a variety of functions, including the formation of transmembrane aqueous channels. The model is discussed in the context of the forces involved in membrane protein folding and the interpretation of sequence data.

In the following, we examine the suggestion that the folding of many integral membrane proteins can be understood in terms of two energetically distinct stages. In the first stage, independently stable helices are formed across the hydrophobic region of the membrane lipid bilayer. In the second, the helices interact with one another to give a functional, globular membrane protein. Similar helix–helix interactions participate in the stabilization of membrane protein oligomers.

Integral membrane proteins are partially buried in the nonpolar environment of the lipid bilayer, where the hydrophobic effect is absent and intrachain hydrogen bonds take on a much greater significance than in water, since the lipid solvent is unable to form them. This energy balance led to the expectation that the transmembrane region of membrane proteins would consist of predominantly hydrophobic segments with regular secondary structure and, more specifically, of bundles of hydrophobic α-helices (Henderson, 1975, 1977). This expectation has been largely fulfilled, although some exceptions do exist (see below).

Methods for predicting transmembrane helical regions from amino acid sequence data rely on scanning the sequence for stretches of residues long enough to span the nonpolar region of a lipid bilayer as α-helices and hydrophobic enough to be expected to be at lower free energy across a membrane than in an aqueous environment [reviewed by Engelman et al. (1986)]. Only two structures are known with enough certainty to provide a test of these predictions: bacterial photosynthetic reaction centers and bacteriorhodopsin. The existence and approximate position in the sequence of the 18 transmembrane helices in these two structures are well predicted by a simple hydrophobicity analysis (Michel et al., 1986a; Engelman et al., 1982; Ovchinnikov et al., 1985). This success is all the more striking given that the helices involved, predicted to be stable as independent entities in the bilayer, actually have more contact with other helices and pigments than they do with lipids. Therefore, we propose a conceptual division of the process of folding of these molecules into factors giving rise to the transmembrane helices and those determining their assembly into the final, folded structure. In this two-stage model (Figure 1), transmembrane helices are regarded as autonomous folding domains.

Similarly, the formation of oligomeric complexes of proteins in membranes will frequently involve the side to side interaction of transmembrane helices, as is seen in the oligomer of the photosynthetic reaction center and in the trimeric association of bacteriorhodopsin molecules. Association of single transmembrane helices is encountered in complexes of the photosynthetic and respiratory chains [see Popot and de Vitry (1990)] and may play a role in the dimerization of some anchored proteins (Bormann et al., 1989). In these instances, the packing of helices at subunit interfaces is like that within the subunits themselves. Thus, understanding helix–helix interactions in bilayers may clarify both folding of individual

‡ Institut de Biologie Physico-Chimique.
§ Yale University.

0006-2960/90/0429-4031$02.50/0

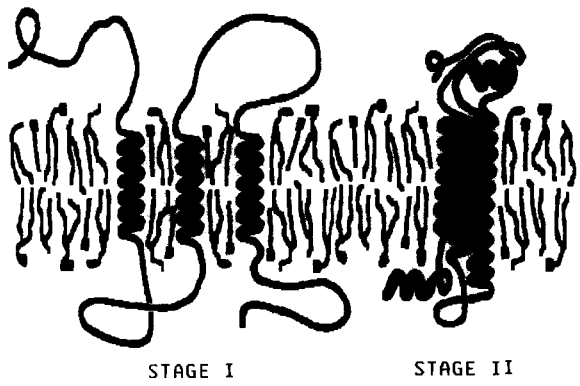

FIGURE 1: Two-stage model for the folding of α-helical integral membrane protein. The first stage is the formation of independently stable trans-bilayer helices, principally in response to the hydrophobic effect and the formation of main-chain hydrogen bonds in the nonaqueous environment. The second stage is the interaction of the helices to form the tertiary fold of the polypeptide. Factors that could contribute to the energetics of stage II are the links between helices, packing of helices and lipid molecules, polar interactions between helices, and, when applicable, association with prosthetic groups or with other proteins.

molecules and factors involved in oligomerization.

The two-stage model posits that the final structure in the transmembrane region results from the accretion of smaller elements (helices), each of which has reached thermodynamic equilibrium with the lipid and aqueous phases before packing. Some rearrangements are to be expected upon assembly (see below), but some others, like flipping helices through the membrane or inserting new segments, are considered as kinetically forbidden. It is therefore implicit in the model that the final structure may be the one with the lowest free energy according to these constraints, but not necessarily that with the lowest overall free energy. This interplay of thermodynamic and kinetic factors has been recently discussed by Finkelstein and Ptitsyn (1987) in the context of a model in which soluble globular proteins fold by packing preformed secondary structure elements. The major difference between soluble and membrane proteins, in this respect, is that transmembrane α-helices in lipids are much more stable than secondary structure elements of the same length in water and more akin to autonomous folding domains. In the same way as folding domains do not unfold and refold when they associate, the two-stage model excludes structures incorporating transmembrane segments that would not be individually stable. While a number of other circumstances can be imagined, such as the stabilization of an otherwise unstable helix or extended segment through a hairpin link to a stable helix (Engelman & Steitz, 1981), it seems that the simpler view may suffice in most of the known cases.

In the following paragraphs, we discuss helix formation in the bilayer and helix–helix associations. We have deliberately restricted ourselves to a consideration of the energetics of polypeptides inserted into a lipid bilayer, leaving aside the mechanism of insertion itself. One of the leading contentions of this discussion, indeed, is that many aspects of integral membrane protein folding can be understood without reference to the actual way insertion is achieved.

Hydrophobic Helices in Lipid Bilayers Are Highly Stable Structures

Helical structure is known to be induced in polypeptides in nonaqueous environments (Singer, 1962, 1971). The large free energy cost of transferring an unsatisfied hydrogen-bond donor or acceptor from an aqueous to a nonpolar environment or of

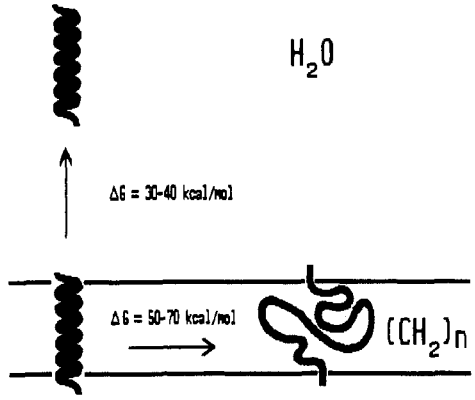

FIGURE 2: Stability of a 20-residue polyalanine transmembrane α-helix. Estimates are indicated of the free energy cost of removing the helix from the lipid phase (top) or breaking all of the hydrogen bonds (right). Modified from Engelman and Steitz (1981).

breaking such a bond in a nonpolar environment suggests that most hydrogen bonds must be satisfied when peptides are inserted into a membrane environment. α-Helices accomplish this in a systematic fashion that links nearby parts of the chain. If we consider a polypeptide with nonpolar side chains traversing a membrane bilayer, its removal from the bilayer is opposed by the hydrophobic effect and amounts to tens of kilocalories. If the peptide is unfolded within the nonpolar environment, breaking all of the hydrogen bonds, even more energy is required (Figure 2).

Thus, the nonpolar transbilayer helix is a very stable structure—so stable, in fact, that it can accommodate several polar groups without becoming unstable. Inserting polar groups into the bilayer may be required for function, as when histidine residues serve as ligands for chlorophylls or hemes. The term "hydrophobic" therefore can be ambiguous, inasmuch as many hydrophobic α-helices may contain a number of polar residues that may also make them "amphipathic". Several different scales and approaches for identifying transmembrane helices on the basis of their hydrophobicity have been developed and appear to be largely successful. In many cases, indeed, the local hydrophobicity is so high that putative transbilayer helices can be identified by inspection. In other cases, it may be impossible to decide on the basis of hydrophobicity alone whether a given segment is more likely to span the membrane or to belong to the hydrophobic core of an extramembrane domain [cf. Eisenberg (1984), Klein et al. (1985), Engelman et al. (1986), and Popot and de Vitry (1990)]. Throughout the present paper, we mean by "hydrophobic α-helices" helices that are hydrophobic enough to partition into lipid bilayers with a strongly favorable free energy change.

If transbilayer helices are the stable building blocks of many membrane proteins and complexes, at least two interesting issues arise. If hydrogen bonding of the main chain and the hydrophobic effect are not important factors in helix–helix association (as they are already taken into account in the formation of the helices themselves), how are the folded structures of the membrane proteins maintained? Further, ion channels contain aqueous transmembrane paths—how can such paths be provided by largely nonpolar helices?

Factors Governing Helix–Helix Association

A membrane protein formed from independently stable helices must be held together by interactions that permit the detailed close packing of the helices, overcoming the entropy that favors helix separation (in the range of 1–10 kcal/mol for a pair of helices). Possibilities for such interactions include

hydrogen bonds, ion pairs, the interaction of helix dipoles, packing differentials involving lipid/protein interactions, interactions with prosthetic groups, and external constraints such as links between the helices or interactions with other molecules.

It has been pointed out that external structures of the subunits may stabilize the transmembrane helix bundle in the photosynthetic reaction center (Yeates et al., 1987). Extramembrane regions may play an important role in helix assembly in proteins such as the Na^+/K^+-ATPase α-subunit or the subunits of the nicotinic acetylcholine receptor, where large extramembrane regions appear to separate putative transmembrane helices (see below). It is of interest to note in this context that the size of a stable extramembrane protein domain may be somewhat less than for globular soluble proteins, since constraint of the polypeptide ends will favor folding. Conversely, links between helices will limit their possible diffusion in the plane of the lipid bilayer so that the unfavorable entropic term in the free energy of association will be reduced. These two entropic effects are analogous to those of disulfide bridges in soluble proteins (Flory, 1956; Pace et al., 1988).

Hydrogen bonds and ion pairs can be used to drive the association of local regions of a pair of helices. Van der Waals interactions would then promote detailed close packing and further stabilize the association. As has been noted, the strength of a hydrogen bond in a nonaqueous milieu is in the range of 4–6 kcal/mol (Allen, 1975). Thus, a single hydrogen bond could, in principle, match the entropic term. Ion pairing or strong hydrogen bonding could provide even larger association energies; these would depend very much on the local environment of the ion pair [see Honig et al. (1986)]. While the detailed structure of bacteriorhodopsin is not yet known, early indications that the interior of the protein might be more polar than its lipid-exposed surface (Engelman & Zaccai, 1980) are consistent with results from hydrophobic labeling (Brunner et al., 1985) and site-directed mutagenesis (Khorana, 1988) and with the orientation of the two sequence segments whose arrangement into the structure has been established with some certainty (Popot et al., 1989). The photosynthetic reaction center transmembrane region features no interhelical salt bridges and few hydrogen bonds (Yeates et al., 1987), but the helices are oriented with their more polar surfaces toward the interior of the complex (Rees et al., 1989).

Interactions with prosthetic groups also provide packing constraints. These will be particularly strong when liganding involves residues belonging to several helices. Such is the case in the photoreaction centers (Michel et al., 1986b) and probably in most of the complexes of the respiratory and photosynthetic chains.

Because α-helices have polar ends arising from charge separation in their peptide linkages ("dipoles"), it has been proposed that antiparallel association between nearest-neighbor helices is energetically favored compared with a parallel association [see Hol (1985) and references cited therein]. As noted by Yeates et al. (1987), the photosynthetic reaction center contains mainly associations of antiparallel helices; the most likely folding model for bacteriorhodopsin also involves mostly antiparallel interactions [see Popot et al. (1989)]. The weakening influences of solvent dielectric ions and counterions at the helix ends, however, can be important (Rogers & Sternberg, 1984; Honig et al., 1986; Gilson & Honig, 1989); parallel helices are actually found in several instances. Stabilization by helix–dipole interactions may be weak in bacteriorhodopsin, which has only short extramembrane loops and therefore fairly exposed helix ends. In the photoreaction center from *Rhodopseudomonas viridis*, helix–dipole interactions might be reinforced when the cytochrome and the H subunit bind to the core and screen the two ends of the α-helix bundle from the solvent.

The contribution of packing effects deserves close examination. It may be difficult for lipid chains to pack well against the surface of an α-helix given the irregular contour presented by protruding side chains. As helices are known to engage in detailed close packing (Yeates et al., 1987; Richards, 1977), a less favorable packing at the helix–lipid interface should act to favor association of helices and association of lipids in separate regions. Further, the restriction of lipid-chain conformations in the vicinity of a comparatively rigid helix may give rise to some entropic preference of lipid molecules to be next to other lipids. The magnitudes of such contributions are difficult to calculate, but a rough estimate can be made for the effect of close packing. The free energy cost of creating cavities in proteins has been estimated by Rashin et al. (1986) to be 60 cal·mol^{-1}·Å^{-3} (within a factor of 2 or so). On this basis, filling a 30–80-Å^3 cavity (the size of a small side chain) could provide a large portion of the free energy required to associate two helices. Destabilizing free energy changes of comparable magnitude have been recently determined after voids were introduced genetically in the interior of a soluble enzyme (Kellis et al., 1988).

Thus, it is seen that a variety of possible interactions may drive helix association. It is likely that each of the factors considered above will prove to play a role but that a different balance will exist in different cases, e.g., depending on the protein's function [cf. Michel et al. (1986a)].

An Experimental Study of Factors in Folding: Bacteriorhodopsin

Strong experimental tests of the two-stage folding model have come from studies of the refolding of proteolytically cleaved and denatured bacteriorhodopsin. Bacteriorhodopsin (BR) consists of a bundle of seven transmembrane helices that surround the retinal prosthetic group (Henderson & Unwin, 1975; Kouyama et al., 1981; Jubb et al., 1984). It is the only membrane protein to have been wholly renatured starting from the completely unfolded polypeptide as well as from two denatured proteolytic fragments (Huang et al., 1981; Liao et al., 1983). The fragments are produced by a single chymotryptic cleavage between the second and third of the seven transmembrane segments. A covalent link between them is not required for the molecule to refold properly. Further, it has proven possible to refold the BR fragments in lipid bilayers under conditions such that the formation of a two-dimensional lattice can subsequently be induced (Popot et al., 1987). The structure of the renatured molecules can be studied crystallographically and is shown to be indistinguishable from that of native BR (Popot et al., 1986). These observations bring further strong support to the contention that the native structure of BR lies at a free energy minimum (Huang et al., 1981). This is an important conclusion, as the anisotropic environment of membrane proteins and the intrinsic asymmetry of their mode of insertion during synthesis make it possible that a native structure might be biosynthetically trapped at a state of higher energy that could not be reached during refolding in vitro.

When the two BR fragments are separately refolded into lipid vesicles, they regain helical structure. Upon fusion of such vesicles, the two fragments interact and bind retinal. The resulting complex has the visible absorption spectrum of native BR (Popot et al., 1987). The kinetics of chromophore regeneration are the same whether retinal is added to native

(chymotryptically cleaved) apoprotein or to the complex formed by reassociating prerefolded fragments. All steps occur in bilayers of *Halobacterium* lipids, in an environment that must be similar to that of the natural plasma membrane.

These observations indicate that, while the isolated fragments have been refolded under conditions vastly different from those prevailing during natural folding in vivo, their structure as isolated entities is close to that which they adopt in native BR. This permits recognition and correct reassociation to occur under conditions (absence of detergent) where major reorientations, e.g., flipping of helices through the bilayer, must be kinetically forbidden. It seems probable that, despite the absence of the rest of the molecule, the correct number of α-helices per fragment must form, that they must span the bilayer of the reconstituted vesicles, and that they must be close to their correct position in the sequence. The first point is directly substantiated by ultraviolet–circular dichroism spectra, which indicate each fragment to have recovered a highly α-helical structure. The other two points are currently under investigation. The refolded fragments in lipid vesicles are highly stable and their folded state must correspond to a deep free energy minimum.

The outcome of these experiments is such as would be expected from the two-stage model, where each transmembrane helix behaves as an independent folding domain. In this case, the link between the second and third helices and the presence of the prosthetic group are not required for fragment assembly to occur. The implication is that the remaining sources of helix–helix interactions, polar forces and packing effects, must dominate.

As mentioned before, at the level of the bilayer there is not much difference between packing helices within a subunit and packing subunits into an oligomer. This is illustrated in vitro by the above experiments, in which refolded BR fragments behave as the two subunits of a heterodimer. Equivalent cases are observed in vivo, either naturally or as the result of genetic experiments, and produce split or "microassembled" proteins that are functional; integral protein complexes in the inner membranes of chloroplasts and mitochondria contain very small subunits apparently comprised of a single transmembrane α-helix and little more [reviewed by de Vitry and Popot (1989) and Popot and de Vitry (1990)].

The formation of separately stable domains by individual transmembrane helices may be related to the fact that most introns in genes coding for polytopic proteins are located in the loops between helices (Jennings, 1989; P. Slonimski, personal communication). Transmembrane helices may have been recombined by exon shuffling in the course of evolution, in response to the need for new or modified functions.

Structures Involved in Transmembrane Aqueous Channels

Aqueous channels present a challenge to the idea that bundles of individually stable helices may form most transmembrane protein structures. It has been thought that channels must have strongly polar linings and that such linings cannot be formed from the weakly polar surfaces found in stable transmembrane helices. The nicotinic acetylcholine receptor (nAChR) provides an instructive example.

Several folding models of the nAChR subunits have been advocated. Most feature either four mainly hydrophobic transmembrane helices ("four-helix" model) or five transmembrane segments, one or two of them carrying numerous charges [reviewed by Popot and Changeux (1984), Hucho (1986), Numa (1989), and Changeux (1990)]. Immunological evidence in favor of models with a charged channel lining and an odd number of transmembrane segments has been presented [see Ratnam et al. (1986b) and references cited therein]. Recent experiments, however, have tended not to support these models. Namely, (i) biochemical data favor an even number of transmembrane segments (McCrea et al., 1986, 1987; DiPaola et al., 1989); (ii) the various charged segments proposed as candidates for the channel lining are probably external to the membrane (Dennis et al., 1988; Dwyer, 1988; Atassi et al., 1987); (iii) electrophysiological measurements indicate that, in its narrowest part, the nAChR channel is uncharged (Dani & Eisenman, 1987); (iv) biochemical (Giraudat et al., 1986, 1987, 1989; Hucho et al., 1986; Oberthür et al., 1986; Revah et al., 1990) and genetic evidence (Imoto et al., 1986, 1988; Leonard et al., 1988) suggests that each subunit contributes the second of its four hydrophobic helices toward lining the channel.

While no folding model of the nAChR subunits can be considered as established, it seems therefore probable that their transmembrane region (and, presumably, that of all chemically gated channels sequenced to date; Grenningloh et al., 1987a,b; Schofield et al., 1987; Hollmann et al., 1989; Wada et al., 1989; Gregor et al., 1989) comprises only hydrophobic α-helices. A simple hypothesis is that the two-stage model applies, but with an additional level of assembly: following folding of the subunits, they would pack without a drastic change in their transmembrane region to form the oligomeric channel. It has often been argued that in large membrane proteins or in channel-forming proteins some sequence segments are not in contact with lipids, and therefore need be neither α-helical, nor hydrophobic, nor long enough to span the bilayer, and that restricting topological models to long hydrophobic transmembrane α-helices would be misleading. In the case of the nAChR, such hypotheses do not seem warranted.

Two differences between BR and the nAChR subunits deserve comment. First, the transmembrane region of the nAChR represents only about one-fourth of its residues, rather than three-fourths in bacteriorhodopsin. From the amino terminus on, the four-helix model of the subunits features a large extracellular region, followed by three closely spaced hydrophobic helices, a cytoplasmic region, and a carboxy-terminal hydrophobic helix. Unlike BR, the extramembrane regions are large enough to contain folding domains of their own (they comprise about 200 and 100 residues, respectively).

A second and important difference is the formation of the channel. When BR monomers pack into trimers, they trap between them half a dozen lipid molecules (Glaeser et al., 1985). Would not we expect the same thing to occur upon assembly of the nAChR subunits, resulting in a "channel" filled with lipids? An answer might lie in the channel's dimensions. While it is known to be blocked by many amphipathic compounds [see Popot and Changeux (1984)], its narrowest width is estimated to be 6–7 Å [Maeno et al., 1977; Dwyer et al., 1980; see also Furois-Corbin & Pullman (1989) and references cited therein]. This is too small to accommodate a phospholipid molecule (but close to the dimension of cholesterol). The overall profile of the channel is not known (although the narrow section has been shown to be short; Dani & Eisenman, 1987; Dani, 1989), but narrowness or poor steric adaptation might exclude membrane lipids. It is worth recalling that ca. 22-residue-long hydrophobic peptides inserted in lipid bilayers do manage to open aqueous channels, which are thought to result from aggregation of independently stable transmembrane α-helices [see, e.g., Molle et al. (1988), Lear et al. (1988), and Oiki et al. (1988)].

To form a stable water-filled channel from fully folded nAChR subunits in a membrane requires that the unfavorable energy of creating a largely hydrophobic surface in contact with water be compensated by the energy of other interactions. Simple calculations show that this can easily be the case. While the actual geometry of the nAChR channel is not known, reasonable assumptions about the area and composition of its lining give an upper bound of 40 kcal/mol as the likely penalty for creating it. Dissociation constants for oligomeric proteins are usually less than 10^{-8}, corresponding to a free energy of dissociation of more than 10 kcal per interface [see, e.g., Chothia and Janin (1975)]. As the receptor has large extramembrane domains, the association of the five subunits could more than compensate for the energy cost of forming the channel, even if the additional contribution from subunit–subunit interactions in the lipid phase is disregarded. The structure of the tryptophan synthase complex (Hyde et al., 1988) shows the presence of a largely hydrophobic tunnel 25 Å long in the oligomeric enzyme complex. The tunnel seems filled with solvent and is large enough for an indole molecule (ca. 7 Å wide) to pass. This striking structural finding provides a clear basis for the contention that oligomeric association energies can create an aqueous channel through a hydrophobic region.

Other Structures

Space limitations prevent a detailed consideration of other structures. It should be noted, however, that, while the hydrophobic helix bundle might be the most common sort of transmembrane region, it is not the only one. Porins, which open aqueous channels in the outer membranes of Gram-negative bacteria, are neither α-helical nor markedly hydrophobic. Rather than being a cluster of individually stable helices, they are thought to be made up of an assembly of β-sheets stabilized by a dense network of hydrogen bonds [see, e.g., Kleffel et al. (1985), Nabedryk et al. (1988), and references cited therein]. Porins are the only integral membrane proteins for which the presence of nonhelical transmembrane secondary structure is well established. Their peculiar structure might be related to difficulties in exporting proteins containing hydrophobic sequence segments to the outer membrane [cf. Davis and Model (1985), MacIntyre et al. (1988), and Popot and de Vitry (1990)]. As individual β-strands, or pairs of β-strands, leave many hydrogen bonds unsatisfied, some concerted mechanism of insertion must be expected.

Voltage-gated channels such as the Na^+, K^+, and Ca^{2+} channels present a special problem, as their structure must include a charged sensor of transmembrane voltage [for a discussion on structural requirements for sensors, see Honig et al. (1986)]. Comparative examination of their sequences suggests that the sensor function could be accomplished by transmembrane α-helices including one positively charged residue every turn in an otherwise hydrophobic segment (Noda et al., 1984; Tempel et al., 1987; Tanabe et al., 1987; Numa, 1989). The rest of the transmembrane region would be made up of numerous hydrophobic helices. If these structural premises are correct, these proteins contain helices that would not be expected to spontaneously partition into the lipid phase. While the protein as a whole might lie at its free energy minimum, some of the transmembrane segments would not. This again would require a concerted mechanism of insertion. For instance, the insertion of a polar helix could be stabilized by interaction with a very hydrophobic helix, as has been proposed in the helical hairpin hypothesis (Engelman & Steitz, 1981).

Model Building

When one attempts to build structural models from a sequence, the safest approach is probably to restrict postulates to hydrophobic transmembrane segments, unless structural and functional data impose a more elaborate hypothesis. The two-stage model, whether or not it describes exactly the folding pathway that is followed in vivo, provides an interesting approach to structural predictions and justifies attempts at building up transmembrane regions by assembling computationally preformed helices [see, e.g., Holm et al. (1987) or Furois-Corbin and Pullman (1989)]. It should be realized, however, that the difficulties of this task are considerable. Even if one assumes that the identification of the transmembrane helices is error free, the exact limits of each cannot be precisely defined, and the dynamics of tilting, bobbing, twisting, and bending [cf. Vogel et al. (1988)] permit a range of helix–helix contacts at the time of assembly. In BR, there is indirect evidence that reassociation of the two chymotryptic fragments involves proline–imide bond cis–trans isomerization (Popot et al., 1987). The presence of prolyl residues in transmembrane segments will result in various types of deviations from perfect α-helices, as in helix C of the reaction center L subunit (Deisenhofer et al., 1985; Michel et al., 1986a) and in helix B of BR (Popot et al., 1989). Given that the C^α positions in real helices deviate from the ideal and that side chains can adopt many different conformations, packing helices is not a trivial computational problem. It can probably best be approached when additional constraints can be resorted to, such as the geometrical constraints introduced by multiple liganding of prosthetic groups, as in cytochrome *c* oxidase [see, e.g., Holm et al. (1987)].

Conclusion

The observations and experiments summarized here can be simply understood by assuming that hydrophobic helices in transmembrane regions of proteins behave as autonomous folding domains, analogous to the larger domains characterized in soluble proteins. This type of structure is probably not particular to electron or proton pumps and may extend to proteins that delimit transmembrane aqueous channels. Membrane proteins that include strongly hydrophilic transmembrane segments do exist, but they may turn out to be the exception.

While undoubtedly an oversimplification, the two-stage model represents a useful way of rationalizing the structure and behavior of integral proteins in terms of assemblies of hydrophobic transmembrane helices. It provides a reasonable basis for model building that emphasizes what can be included in models over what cannot be excluded.

References

Allen, L. C. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4701–4705.

Atassi, M. Z., Mulac-Jericevic, B., Yokoi, T., & Manshouri, T. (1987) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 46*, 2538–2547.

Bormann, B. J., Knowles, W. J., & Marchesi, B. T. (1989) *J. Biol. Chem. 264*, 4033–4037.

Brunner, J., Franzusoff, A. J., Luscher, B., Zugliani, C., & Semenza, G. (1985) *Biochemistry 24*, 5422–5430.

Changeux, J.-P. (1990) *Fidia Res. Found. Neurosci. Award Lect.* (in press).

Chothia, C., & Janin, J. (1975) *Nature (London) 256*, 705–708.

Dani, J. A. (1989) *J. Neurosci. 9*, 884–892.

Dani, J. A., & Eisenman, G. (1987) *J. Gen. Physiol. 89*, 959–983.
Davis, N. G., & Model, P. (1985) *Cell 41*, 607–614.
Deisenhofer, J., Epp, O., Miki, K., Huber, R., & Michel, H. (1985) *Nature (London) 318*, 618–624.
Dennis, M., Giraudat, J., Kotzyba-Hibert, F., Goeldner, M., Hirth, C., Chang, J.-Y., Lazure, C., Chrêtien, M., & Changeux, J.-P. (1988) *Biochemistry 27*, 2346–2357.
de Vitry, C., & Popot, J.-L. (1989) *C. R. Acad. Sci., Ser. 3 309*, 709–714.
DiPaola, M., Czajkôwski, C., & Karlin, A. (1989) *J. Biol. Chem. 264*, 15457–15463.
Dwyer, B. P. (1988) *Biochemistry 27*, 5586–5592.
Dwyer, T. M., Adams, D. J., & Hille, B. (1980) *J. Gen. Physiol. 75*, 469–492.
Eisenberg, D. (1984) *Annu. Rev. Biochem. 53*, 595–623.
Engelman, D. M., & Zaccai, G. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5894–5898.
Engelman, D. M., & Steitz, T. A. (1981) *Cell 23*, 411–422.
Engelman, D. M., Goldman, A., & Steitz, T. A. (1982) *Methods Enzymol. 88*, 81–88.
Engelman, D. M., Steitz, T. A., & Goldman, A. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 321–353.
Finkelstein, A., & Ptitsyn, O. B. (1987) *Prog. Biophys. Mol. Biol. 50*, 171–190.
Flory, P. J. (1956) *J. Am. Chem. Soc. 78*, 5222–5235.
Furois-Corbin, S., & Pullman, A. (1989) *Biochim. Biophys. Acta 984*, 339–350.
Gilson, M. K., & Honig, B. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1524–1528.
Giraudat, J., Dennis, M., Heidmann, T., Chang, J.-Y., & Changeux, J.-P. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2719–2723.
Giraudat, J., Dennis, M., Heidmann, T., Haumont, P. T., Lederer, F., & Changeux, J.-P. (1987) *Biochemistry 26*, 2410–2418.
Giraudat, J., Galzi, J.-L., Revah, F., Changeux, J.-P., Haumont, P.-Y., & Lederer, F. (1989) *FEBS Lett. 253*, 190–198.
Glaeser, R. M., Jubb, J. S., & Henderson, R. (1985) *Biophys. J. 48*, 775–780.
Gregor, P., Mano, I., Maoz, I., McKeown, M., & Teichberg, V. (1989) *Nature (London) 342*, 689–692.
Grenningloh, G., Rienitz, A., Schmitt, B., Methfessel, C., Zensen, M., Beyreuther, K., Gundelfinger, E. D., & Betz, H. (1987a) *Nature (London) 328*, 215–220.
Grenningloh, G., Gundelfinger, E. D., Schmitt, B., Betz, H., Darlison, M. G., Barnard, E. A., Schofield, P. R., & Seeburg, P. H. (1987b) *Nature (London) 330*, 25–26.
Henderson, R. (1975) *J. Mol. Biol. 93*, 123–138.
Henderson, R. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 87–109.
Henderson, R., & Unwin, P. N. T. (1975) *Nature (London) 257*, 28–32.
Hol, W. G. J. (1985) *Prog. Biophys. Mol. Biol. 45*, 149–195.
Hollmann, M., O'Shea-Greenfield, A., Rogers, S. W., & Heinemann, S. (1989) *Nature (London) 342*, 643–648.
Holm, L., Saraste, M., & Wikström, M. (1987) *EMBO J. 6*, 2819–2823.
Honig, B. H., Hubbell, W. L., & Flewelling, R. F. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 163–193.
Huang, K.-S., Bayley, H., Liao, M.-J., London, E., & Khorana, H. G. (1981) *J. Biol. Chem. 256*, 3802–3809.
Hucho, F. (1986) *Eur. J. Biochem. 158*, 211–226.
Hucho, F., Oberthür, W., & Lottspeich, F. (1986) *FEBS Lett. 205*, 137–142.
Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., & Davies, D. R. (1988) *J. Biol. Chem. 263*, 17857–17871.
Imoto, K., Methfessel, C., Sakmann, B., Mishina, M., Mori, Y., Konno, T., Fukuda, K., Kurasaki, M., Bujo, H., Fujita, Y., & Numa, S. (1986) *Nature (London) 324*, 670–674.
Imoto, K., Busch, C., Sakmann, B., Mishina, M., Konno, T., Nakai, J., Bujo, H., Mori, Y., Fukuda, K., & Numa, S. (1988) *Nature (London) 335*, 645–648.
Jennings, M. L. (1989) *Annu. Rev. Biochem. 58*, 999–1027.
Jubb, J. S., Worcester, D. L., Crespi, H. L., & Zaccai, G. (1984) *EMBO J. 3*, 1455–1461.
Kellis, J. T., Jr., Nyberg, K., Sali, D., & Fersht, A. R. (1988) *Nature (London) 333*, 784–786.
Khorana, H. G. (1988) *J. Biol. Chem. 263*, 7439–7442.
Kleffel, B., Garavito, R. M., Baumeister, W., & Rosenbusch, J. P. (1985) *EMBO J. 4*, 1589–1592.
Klein, P., Kanehisa, M., & DeLisi, C. (1985) *Biochim. Biophys. Acta 815*, 468–476.
Kouyama, T., Kimura, Y., Kinosita, K., Jr., & Ikegami, A. (1981) *J. Mol. Biol. 153*, 337–359.
Lear, J. B., Wasserman, Z. R., & DeGrado, W. F. (1988) *Science 240*, 1177–1181.
Leonard, R. J., Labarca, C. G., Charnet, P., Davidson, N., & Lester, H. A. (1988) *Science 242*, 1578–1581.
Liao, M.-J., London, E., & Khorana, H. G. (1983) *J. Biol. Chem. 258*, 9949–9955.
MacIntyre, S., Freudl, R., Eschbach, M.-L., & Henning, U. (1988) *J. Biol. Chem. 263*, 19053–19059.
Maeno, T., Edwards, C., & Anraku, M. (1977) *J. Neurobiol. 8*, 173–184.
McCrea, P. D., Popot, J.-L., & Engelman, D. M. (1986) *Biophys. J. 49*, 355a.
McCrea, P. D., Popot, J.-L., & Engelman, D. M. (1987) *EMBO J. 6*, 3619–3626.
Michel, H., Weyer, K. A., Gruenberg, H., Dunger, I., Oesterhelt, D., & Lottspeich, F. (1986a) *EMBO J. 5*, 1149–1158.
Michel, H., Epp, O., & Deisenhofer, J. (1986b) *EMBO J. 5*, 2445–2451.
Molle, G., Dugast, J. Y., Duclohier, H., Daumas, P., Heitz, F., & Spach, G. (1988) *Biophys. J. 53*, 193–203.
Nabedryk, E., Garavito, R. M., & Breton, J. (1988) *Biophys. J. 53*, 671–676.
Noda, M., Shimizu, S., Tanabe, T., Takai, T., Kayano, T., Ikeda, T., Takahashi, H., Nakayama, H., Kanaoka, Y., Minamino, N., Kangawa, K., Matsuo, H., Raftery, M. A., Hirose, T., Inayama, S., Hayashida, H., Miyata, T., & Numa, S. (1984) *Nature (London) 312*, 121–127.
Numa, S. (1989) *Harvey Lect. 83*, 121–165.
Oberthür, W., Muhn, P., Baumann, H., Lottspeich, F., Wittmann-Liebold, B., & Hucho, F. (1986) *EMBO J. 5*, 1815–1819.
Oiki, F., Danho, W., Madison, V., & Montal, M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8703–8707.
Ovchinnikov, Y. A., Abdulaev, N. G., Vasilov, R. G., Vturine, I. Y., Kuryatov, A. B., & Kiselev, A. V. (1985) *FEBS Lett. 179*, 343–350.
Pace, C. N., Grimsley, G. R., Thompson, J. A., & Barnett, B. J. (1988) *J. Biol. Chem. 263*, 11820–11825.
Popot, J.-L., & Changeux, J.-P. (1984) *Physiol. Rev. 64*, 1162–1239.
Popot, J.-L., & de Vitry, C. (1990) *Annu. Rev. Biophys. Biophys. Chem. 19*, 369–403.
Popot, J.-L., Trewhella, J., & Engelman, D. M. (1986) *EMBO J. 5*, 3039–3044.

Popot, J.-L., Gerchman, S.-E., & Engelman, D. M. (1987) *J. Mol. Biol. 198*, 655–676.
Popot, J.-L., Engelman, D. M., Gurel, O., & Zaccai, G. (1989) *J. Mol. Biol. 210*, 829–847.
Rashin, A. A., Iofin, M., & Honig, B. H. (1986) *Biochemistry 25*, 3619–3625.
Ratnam, M., Le Nguyen, D., Rivier, J., Sargent, P. B., & Lindstrøm, J. (1986) *Biochemistry 25*, 2633–2643.
Rees, D. C., Komiya, H., Yeates, T. O., Allen, J. P., & Feher, G. (1989) *Annu. Rev. Biochem. 58*, 607–633.
Revah, F., Galzi, J.-L., Giraudat, J., Haumont, P.-Y., Lederer, F., & Changeux, J.-P. (1990) *Proc. Natl. Acad. Sci. U.S.A.* (in press).
Richards, F. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.
Rogers, N. K., & Sternberg, M. J. E. (1984) *J. Mol. Biol. 174*, 527–542.
Schofield, P. R., Darlison, M. G., Fujita, N., Burt, D. R., Stephenson, F. A., Rodriguez, H., Rhee, L. M., Ramachandran, J., Reale, V., Glencorse, T. A., Seeburg, P. H., & Barnard, E. A. (1987) *Nature (London)* 328, 221–227.
Singer, S. J. (1962) *Adv. Protein Chem. 17*, 1–68.
Singer, S. J. (1971) in *Structure and Function of Biological Membranes*, pp 145–222, Academic Press, New York.
Tanabe, T., Takeshima, H., Mikami, A., Flockerzi, V., Takahashi, H., Kangawa, K., Kojima, M., Matsuo, H., Hirose, T., & Numa, S. (1987) *Nature (London) 328*, 313–318.
Tempel, B. A., Papazian, D. M., Schwartz, T. L., Jan, Y. N., & Jan, L. Y. (1987) *Science 237*, 770–775.
Vogel, H., Nillson, L., Rigler, R., Voges, K.-P., & Jung, G. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5067–5071.
Wada, K., Dechesne, C. J., Shimasaki, S., King, R. G., Kusano, K., Buonanno, A., Hampson, D. R., Banner, C., Wenthold, R. J., & Nakatani, Y. (1989) *Nature (London) 342*, 684–689.
Yeates, T. O., Komiya, H., Rees, D. C., Allen, J. P., & Feher, G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6438–6442.

Accelerated Publications

Terminal Deoxynucleotidyltransferase Containing Megadalton Complex from Young Rat Thymus Nuclei: Identification and Characterization†

V. N. Pandey,‡ V. P. Dave,‡ M. S. Patil,‡ D. S. Pradhan,‡ S. B. Amrute,§ and M. J. Modak*,§

Biochemistry Division, Bhabha Atomic Research Centre, Bombay 400085, India, and Department of Biochemistry and Molecular Biology, New Jersey Medical School, University of Medicine and Dentistry of New Jersey, Newark, New Jersey 07103-2757

Received February 1, 1990; Revised Manuscript Received February 27, 1990

ABSTRACT: Nuclear matrix prepared from 2–3 week old rat thymuses contains tightly bound TdT activity which has been quantitatively solubilized with nonionic detergent and sonication. TdT is contained in a discrete complex with a sedimentation value of 23 S. The complex is retained on an anti-TdT antibody column and contains DNA ligase and 3′–5′ exonuclease activities as well as DNA and several other proteins but is devoid of replicative DNA polymerases. Such a type of multienzyme complex is absent from the nuclear extracts of thymus prepared from older rats and also from liver and spleen extracts of young and old rats.

Terminal deoxynucleotidyltransferase (TdT;[1] EC 2.7.7.31) is a unique DNA polymerase that catalyzes the polymerization of deoxyribonucleotides onto the 3′-OH end of DNA without template direction (Bollum, 1974). The restricted presence of TdT in prelymphocytes implies that TdT is closely related to lymphocyte development in both thymus gland and bone marrow (Baltimore, 1974; Chang & Bollum, 1987). Extensive biochemical characterization including active site determination of TdT has been carried out (Bollum, 1974; Bhalla et al., 1977; Modak, 1978, 1979; Modak & Gillerman-Cox, 1982; Pandey & Modak, 1987a,b, 1988a,b, 1989; Chang & Bollum, 1987). However, its biological role remains to be elucidated. Recent studies have provided some evidence to suggest that TdT may be involved in immunoglobulin (Ig) or T cell receptor (TCR) gene rearrangement/recombination to generate the diversified immune response observed in vertebrates (Alt & Baltimore, 1982; Desiderio et al., 1984; Roth et al., 1986; Alt et al., 1987). The process of TdT-mediated addition of nonconserved sequences, however, is only one of many steps in the process of recombination, and hence, the concerted action of several enzymes/proteins may be required at the site of rearrangement/recombination. In this context, various components may be expected to exist in the form of a multienzyme/protein complex. In addition, logistically one may not expect TdT to be free in the nucleoplasm since it has the

† This research was supported in parts by grants from New Jersey State Commission for Cancer Research and National Science Foundation (DMB-8715829) and by a fellowship to V.P.D. by the Atomic Energy Commission, India.
* Address correspondence to this author.
‡ Bhabha Atomic Research Centre.
§ University of Medicine and Dentistry of New Jersey.

[1] Abbreviations: TdT, terminal deoxynucleotidyltransferase; dNTP, deoxynucleoside triphosphates; DTT, dithiothreitol; PMSF, phenylmethane sulfonyl fluoride; Ig, immunoglobulin; TCR, T cell receptor; VDJ, variable, diversity, and joining segments of Ig or TCR genes; NP40, Nonidet P40; SDS, sodium dodecyl sulfate.

0006-2960/90/0429-4037$02.50/0