Journal of Receptor, Ligand and Channel Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Hydrophobicity profiles in G protein-coupled receptor transmembrane helical domains

This article was published in the following Dove Press journal:
Journal of Receptor, Ligand and Channel Research
30 November 2010
Number of times this article has been viewed

Chiquito J Crasto

Division of Research, Department of Genetics, University of Alabama at Birmingham, Alabama, USA

Abstract: The lack of a crystallographically derived structure for all but three G (TP [guanosine triphosphate]-binding) protein-coupled receptor (GPCRs) proteins necessitates the use of computationally derived methods to determine their structures. Computational methodologies allow a mechanistic glimpse into GPCR–ligand interactions at a molecular level to better understand the initial steps leading to a protein's biologic functions, ie, protecting the ligands that activate, deactivate, or inhibit the protein, stabilizing protein structure in the membrane's lipid bilayer, and ensuring that the hydrophilic environment within the GPCR-binding pocket is maintained. Described here is a formalism that quantifies the amphiphilic nature of a helix, by determining the effective hydrophobicity (or hydrophilicity) at specific positions around it. This formalism will enable computational protein modelers to position helices so that the functional aspects of GPCRs are adequately represented in the model. Hydro-Eff®, an online tool, allows users to calculate effective helical hydrophobicities.

Keywords: GPCR, effective hydrophobicity, amphiphilicity

Introduction

G (TP-binding) protein-coupled receptor (GPCR) proteins are ubiquitous. They traduce cell membranes, span several tissues, and are responsible for myriad functions.[1,2] GPCRs serve as conduits of material within the cell from outside. Alternatively, their activation following ligand binding serves to initiate intracellular processes (eg, G-protein coupling and signal transduction).[3,4] The publication of initial and subsequent revised drafts of the human genome brought to light the ubiquity of GPCRs in membrane-related function.[5] Olfactory receptors, which constitute superfamilies in mammalian genomes, are an example of GPCRs.[6–8]

Structure of GPCRs

GPCRs are believed to possess a tertiary structure comprising an assembly of seven transmembrane helical domains connected by three intracellular and three extracellular loops, and possessing an extracellular N-terminus and an intracellular C-terminus.[9,10] While this is typical, anomalies do exist; a recent publication has identified a functional olfactory receptor with an anomalous structure of six transmembrane regions and an extracellular C-terminus.[11]

Three GPRCs have been structurally analyzed by X-ray crystallography, ie, rhodopsin,[12] the beta-adrenergic receptor,[13] and the adenosine A_{2A} receptor.[14] All structures have been solved at relatively high resolution, with the resolution for

Correspondence: Chiquito J Crasto
752-A Kaul, 1530 3rd Avenue S,
Birmingham, AL 35294, USA
Tel +1 205 996 7083
Fax +1 205 975 5689
Email chiquito@uab.edu

the structure of rhodopsin improving over several studies.[9] During structure solution of the beta-adrenergic receptor, researchers used molecular replacement methods relying on rhodopsin to serve as a structural template for select regions where the X-ray reflection-derived electron density was not available.[13] In the case of the adenosine A_{2A} receptor, portions of the protein were replaced with lysozyme to facilitate crystallization,[14] and the protein was complexed with an antagonist. Such measures are used to alleviate the instability of these membrane-bound proteins (primarily because of solubility issues) during purification and crystallization.[15]

Computational modeling methodologies, such as homology modeling, have been utilized to assess the structure of GPCRs. In homology modeling,[16,17] the structure of rhodopsin has most often been used as a template to model the target GPCR. However, it is likely that the structures of beta-adrenergic and adenosine receptors would also be used in future homology modeling protocols.[18]

There is significant diversity among GPCR sequences.[19,20] Therefore, unless there is strong sequence similarity, as a precursor to homology modeling, the helical regions of "target" GPCR have been structurally, and not sequentially, matched to the rhodopsin "template". Transmembrane helical regions in the test GPCRs have been identified through secondary structure prediction methods,[21,22] or, as has more recently been done, through hidden Markov models (HHMs), which identify the probability of amino acid sequences being transmembrane helices.[23,24] Earlier methodologies have addressed the lack of homology between GPCRs and rhodopsin by building canonical or idealized helices and "mounting" them over the low-resolution electron densities of the transmembrane region of rhodopsin, derived from electron diffraction experiments.[19,25] Each method, whether using high- or low-resolution structures of rhodopsin, presents unique advantages and challenges. Using high-resolution structures as a homology-modeling template allows for better side-chain and rotameric positioning. However, they introduce certain template-specific structural artifacts into the final structure, which have to be addressed in the steps following homology modeling. Homology modeling is also hindered due to the variable lengths of helices and loops between the target and the template. The average length of the helices in rhodopsin is roughly 30 amino acid residues,[12] and that for the beta-adrenergic receptor is about 20 amino acid residues.[13] More rigorous procedures that start from homology modeling have been used to model GPCRs, especially olfactory receptors.[11,20,25]

Rationalizing the hydrophilic core of the tertiary protein structure

Individual amino acid residues that point towards the interior of the protein are responsible for ligand binding through a combination of covalent, electrostatic, and van der Waals interactions. A model of a GPCR should preserve these binding features and those that rationalize the protein's core. Once the transmembrane helical scaffold has been constructed via modeling, helices have to be rotated such that, presumably, the hydrophilic side of the helix is pointed towards the interior of the protein, ie, where ligand binding is expected to take place. This positions the hydrophobic side of the helix towards the hydrophobic lipid bilayer.

This paper describes a formalism that quantifies helical amphiphilicity by determining the effective hydrophobicities at different positions in a helix. The effective hydrophilicity is calculated at specific angles around the helix, which in this methodology is represented as a helical wheel. The effective hydrophobicity includes not only the hydrophobicity of the amino acid residue at that angle, but also contributions from the hydrophobicities of all other residues in the helix. The effective hydrophobicity thus takes into account how the electronic nature (eg, polarity) of surrounding amino acid residues will affect the hydrophobicity at a specific point (or angle) for a helix.

The first comprehensive review of the structural aspects of GPCRs and membrane proteins was published by Eisenberg.[26] Eisenberg arrived at a theoretic formalism for hydrophobic moments of membrane protein helices. This hydrophobic moment was a single value for the entire helix and quantified as a number per residue. It is a vector sum over all angles of side chains and includes hydrophobicities of the residues involved.[27] The larger this value, the larger the amphiphilicity of the helix.[26] Programs such as the PERSCAN software (currently unavailable) have used Fourier transform methodologies based on differences between substitution frequencies of buried residues in water-soluble proteins and lipid-accessible residues to determine helical amphiphilicities.[28] Recent reviews[29–31] of research related to membrane protein structures have made inferences based on observations of the crystallographically derived structures of GPCRs, which were not available at the time of the Eisenberg review.

The work described here revisits helical amphiphilicities in GPCRs, in light of published X-ray-determined structures for rhodopsin, the beta-adrenergic receptor and the adenosine A_{2A} receptor.

Methods

Effective hydrophobicity

The effective hydrophobicity, Θ_θ, for an alpha helix is determined by Equation 1.

$$\Theta_\theta = \sum_{i=o}^{360-\theta} \mu_\theta \cdot \cos i \qquad (1)$$

This value is calculated at angles θ, ie, the angles that residue side chains occupy on a helical wheel; μ is the hydrophobicity value for each amino acid residue that resides at angle θ. Although residues occupy 20° intervals on a helical wheel, sequentially the amino acids are positioned at 100° intervals along the helical turns. For example, for a helix FGPTGCNLEGFF, the angles on the wheel will be occupied thus: F(0°), G (100°), P (200°), T (300°), G (40° [60° + 300° = 360°≡ 0° + 40° = 40°]), C (140°), N (240°), L (340°), E (80°), G (180°), F (280°), and F (20°). To illustrate how Equation 1 works, consider a residue at an angle θ = 100°. This is the angle in question, ie, the angle at which the effective hydrophobicity will be measured. The first term in Equation 1 is $\mu_{100} \times \cos 0$. The vector summation in Equation 1 is then over cosine contributions (projections) from other residues at other angles to the angle being considered. Since residues occupy, in sequence, angles at intervals of 100° in a helical wheel, the next residue in this sequence is at angle 200°. The contribution from this residue to the residue (at the angle in question), ie, the second term in equation 1, is $\mu_{200} \times \cos 100$. The contribution from the third term, 200° away, is $\mu_{300} \times \cos 200$, and so forth. Figure 1 illustrates how Equation 1 is used to determine effective hydrophobicities at angles on a helical wheel.

If the sequence is long enough that another residue lies at angle 100° (approximately six turns lower along the helical axis), then this hydrophobicity is added to the one in question times the cosine of the angle 0°. For helices of sequences containing greater than 18 residues, additional residues are superimposed on previously occupied positions on the helical wheel. For example, in Figure 1, the Glu at angle 0° and the Gly at angle 360° are superimposed.

The μ values used in the equation can be derived from hydrophobicity scales, reviewed extensively by Eisenberg.[27] These hydrophobicities have been variously determined, ie, through thermodynamic calculations of the free energy changes with an amino acid in an aqueous versus a hydrophobic environment, or, semiempirically, based on a survey of the likelihoods of amino acid residues being buried versus exposed, or a combination of both. The scale most often used (and used in this paper) is the Eisenberg consensus hydrophobicity scale,[27] which were obtained by a simple averaging of hydrophobicities determined, using different conceptual methodologies, by Chothia,[32] Janin,[33] and Von Heijne and Blomberg.[34]

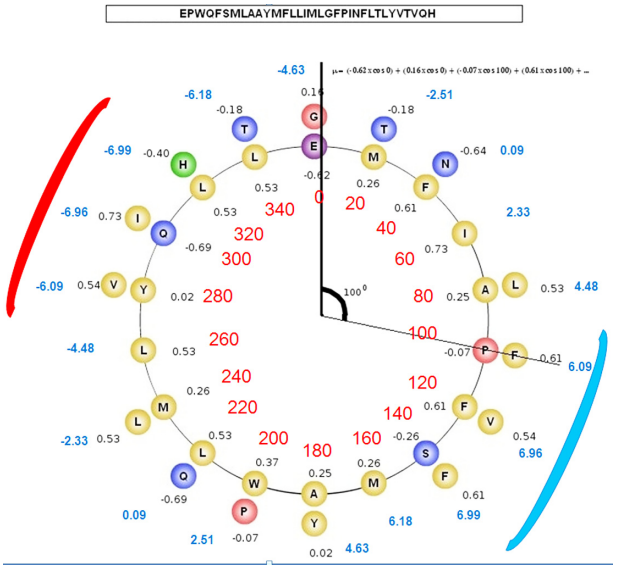

Figure 1 Equation 1 was used to determine effective hydrophobicity. The helical wheel for a fictitious sequence (in the inset box) was created using an Internet tool (see http://www-nmr.cabm.rutgers.edu/bioinformatics/Proteomic_tools/Helical_wheel/). The calculation for μ is shown for the first few residues. The numbers for each residue reflect the hydrophobicities for the residues as taken from Eisenberg's consensus hydrophobicity scales. The angles at which effective hydrophobicities, Θ_θ, are determined are in red on the inside of the wheel. The effective hydrophobicities in blue are on the outside of the helical wheel. The red curve shows the hydrophilic regions as identified by Hydro-Eff®. The blue curve shows the hydrophobic region for this fictitious helical sequence.

Hydro-Eff®: An Internet-accessible tool to determine effective hydrophobicities

An Internet-based tool, Hydro-Eff® (see http://bioinfo.genetics.uab.edu/hydro-eff.pl) allows users to calculate Θ_θ. Figures 2 and 3 illustrate the web user interface and the results page. The web page contains information about the concepts behind Hydro-Eff. The user can enter a sequence in the following format: Helix1:EPWQFS …, in the text box. The "Helix1" or user-preferred designation is necessary to identify helices, because Hydro-Eff allows the user to enter unlimited helical sequences in the text box. For example, the input for GPCRs will typically consist of seven helices. Also, important is the use of the ":" that separates the helical designation from the sequence. This format is required by the Hydro-Eff program.

Hydro-Eff software is written using PERL (Practical Extraction and Report Language). It uses the CGI.pm module to allow access to Hydro-Eff via the Internet. The front page and the results page are annotated with a description of the program with instructions for its use. The author's email address is included, so that users can contact the author for

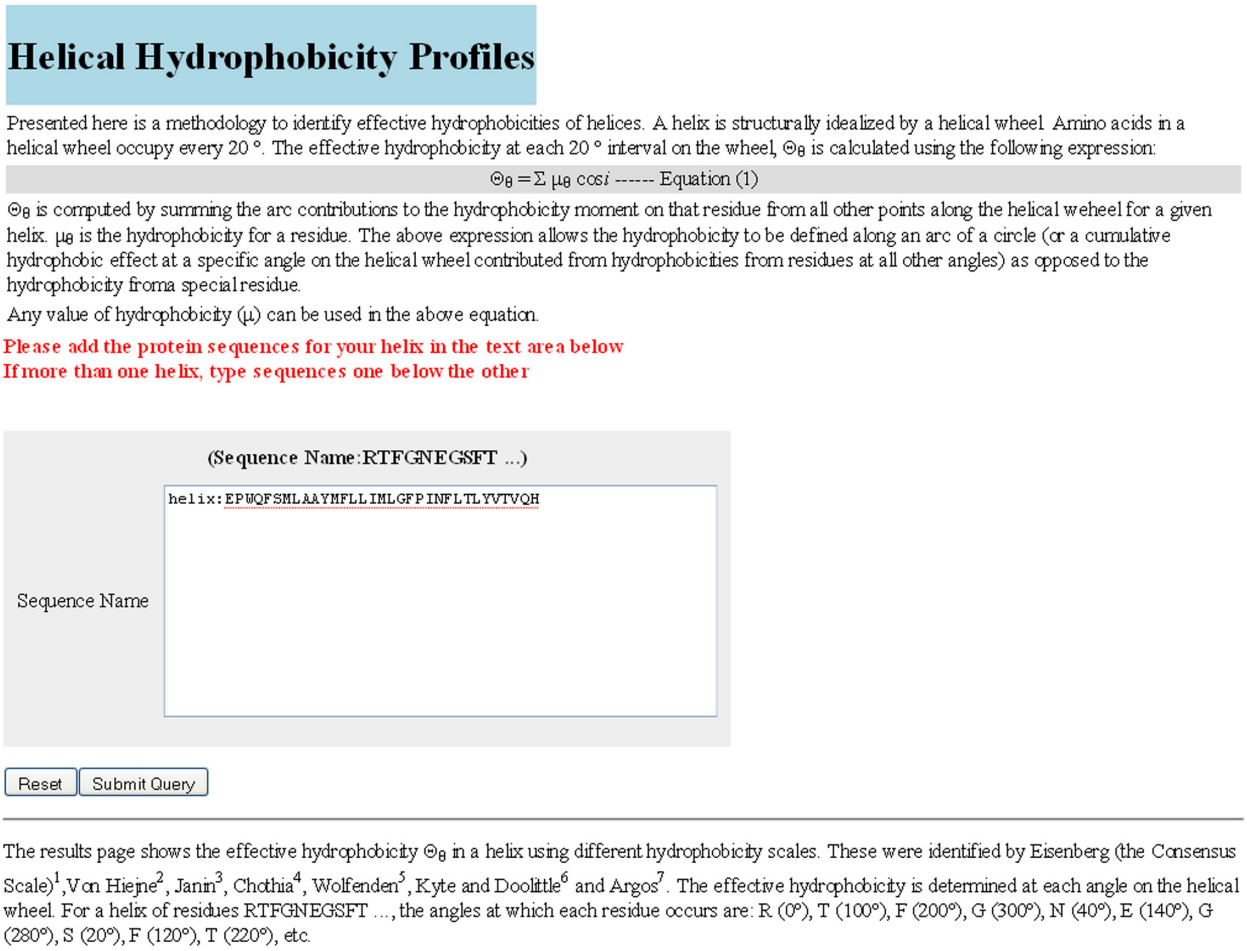

Figure 2 A screen capture of the Web page for the Hydro-Eff® tool. The user can enter one or more sequences in the text box. The text on the web page explains the rationale for Equation 1 and how it can be used to determine effective hydrophobicities. Hydro-Eff can be accessed at http://bioinfo.genetics.uab.edu/hydro-eff.pl. Accessing this program requires an Internet browser.

help in using Hydro-Eff. To execute Hydro-Eff, only an Internet connection and a browser are needed. All calculations of Θ_θ are performed on the server side.

The results are tabulated such that the Θ_θ values are obtained at specific angles on the helical wheel. The Θ_θ values are calculated using the hydrophobicities (μ) determined by Eisenberg (the consensus scale),[26,27] Von Hiejne,[35] Janin,[33] Chothia,[32] Kyte and Doolittle,[36] and Argos et al.[37] The values of μ are taken from Table 2 of Eisenberg's review.[27] The values for μ published by Tanford and Nozaki[38] and Segrest and Feldmann[39] are not included in the calculations for Θ_θ because μ values for the polar residues Asp, Arg, Lys, and His were not provided by this method. For the same reason, Wolfenden's scales which do not include a μ value for Pro, are not included in the calculations.[40]

Results

In Figure 1, the helical wheel was generated using an Internet-based tool (see http://www-nmr.cabm.rutgers.edu/bioinformatics/Proteomic_tools/Helical_wheel/) developed by John K Everett. Figure 1 shows that the amino acid residues are positioned at the 20° angles on a helical wheel for a fictitious helix of sequence: EPWQFSMLAAYMFLLIMLGFPINFLTLYVTVQH. The first residue (Glu) is at angle 0°. The next residue (Pro) is at angle 100°. The third residue (Trp) is at angle 200°, and so on. Consider the summation in equation 1. At angle 0°, the hydrophobicity of Glu (–0.62) contributes to Θ. Because the angle it makes with itself is 0°, the cos θ contribution is 1, and therefore the contribution is from hydrophobicity only. For the next angle at 100°, consider the vector passing through angle 100° on the wheel represented by the Pro side chain. If this vector were resolved, the cos100 projection would fall on the residue at 0°, namely at Glu. This is continued at every 100° interval until Θ_θ at all angles and all residues are covered. Figure 1 shows the partial calculation of the effective hydrophobicity, Θ_θ at angle 0. At this angle, two residues

The results page shows the effective hydrophobicity Θ_θ in a helix using different hydrophobicity scales. These were identified by Eisenberg (the Consensus Scale),Von Hiejne, Janin, Chothia, Wollenden, Kyte and Doolittle and Argos. The effective hydrophobicity is determined at each angle on the helical wheel. For a helix of residues RTFGNEGSFT ..., the angles at which each residue occurs are: R (0°), T (100°), F (200°), G (300°), N (40°), E (140°), G (280°), S (20°), F (120°), T (220°), etc.

helix:EPWQFSMLAAYMFLLIMLGFPINFLTLYVTVQH

Angle	Consensus	von Heijne	Janin	Chothia	Wolfenden	Kyte	Argos
0	-1.313	-12.488	-0.957	0.760	-7.689	-1.775	-2.582
20	-0.892	-10.350	-0.508	1.310	-4.974	-0.536	-2.667
40	-0.364	-6.965	0.002	1.702	-1.660	0.768	-2.430
60	0.208	-2.739	0.512	1.889	1.855	1.980	-1.900
80	0.755	1.818	0.960	1.848	5.145	2.952	-1.141
100	1.210	6.155	1.292	1.584	7.816	3.569	-0.244
120	1.520	9.749	1.468	1.129	9.543	3.755	0.682
140	1.647	12.168	1.468	0.538	10.120	3.488	1.526
160	1.575	13.119	1.290	-0.118	9.476	2.800	2.186
180	1.313	12.488	0.957	-0.760	7.689	1.775	2.582
200	0.892	10.350	0.508	-1.310	4.974	0.536	2.667
220	0.364	6.965	-0.002	-1.702	1.660	-0.768	2.430
240	-0.208	2.739	-0.512	-1.889	-1.855	-1.980	1.900
260	-0.755	-1.818	-0.960	-1.848	-5.145	-2.952	1.141
280	-1.210	-6.155	-1.292	-1.584	-7.816	-3.569	0.244
300	-1.520	-9.749	-1.468	-1.129	-9.543	-3.755	-0.682
320	-1.647	-12.168	-1.468	-0.538	-10.120	-3.488	-1.526
340	-1.575	-13.119	-1.290	0.118	-9.476	-2.800	-2.186

Figure 3 A screen capture of results of effective hydrophobicity at different angles (at 20° intervals) on an idealized helical wheel. The user can access results for effective hydrophobicities determined using seven different hydrophobicity scales. The text on the page also provides some information about the different hydrophobicity scales. Hydro-Eff® results are accessible via an Internet browser.

are superimposed on the idealized wheel, E and G. The hydrophobicity values μ for these are –0.62 for glutamic acid and 0.16 for glycine (taken from the Eisenberg consensus hydrophobicity scales[26]). Thus, the effective hydrophobicity extends along the axis of the helix but is positioned at a specific angle. The effective hydrophobicity at angle 0° is determined as (the following is incomplete, but is used to illustrate how the calculation occurs, shown in Figure 1):

$$\Theta_0 = (-0.62 \times \cos 0) + (0.16 \cos 0) + (0.26\ [\text{for M at } 20°] \times \cos 20) + (-0.18\ [\text{for T at } 20°]) + (0.61\ [\text{F at } 40°] \times \cos 40) + (-0.64\ [\text{N at } 40°]) \times \cos 40 + (0.73\ [\text{I at } 60°] \times \cos 60) + \cdots$$

From a biochemical standpoint, one would expect the hydrophobicity at one angle on the wheel to be perturbed by neighboring and surrounding residues because of the electronic withdrawing and donating effect of their acidic, basic, or neutral natures. As the angle increases from the angle in question, the decreasing value of the cosine of that angle will reduce the perturbing effect of the side chain for that residue.

In Figure 1, the effective hydrophobicity values are on the outside (in blue) and the angles of the helical wheel are on the inside (in red). The other numbers represent the μ values. The most hydrophilic region illustrated by a red curve is between angles 260° and 360° and is centered between angles 300° and 320°. Figure 1 shows that histidine (320°), glutamic acid (0°), and glutamine (300°) contributed to the hydrophilicity of this side of the helix. The large number of nonpolar residues ensured that the most hydrophobic regions are between angles 80° and 180°, represented by a blue curve. Thus, equation 1 allows the user to quantify the amphiphilicity of a helix.

Hydro-Eff results

The results page (Figure 3) shows the tabulated effective hydrophobicities Θ_θ for a helix using six different hydrophobicity scales, ie, Eisenberg,[27] Von Hiejne,[35] Janin,[33] Chothia,[32] Kyte and Doolittle,[36] and Argos.[37] The values for μ are taken from Table 2 by Eisenberg and colleagues.[26] The tabulated results are read, not in terms of the

peptide sequence, but in terms of the effective hydrophobicity at a specific angular location on the helical wheel. Indeed, the contribution from the residues at that specific angle is the highest because the hydrophobicity μ is multiplied by cos 0 = 1.

Discussion

In this paper, a formalism that quantifies the amphiphilic nature of an alpha helix is introduced. This formalism will assist computational biologists and crystallographers in identifying helical amphiphilicities. The availability of crystal structures of three GPCRs allows us to test whether Equation 1 properly represents transmembrane helical positioning for the asymmetric unit within the crystallographic unit cell, and possibly the biologic environment.

Equation 1 provides a mathematic basis that includes the biologic imperative of helical positioning. Hydrophobicity contribution from a side chain at a specific angle will be significantly reduced, enhanced, or otherwise influenced, because a bulky electron density withdrawing or donating side chain is present at neighboring positions (angles on the helical wheel). The cosine vector contribution represents the contribution of a neighboring amino acid residue side chain. As the cosine values decrease with increasing angle, the contribution (biologic and vector) from the cosine component of the angle of the residue side chain is likely to decrease for residue at angles removed from the angle at which the effective hydrophobicity is being calculated.

The structures of rhodopsin (mostly) and the beta-adrenergic receptor have been used to model GPCRs,[11,25,41–50] with a recent publication listing use of the adenosine A_{2A} receptor.[18] If there is weak homology between the GPCR and rhodopsin or beta-adrenergic receptor (the template used during homology modeling), homology modeling can only be the first step in the modeling protocol. Homology modeling can only be used to establish basic helical positions, much like the low-resolution helical densities from electron diffraction experiments. This affects how the helices are rotated to maintain a hydrophilic core. Alternative protocols have been designed that determine helical rotations in the GPCR transmembrane assembly based purely on lowest system energies.[25] These do not take hydrophobicities into account.

The polar core of a transmembrane assembly is typically centered on one or more highly polar amino acid residues, such as arginine, lysine, histidine, aspartic acid, and glutamic acid. Deploying Equation 1 precludes the need to identify residues that might contribute to helices individually. These are centered on a specific angle on the helical wheel, as illustrated by the red and blue curves in Figure 1, and not on a specific amino acid residue.

Effective hydrophobicities for transmembrane domains in rhodopsin and beta-adrenergic receptors

Mutagenesis experiments and functional analysis experiments identify agonists and antagonists that bind to these receptors and activate or inhibit their function.[51,52] Binding ligands may form covalent or electrostatic bonds with specific amino acids.[53] Because ligand binding is followed by GPCR activation, it is likely that other nonpolar interactions also facilitate the activation. The evidence for the role of specific residues in binding is that function diminishes or is enhanced by mutating these residues using residues of varying polarity and size and length of side chains.[54]

Hydro-Eff was used to determine the effective hydrophobicities of rhodopsin and the beta-adrenergic and adenosine receptors. The helical sequences were identified by observing the structures of asymmetric units of the three protein structures: PDB code 1U19 for rhodopsin; PDB code 2R4R for the beta-adrenergic receptor; and PDB code 3EML for the adenosine A_{2A} receptor. Here, the structure of the beta-adrenergic receptor with the PDB identifier 2R4R is used. Another structure of the beta-adrenergic receptor with a tobacco etch virus cleavage site has also been published (PDB identifier 2R4S).[13] Both structures were inspected, and the peptide sequences that constitute helices in both structures are not different. Because the Hydro-Eff methodology relies on sequences that make up a helix, any results stemming from the calculation of effective hydrophobicities will be identical for both structures.

The hydrophilic regions predicted by Hydro-Eff are in green and the polar residues are in red, with side chains shown. Figures 4–6 illustrate how polar residues significantly impact the amphiphilicities (through their μ values), although these residues do not often exactly coincide with the effective hydrophobicity for that angle. In Figures 4 and 5, where the effective hydrophobicities coincide with polar residues, the side chains are shown and the residue is colored green.

For rhodopsin, the hydrophilic transmembrane regions I, VI, and II, as determined using Hydro-Eff, are pointed towards the lipid membrane. The polar sides of the transmembrane regions III and VII are pointed towards the interior. The polar transmembrane regions IV and V are pointed towards each other (Figure 4). For the beta-adrenergic receptor, the polar side of the transmembrane helices II, III, IV, and V are pointed

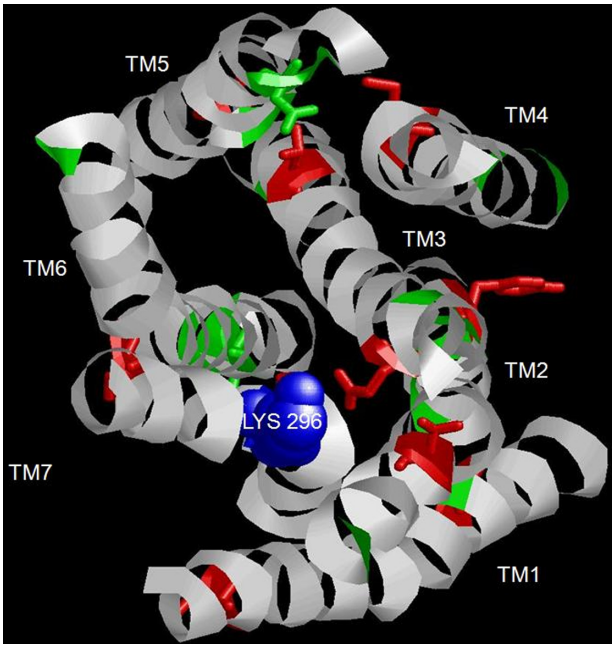

Figure 4 Amino acids and side chains for the residues for the transmembrane helices on which the effective hydrophobicity residues for rhodopsin. The green regions (without side chains) in the helices show the direction of the effective hydrophobicity for that helix. The amino acid residue and side chains, highlighted in red, show charged residues in the helices of the protein (Arg, Lys, Asp, Glu, His). This is to illustrate that although the effective hydrophobicities do not necessarily reside on the most charged residue, equation 1 takes into account the charged residues in its determination of Θ_θ which is pointed in roughly the same direction as the polar side chains. Amino acids in green with side chains showing are those where the effective hydrophobicity Θ_θ resides specifically on a polar amino acid residue. The space-filled blue residue is Lys296 implicated in a covalent bonding with retinal. This residue is pointed into the interior of the protein, where ligand binding is likely to take place.

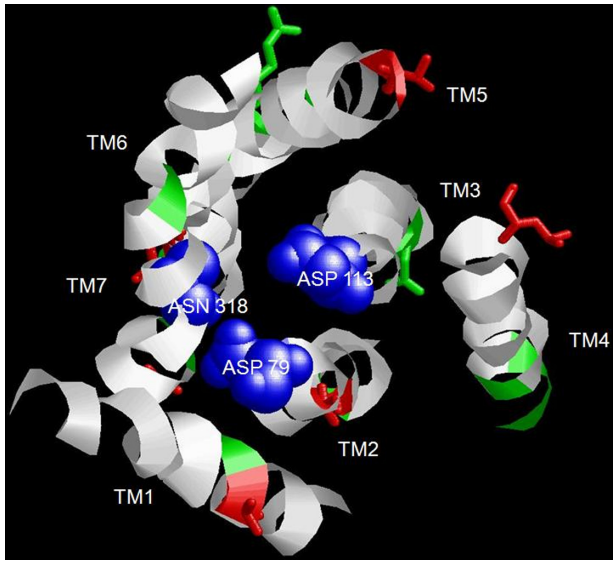

Figure 5 Amino acids and side chains for residues of transmembrane helices on which are located the effective hydrophobicity residues for the beta adrenergic receptor. The green regions on the helices show the direction of the effective hydrophobicity for that helix. The amino acid residue and side chains, highlighted in red, show charged residues in the helices of the protein (Arg, Lys, Asp, Glu, His). This is to illustrate that although the effective hydrophobicities do not necessarily reside on the most charged residue, equation 1 takes into account the charged residues in its determination of Θ_θ which is pointed in roughly the same direction as the polar side chains. Amino acids in green with side chains showing are those where the effective hydrophobicity Θ_θ resides specifically on a polar amino acid residue. The residues represented by blue space-filled side chains are implicated in ligand binding.

towards the interior of the protein. Residues on transmembrane regions III and IV are implicated in ligand binding. Transmembrane regions I, VI, and VII are pointed away from the interior of the protein and mostly towards other transmembrane helices (Figure 5).

The blue, space-filled side chains in Figures 4 and 5 represent the primary binding interaction between the ligand and receptor. For rhodopsin, the ligand is retinal (aldehyde) and forms a covalent Schiff base interaction with Lys296.[55] Figure 4 shows that this lysine is positioned in the hydrophilic interior, making it amenable to ligand binding; the hydrophilicity is centered on Ser298, pointed towards the interior of the protein, away from the bilayer, allowing lysine to be accessible to ligand binding. For the beta-adrenergic receptor, the binding ligands are from the alcohol family, ie, antagonists (alprenolol and propranolol) and agonists (isoproterenol and ephedrine). Asp79, Asp113, and Asn318 are implicated in binding.[56,57] These residues mostly coincide with the position of effective hydrophobicities.

One reason for the slight shift of the effective hydrophobicities from polar residues implicated in binding for both structures is probably that these residues are found at the junctions of the helices and extracellular loops, and are therefore subject to larger displacements at the alpha-carbon backbone than other residues in the helices. To test this, in the case of the adenosine A_{2A} receptor,[58] prior to using Hydro-Eff, the helical sequences were truncated by one turn at the N- and C-ends of the helix. Figure 6 shows that for all transmembrane helices, the green regions depicting the angles of highest effective hydrophilicity were pointed towards the center of the transmembrane assembly, which confirms that Equation 1 can be used to identify helical amphiphilicity and properly position helices in the GPCR transmembrane assembly. In most cases, the effective hydrophobicity was in the general direction of most polar residues on the helix, the side chains being identified in red. This indicates that the dynamic behavior of the protein at the helix-loop junctions does not identify side chain positioning with certainty, and should not be considered while using Hydro-Eff.

Figures 4–6 illustrate that while it is not necessary for polar residues to be the sole determinant of a helix's amphiphilicity. The side chains of these polar residues when

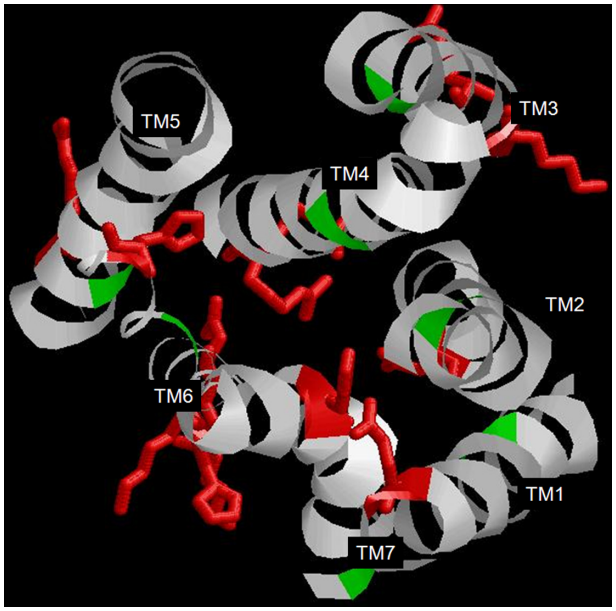

Figure 6 Amino acids and side chains for residues of transmembrane helices on which are located the effective hydrophobicity residues for the adenosine A_{2A} receptor. The green regions in the helices show the direction of the effective hydrophobicity for that helix. The amino acid residue and side chains, highlighted in red, show charged residues in the helices of the protein (Arg, Lys, Asp, Glu, His). This is to illustrate that although the effective hydrophobicities do not necessarily reside on the most charged residue, Equation 1 takes into account the charged residues in its determination of Θ_θ, and the effective hydrophobicities are pointed in roughly the same direction as the polar side chains. Amino acids in green with side chains showing are those where the Θ_θ resides specifically on a charged residue. The blue space-filled side chains are for residues that are implicated in ligard binding. The figure shows that these residues are pointed into the interior of the protein where ligand binding is likely to take place.

positioned into the interior of the protein, by virtue of their hydrophobicity values, are the highest value-contributors to the effective hydrophobicities of the helix. If polar residues are on directly opposite sides, then the amphiphilicity determined through calculations from Equation 1 help in positioning the helices. It is possible that, in addition to ligand binding, some polar residues also contribute toward stabilizing the helical bundle, thus protecting the ligand. It is also likely that, in addition to the polar residues, ligand binding is also stabilized by van der Waals interactions from the side chains of nonpolar amino acid residues. Mustafia and Palczewski have extensively assessed the binding in rhodopsin, the beta-adrenergic receptor, and the adenosine A_{2A} receptor, based on the receptor structures.[59] These are considerations that go beyond sequence and structure, ie, to specific biologic functions associated with specific GPCRs.

Using effective hydrophobicities as an aid to modeling GPCRs

As has been seen from our calculations of effective hydrophobicity for rhodopsin, the beta-adrenergic receptor, and the adenosine A_{2A} receptor, as well as the discussions in the aforegoing paragraphs, the hydrophilic side does not always face the binding region in the interior of the pocket, although the residues implicated in binding are positioned where they would be amenable to the interacting ligand.

As observations from Figures 4 and 5 indicate, the residues that have polar side chains pointed outwards are at the junction of the helices and loops either on the extracellular side or in the cytoplasm. Therefore, around the helix–loop junctions, a polar amino acid residue, the side chain of which is pointed away from the protein's core will not necessarily violate polar-nonpolar interactions by being buried in the lipid bilayer. These are the transmembrane amino acid residues that are responsible for ligand intake prior to binding, or on the cytoplasmic side, are those responsible for G-protein coupling or residues involved in catalyzing signal transduction or other intracellular processes. One example is the MAYDRY peptide sequence motif, which is found at the junction of the third transmembrane helix and the second intracellular loop. This sequence (or a sequentially similar) motif is ubiquitous for all GPCRs. The polar residues, Asp and Arg at this junction will not be in the lipid bilayer.

Figure 7 shows the transmembrane helices for rhodopsin truncated to regions that are possibly embedded in the plasma membrane. The transmembrane amphiphilicities were then re-determined using Hydro-Eff. Figure 7 was created for two reasons, ie, to assess how the helical amphiphilicity

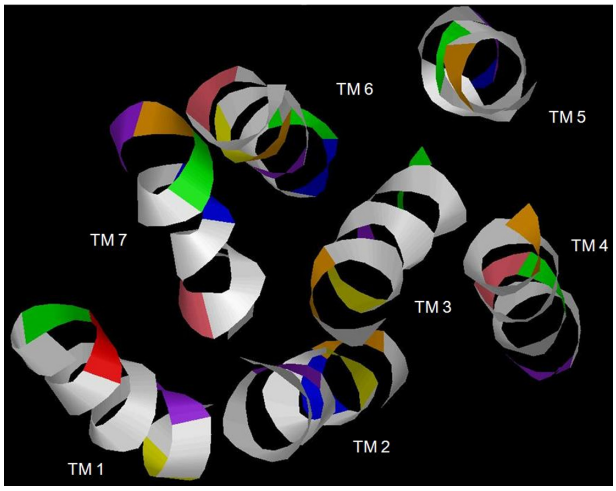

Figure 7 Structure of rhodopsin with helices truncated to regions that will be embedded in the plasma membrane. The effective hydrophobicities were then determined again using all the scales determined in Hydro-Eff®. The different colors indicate Θ_θ for different hydrophobicity scales, ie, Eisenberg consensus scales (red), Von Heinje (blue), Janin (yellow), Chothia (green), Wolfenden (pink), Kyte (orange), and Argos (purple). Despite some outliers, the figures suggest that during computational protein modeling, the helices should be rotated based on hydrophobicities of helices determined on truncated sequences that are likely to be embedded in the plasma membrane. Effective hydrophilicities are not necessarily pointed towards the interior of the protein, but might be pointed towards other helices, which might sustain the helical system.

determined by Hydro-Eff would allow the observation of helical position if residues at the helix–loop boundaries were not considered in the calculations using Equation 1 and to test how the Hydro-Eff system works for scales other than the Eisenberg consensus scales.

For each of the rhodopsin helices, Θ_θ was determined based on the hydrophobicity scales of, in addition to the Eisenberg consensus scales (red), Von Heinje (blue), Janin (yellow), Chothia (green), Wolfenden (pink), Kyte (orange), and Argos (purple). The colors represent the residues on which the effective hydrophobicity is centered.

For rhodopsin, the effective hydrophobicities using different scales do not always exactly agree. However, they correctly represent the amphiphilic nature of helices, and are pointed in the expected directions, ie, towards the hydrophilic interior. In cases where all the colors representing the different hydrophobicity scales are not observed, this is where the effective hydrophobicity Θ_θ from different scales coincide.

The transmembrane regions 1, 2, 3, 5, 6, and 7 in Figure 7 point to the interior. The effective hydrophobicities for TM4 variably point towards TM2 and TM5, depending on the scales used. While there is general consensus among the methods used to represent the amphiphilic side of the helix, there are some outliers when using different scales. In Figure 7, the effective hydrophobicities for TM1 using the Janin scales,[33] TM4 using Argos,[38] scales and TM7 using Wolfenden's [34] scales, are pointed away from the protein's interior.

Consequences for protein–ligand binding

Determining amphiphilicities is crucial to membrane protein modeling, because properly representing the interior of the protein is necessary if accurate predictions are to be made with regard to ligand binding. This is also consequential to postbinding activities, such as signal transduction. Consider the computationally-driven interaction between an odorant molecule and an olfactory receptor (which are Class A GPCRs).[19,43,47] Erroneous predictions of binding can result if the helices are not rotated correctly to maintain the hydrophilicity of the binding pocket. Equation 1 thus represents a reassessment of the paradigm that allows a protein computational modeler to position the helices.

Making predictions and inferences from computational studies should necessarily go hand in hand with experimental (mutagenesis) studies. Using computational methodologies to identify and predict putative agonists and antagonists among binding ligands have been attempted. These inferences, following static ligand docking, implicate amino acid residues that are at hydrogen bonding distances from the ligand key to binding and activation. Several computational studies, with support from experimental functional analyses, have been reported of the use of binding affinities to predict agonists versus antagonists.[25,44–46,49] Lai et al published the first study of ligand–olfactory interactions using molecular dynamic simulation studies which better mimic the dynamic biologic environment.[20] These studies have identified residues, in addition to those identified through static docking, as necessary for GPCR binding and activation.

Jaakola et al, who used the antagonist ZM241385 bound to the adenosine A_{2A} receptor to crystallize the complex, make the point that the dynamic nature of the binding pocket has to be considered before making inferences about the nature of ligand-binding.[58]

Computational docking seeks to place a ligand in a region in the protein that is least hindered and can accommodate the binding ligand. Mutagenesis studies identify residues that result in activation or inhibition, and computational studies can identify the regions that house these residues. Studies (currently submitted) led by the author point to two specific binding regions in two olfactory receptors, ie, one which preferably binds ligands that serve as activators of olfactory receptors and one which binds ligands that inhibit the receptor. These results stem from experimental functional data for these olfactory receptors.[60]

It is critical then that for any computational modeling protocol to be successful, specifically in the case of GPCRs, that the transmembrane helices be properly identified and modeled, and the biochemical nature of the protein's interior and exterior (its hydrophilicity) be properly maintained. Systems such as the Hydro-Eff are then valuable in this endeavor.

Conclusion

A formalism that quantifies the amphiphilicity of a helix is developed in this paper. For a specific angle on the helical wheel, Hydro-Eff considers hydrophobicity contributions from all other residues in the helix at the angle in question. The web page for the Hydro-Eff tool allows users to input peptide sequences for several helices and obtain tabulated results using differing hydrophobicity scales. Because each hydrophobicity scale was determined based on differing concepts and methodologies, a modeler will have to take into account the function of the protein before deciding which scale to consider. Indeed, results differ, as seen in Figure 7.

Another issue that a prospective modeler would have to consider is how much of the helix would contribute to amphiphilicity. Figures 4 and 5, when contrasted with Figures 6 and 7 in terms of amphiphilicity, indicate that Hydro-Eff predicts the hydrophilic side more effectively

if residues that border the boundary between the helices and the interhelical loops are not used in determination of Θ_θ. This would be important when positioning helices following homology modeling, especially when there is weak homology between GPCRs to be modeled and GPCRs for which structures have already been determined.

Acknowledgments

This work is partially supported by 1R21DC011068-01 (National Institute on Deafness and Other Communication Disorders, National Institutes of Health), a Faculty Development Grant awarded to the author at University of Alabama at Birmingham, University of Alabama at Birmingham Comprehensive Cancer Center, Howell and Elizabeth Heflin Center for Genome Sciences, and the CTSA University of Alabama at Birmingham Center of Clinical and Translational Science.

Disclosure

The author reports no conflict of interest in this work.

References

1. Hamm HE. The many faces of G protein signaling. *J Biol Chem*. 1998;273(2):669–672.
2. Ji TH, Grossmann M, Ji I. G protein-coupled receptors. I. Diversity of receptor-ligand interactions. *J Biol Chem*. 1998;273(28):17299–17302.
3. Gilman AG. G proteins: Transducers of receptor-generated signals. *Annu Rev Biochem*. 1987;56:615–649.
4. Engelhardt S, Rochais F. G proteins: More than transducers of receptor-generated signals? *Circ Res*. 2007;100(8):1109–1111.
5. Bjarnadottir TK, Gloriam DE, Hellstrand SH, Kristiansson H, Fredriksson R, Schioth HB. Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse. *Genomics*. 2006;88(3):263–273.
6. Glusman G, Yanai I, Rubin I, Lancet D. The complete human olfactory subgenome. *Genome Res*. 2001;11(5):685–702.
7. Gilad Y, Lancet D. Population differences in the human functional olfactory repertoire. *Mol Biol Evol*. 2003;20(3):307–314.
8. Buck L, Axel R. A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell*. 1991;65(1):175–187.
9. Palczewski K, Kumasaka T, Hori T, et al. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*. 2000;289(5480):739–745.
10. Stenkamp RE, Teller DC, Palczewski K. Crystal structure of rhodopsin: A G-protein-coupled receptor. *Chembiochem*. 2002;3(10):963–967.
11. Lai PC, Bahl G, Gremigni M, et al. An olfactory receptor pseudogene whose function emerged in humans: A case study in the evolution of structure-function in GPCRs. *J Struct Funct Genomics*. 2008;9(1–4):29–40.
12. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. *J Mol Biol*. 2004;342(2):571–583.
13. Rasmussen SG, Choi HJ, Rosenbaum DM, et al. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*. 2007;450(7168):383–387.
14. Warne T, Serrano-Vega MJ, Baker JG, et al. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature*. 2008;454(7203):486–491.
15. Hartmut M. Crystallization of membrane proteins. *Trends Biochem Sci*. 1983;8(2):56–59.
16. Sanchez R, Sali A. Comparative protein structure modeling. Introduction and practical examples with modeller. *Methods Mol Biol*. 2000;143:97–129.
17. Eswar N, Webb B, Marti-Renom MA, et al. Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci*. 2007;Chapter 2:Unit 29.
18. Tebben AJ, Schnur DM. Beyond rhodopsin: G protein-coupled receptor structure and modeling incorporating the beta2-adrenergic and adenosine A(2A) crystal structures. *Methods Mol Biol*. 2010;672:359–386.
19. Singer MS. Analysis of the molecular basis for octanal interactions in the expressed rat 17 olfactory receptor. *Chem Senses*. 2000;25(2):155–165.
20. Lai PC, Singer MS, Crasto CJ. Structural activation pathways from dynamic olfactory receptor-odorant interactions. *Chem Senses*. 2005;30(9):781–792.
21. Cole C, Barber JD, Barton GJ. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res*. 2008;36(Web Server issue):W197–W201.
22. Persson B, Argos P. Prediction of membrane protein topology utilizing multiple sequence alignments. *J Protein Chem*. 1997;16(5):453–457.
23. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J Mol Biol*. 2001;305(3):567–580.
24. Tusnady GE, Simon I. The HMMTOP transmembrane topology prediction server. *Bioinformatics*. 2001;17(9):849–850.
25. Vaidehi N, Floriano WB, Trabanino R, et al. Prediction of structure and function of G protein-coupled receptors. *Proc Natl Acad Sci U S A*. 2002;99(20):12622–12627.
26. Eisenberg D, Schwarz E, Komaromy M, Wall R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol*. 1984;179(1):125–142.
27. Eisenberg D. Three-dimensional structure of membrane and surface proteins. *Annu Rev Biochem*. 1984;53:595–623.
28. Donnelly D. Modelling alpha-helical transmembrane domains. *Biochem Soc Trans*. 1993;21(1):36–39.
29. Hanson MA, Stevens RC. Discovery of new GPCR biology: One receptor structure at a time. *Structure*. 2009;17(1):8–14.
30. Lagerstrom MC, Schioth HB. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov*. 2008;7(4):339–357.
31. Michino M, Abola E, Brooks CL 3rd, Dixon JS, Moult J, Stevens RC. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat Rev Drug Discov*. 2009;8(6):455–463.
32. Chothia C. The nature of the accessible and buried surfaces in proteins. *J Mol Biol*. 1976;105(1):1–12.
33. Janin J. Surface and inside volumes in globular proteins. *Nature*. 1979;277(5696):491–492.
34. von Heijne G, Blomberg C. Trans-membrane translocation of proteins. The direct transfer model. *Eur J Biochem*. 1979;97(1):175–181.
35. von Heijne G. Membrane proteins: The amino acid composition of membrane-penetrating segments. *Eur J Biochem*. 1981;120(2):275–278.
36. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. 1982;157(1):105–132.
37. Argos P, Rao JK, Hargrave PA. Structural prediction of membrane-bound proteins. *Eur J Biochem*. 1982;128(2–3):565–575.
38. Nozaki Y, Tanford C. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale. *J Biol Chem*. 1971;246(7):2211–2217.
39. Segrest JP, Feldmann RJ. Membrane proteins: Amino acid sequence and membrane penetration. *J Mol Biol*. 1974;87(4):853–858.
40. Wolfenden R, Andersson L, Cullis PM, Southgate CC. Affinities of amino acid side chains for solvent water. *Biochemistry*. 1981;20(4):849–855.
41. Tikhonova IG, Fourmy D. The family of G protein-coupled receptors: An example of membrane proteins. *Methods Mol Biol*. 2010;654:441–444.

42. Trabanino RJ, Hall SE, Vaidehi N, Floriano WB, Kam VW, Goddard WA 3rd. First principles predictions of the structure and function of g-protein-coupled receptors: Validation for bovine rhodopsin. *Biophys J*. 2004;86(4):1904–1921.
43. Kalani MY, Vaidehi N, Hall SE, et al. The predicted 3D structure of the human D2 dopamine receptor and the binding site and binding affinities for agonists and antagonists. *Proc Natl Acad Sci U S A*. 2004;101(11):3815–3820.
44. Hummel P, Vaidehi N, Floriano WB, Hall SE, Goddard WA 3rd. Test of the binding threshold hypothesis for olfactory receptors: Explanation of the differential binding of ketones to the mouse and human orthologs of olfactory receptor 912–993. *Protein Sci*. 2005;14(3):703–710.
45. Hall SE, Floriano WB, Vaidehi N, Goddard WA 3rd. Predicted 3-D structures for mouse I7 and rat I7 olfactory receptors and comparison of predicted odor recognition profiles with experiment. *Chem Senses*. 2004;29(7):595–616.
46. Freddolino PL, Kalani MY, Vaidehi N, et al. Predicted 3D structure for the human beta 2 adrenergic receptor and its binding site for agonists and antagonists. *Proc Natl Acad Sci U S A*. 2004;101(9):2736–2741.
47. Floriano WB, Vaidehi N, Zamanakos G, Goddard WA 3rd. Hier VLS hierarchical docking protocol for virtual ligand screening of large-molecule databases. *J Med Chem*. 2004;47(1):56–71.
48. Floriano WB, Vaidehi N, Goddard WA 3rd, Singer MS, Shepherd GM. Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor. *Proc Natl Acad Sci U S A*. 2000;97(20):10712–10716.
49. Floriano WB, Vaidehi N, Goddard WA 3rd. Making sense of olfaction through predictions of the 3-D structure and function of olfactory receptors. *Chem Senses*. 2004;29(4):269–290.
50. Floriano WB, Hall S, Vaidehi N, Kim U, Drayna D, Goddard WA 3rd. Modeling the human PTC bitter-taste receptor interactions with bitter tastants. *J Mol Model*. 2006;12(6):931–941.
51. Araneda RC, Kini AD, Firestein S. The molecular receptive range of an odorant receptor. *Nat Neurosci*. 2000;3(12):1248–1255.
52. Ladds G, Goddard A, Davey J. Functional analysis of heterologous GPCR signalling pathways in yeast. *Trends Biotechnol*. 2005;23(7):367–373.
53. Owens SL, Fitzke FW, Inglehearn CF, et al. Ocular manifestations in autosomal dominant retinitis pigmentosa with a Lys-296-Glu rhodopsin mutation at the retinal binding site. *Br J Ophthalmol*. 1994;78(5):353–358.
54. Decaillot FM, Befort K, Filliol D, Yue S, Walker P, Kieffer BL. Opioid receptor random mutagenesis reveals a mechanism for G protein-coupled receptor activation. *Nat Struct Biol*. 2003;10(8):629–636.
55. Matsumoto H, Yoshizawa T. Rhodopsin regeneration is accelerated via noncovalent 11-cis retinal-opsin complex – a role of retinal binding pocket of opsin. *Photochem Photobiol*. 2008;84(4):985–989.
56. Strader CD, Sigal IS, Register RB, Candelore MR, Rands E, Dixon RA. Identification of residues required for ligand binding to the beta-adrenergic receptor. *Proc Natl Acad Sci U S A*. 1987;84(13):4384–4388.
57. Mersmann HJ, McNeel RL. Ligand binding to the porcine adipose tissue beta-adrenergic receptor. *J Anim Sci*. 1992;70(3):787–797.
58. Jaakola VP, Griffith MT, Hanson MA, et al. The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science*. 2008;322(5905):1211–1217.
59. Mustafi D, Palczewski K. Topology of class A G protein-coupled receptors: insights gained from crystal structures of rhodopsins, adrenergic and adenosine receptors. *Mol Pharmacol*. 2009;75(1):1–12.
60. Malnic B, Hirono J, Sato T, Buck LB. Combinatorial receptor codes for odors. *Cell*. 1999;96(5):713–723.