# Analysis and refinement of criteria for predicting the structure and relative orientations of transmembranal helical domains

Juan A. Ballesteros and Harel Weinstein
Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York, New York 10029 USA

## SUMMARY

We are interested in modeling the membrane-spanning domain of the serotonin 5-HT$_{1A}$ G-protein coupled receptor. This superfamily of proteins is predicted to share the topology of the seven transmembrane helices of bacteriorhodopsin (BR), even though no significant sequence homology had been identified (1). We found significant homologies by allowing for helix shuffling corresponding to minimal exon shuffling during evolution (2). Consequently, our strategy for building the model for the 5-HT$_{1A}$ receptor has been to construct hypotheses concerning helix–helix interactions, their orientations, and arrangement in bundles surrounded by lipid, based on the 3.5 Å resolution structure of BR (3). Inferences resulting from such models were tested against the 2.3 Å resolution structure of the photosynthetic reaction center (PRC) from *Rhodobacter Viridis* (4). These comparisons led us to a reevaluation of current methods for the identification and topological orientation of membrane-embedded α-helices. We find that methods used currently (5) in the construction of helical transmembrane domains could be misleading if used indiscriminately. These methods include the hydrophobicity profile, the hydrophobic moment, helix amphiphilicity, and charge neutralization. A refinement is proposed here, based on empirical observations, molecular modeling, and physicochemical considerations designed to overcome some of the shortcomings inherent in the use of the above mentioned methods. Here we present the analysis of two of the motifs identified in our study that led to the proposed refinements: the distribution of acidic and basic residues in the transmembranal domains, and the kink induced by a Pro residue in an α-helix.

## BASIC AND ACIDIC RESIDUES

*The specific distribution of basic and acidic residues we identified in BR could provide a biophysical basis for the "positive-inside" rule determining α-helix (Hx) topology (6).* There is a clear asymmetric distribution of acidic/basic residues within the transmembrane α-helical domains. The basic amino acids (AA) are localized at cytoplasmatic boundaries, whereas the acidic AA are more prominent at protein–protein interfaces. In BR, seven out of eight basic AA are at the cytoplasmatic terminus; all seven are protruding into the membrane phase, possibly, to interact with negatively charged lipid head groups (Lys in Hx 6 has been identified as such by Henderson [3]). This observation for BR is verified by the analysis of PRC. In both BR and PRC, the highly polarized distribution suggests that the basic residues function to anchor the Hx to the cytoplasmatic side of the membrane by means of ion pairs. This hypothesis is supported by the higher content of negatively charged lipid head groups in the cytoplasmic layer. Notably, in the Protein Data Bank coordinates for BR, most of these Arg/Lys were found to have relatively low values of the temperature factors. The observations regarding the position and function of charged amino acids, and the related hypothesis regarding their role in the stabilization of the transmembrane domain, affects current prediction methods in the following manner:

(*a*) the criterion of mutual neutralization between acidic and basic residues (7) is weakened, and may not hold in most cases.

(*b*) Taking into consideration the possible anchoring role of Arg/Lys residues affects the common prediction that position Arg/Lys at the cytoplasmatic side outside the membrane-spanning domain. Such predictions are likely from the "+inside rule," and from any method measuring the hydrophobic moment or hydrophobic-profile with windows > 1, due to the high hydrophilicity of the Arg/Lys residues. However, due to the length of the Arg/Lys side chains, these AA can appear 1.–1.5 turns within the membrane spanning domain, as is the case for three out of seven such residues in BR. If these observations regarding the Arg/Lys residues are ignored, then current methods for the prediction of Hx transmembrane domains may shift the membrane-spanning region artificially to exclude what should be intramembranal turns.

(*c*) Arg/Lys within the membrane-spanning domain are generally predicted to be on protein–protein inter-

0006-3495/92/04/107/03 $2.00

faces due to their strong hydrophilicity. However, if they serve in the mode described above, this positioning is erroneous, and methods such as the hydrophobic moment would yield either the wrong orientation, or an incorrectly weaker hydrophobic profile.

Taken together, these considerations can correct some of the misleading features of previously used criteria. They constitute a new, powerful guide for identifying the cytoplasmatic end of a transmembrane helix. Importantly, this criterion can eventually be explored quantitatively from computational simulations.

## PROLINE-KINKS IN HX

*A property of the Proline-kinks in Hx, combined with the role of the free carbonyls at the kink region, provides special structural clues regarding requirements for ligand binding to the receptor, and/or the ensuing receptor activation mechanism.* Because the kink-region does not have a 3.6 AA/turn periodicity, one would expect a twisting of the faces before/after the Pro-kink relative to the straight Hx case. Such an effect would vitiate the profiles for amphiphilicity (8) and hydrophobic moments of Pro-containing Hx because they all assume regular and straight Hx. Yet, such profiles are commonly used for the prediction/modeling of transmembrane Hx without special attention to the effects of Pro-kinks. The analysis of the degree of face twisting in known protein structures shows only subtle twisting. A rotation around the backbone dihedral angles ($\phi$, $\psi$) for Hx could conserve the azimuthal location of the C-$\alpha$'s, with very little twisting (G. Nemethy, personal communication). This rearrangement is favored by the lower energy required to modify a dihedral angle, compared to other internal coordinates, to attain a kinked Hx. Moreover, the conformation of Pro residues and Pro-kinks in Hx has been shown to be highly variable with a relatively wide energy minimum (9). Therefore, any specific interaction such as H-bonding to backbone Hx groups could affect the orientation of the axis of the dihedral angles associated with the residue, thus inducing the face twisting mentioned above. Clear candidates for such interaction are presented below.

A Pro residue within an Hx disrupts two 1–4 backbone H-bonds, $C = O_{i-4} \ldots CN_i$ due to the imide bond, and $C = O_{i-3} \ldots HN_{i+1}$ due to the axial kink (9). This implies up to three putatively H-bond free reactive sites at the kink region, the C = O's being unusually exposed (10). Most charged and polar sidechains in transmembrane domains are likely to be involved in extensive H-bonding due to the low dielectric constant of the transmembranal domain. Consequently, we would expect the three types of reactive sites in the backbone to make up a significant percentage of the total number of free H-bonding donors or acceptors within the transmembrane region. The involvement of such available polar sites in the function of the proteins has been proposed recently based on mutagenesis studies of the Pro residues in *lac permease* transporter (11), and of a bacterial $Ca^{2+}$-channel crystallized in nonpolar solvents (12), as well as on the basis of theoretical considerations (13).

It is important to note that the distortion introduced by a Pro-kink provides a dynamically responsive mechanism through which information about ligand binding could be propagated in the transmembrane structure. This mechanism requires a low activation energy, thus providing a reasonable mode for signal transduction. Such hypotheses anchored in the structural and dynamic properties of the receptor protein are amenable to exploration by current methods of theoretical biophysics.

The work was supported in part by grants from the National Institute on Drug Abuse: DA-06620, and Research Scientist Award: DA-00060 (Dr. H. Weinstein), and by a Fulbright/Ministry of Education and Science, Spain (MEC) scholarship (J. A. Ballesteros).

## REFERENCES

1. Henderson, R., and G. F. X. Schertler. 1990. The structure of bacteriorhodopsin and its relevance to the visual opsins and other seven-helix G-protein coupled receptors. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 326:379–389.
2. Pardo, L., J. A. Ballesteros, R. Osman and H. Weinstein. 1992. On the use of the transmembranal domain of bacteriorhodopsin as a template for modeling the 3-D structure of G-protein coupled receptors. *Proc. Natl. Acad. Sci. USA.* In press.
3. Henderson, R., J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckman, and K. H. Downing. 1990. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* 213:899–929.
4. Deisenhofer, J., and H. Michel. 1989. The photosyntetic reaction center from the purple bacterium rhodopseudomonas viridis. *Science (Wash. DC).* 245:1463–1473.
5. Fasman, G. D., and W. A. Gilbert. 1990. The prediction of transmembrane protein sequences and their conformation: an evaluation. *Trends Biol. Sci.* 15:89–91.
6. Von Heijne, G., and C. Manoil. 1990. Membrane proteins: from sequence to structure. *Protein Eng.* 4:109–112.
7. Engelman, D. M., R. Henderson, A. D. McLachlan, and B. A. Wallace. 1980. Path of the polypeptide in bacteriorhodopsin. *Proc. Natl. Acad. Sci. USA.* 77:2023–2027.
8. Segrest, J. P., H. De Loof, J. G. Dohlman, C. G. Brouillette, and G. M. Anantharamaiah. 1990. Amphiphatic helix motif: classes and properties. *Proteins Struct. Funct. Genet.* 8:103–117.
9. Piela, L., G. Nemethy, and H. A. Scheraga. 1987. Pro-induced constraints in $\alpha$-helices. *Biopolymers.* 26:1587–1600.
10. Woolfson, D. N., and D. H. Williams. 1990. The influence of Pro residues on $\alpha$-helical structure. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 277:185–188.

11. Consler, T. G., O. Tsolas, and H. R. Kaback. 1991. Role of proline residues in the structure and function of a membrane transport protein. *Biochemistry.* 30:1291–1298.
12. Karle, I. L., J. L. Flippen-Anderson, S. Adarwalla, and P. Balaram. 1991. Crystal Structure of [Leu]zervamicin, a membrane ion channel peptide: Implications for gating mechanisms. *Proc. Natl. Acad. Sci. USA.* 88:5307–5311.
13. Williams, K. A., and C. M. Deber. 1991. Proline residues in transmembrane helices: Structural or dynamic role? *Biochemistry.* 30:8919–8923.