The Transmembrane 7-α-Bundle of Rhodopsin: Distance Geometry Calculations with Hydrogen Bonding Constraints

Irina D. Pogozheva, Andrei L. Lomize, and Henry I. Mosberg
College of Pharmacy, University of Michigan, Ann Arbor, Michigan 48109 USA

ABSTRACT A 3D model of the transmembrane 7-α-bundle of rhodopsin-like G-protein-coupled receptors (GPCRs) was calculated using an iterative distance geometry refinement with an evolving system of hydrogen bonds, formed by intramembrane polar side chains in various proteins of the family and collectively applied as distance constraints. The α-bundle structure thus obtained provides H bonding of nearly all buried polar side chains simultaneously in the 410 GPCRs considered. Forty evolutionarily conserved GPCR residues form a single continuous domain, with an aliphatic "core" surrounded by six clusters of polar and aromatic side chains. The 7-α-bundle of a specific GPCR can be calculated using its own set of H bonds as distance constraints and the common "average" model to restrain positions of the helices. The bovine rhodopsin model thus determined is closely packed, but has a few small polar cavities, presumably filled by water, and has a binding pocket that is complementary to 11-*cis* (6-*s*-*cis*, 12-*s*-*trans*, C=N *anti*)-retinal or to all-*trans*-retinal, depending on conformations of the Lys^{296} and Trp^{265} side chains. A suggested mechanism of rhodopsin photoactivation, triggered by the *cis*-*trans* isomerization of retinal, involves rotations of Glu^{134}, Tyr^{223}, Trp^{265}, Lys^{296}, and Tyr^{306} side chains and rearrangement of their H bonds. The model is in agreement with published electron cryomicroscopy, mutagenesis, chemical modification, cross-linking, Fourier transform infrared spectroscopy, Raman spectroscopy, electron paramagnetic resonance spectroscopy, NMR, and optical spectroscopy data. The rhodopsin model and the published structure of bacteriorhodopsin have very similar retinal-binding pockets.

INTRODUCTION

The rhodopsin-like G protein-coupled receptors (GPCRs) are a family of several hundred integral membrane proteins that transduce chemical and optical signals across the cellular membrane (Watson and Arkinstall, 1994). Topographic mapping of GPCRs using lectin binding, limited proteolysis (Martynov et al., 1983; Yarden et al., 1986; Dohlman et al., 1987), chemical modification (Barclay and Findlay, 1984), insertional mutagenesis (Borjigin and Nathans, 1994), and theoretical studies with different hydrophobicity or propensity scales (Hargrave et al., 1984; Baldwin, 1993; Jones et al., 1994; Persson and Argos, 1994; Rost et al., 1995) consistently identify seven transmembrane α-helices in amino acid sequences of GPCRs and place the N-terminus and C-terminus on the extracellular and intracellular sides of the membrane, respectively. The 7-α-bundle organization of GPCRs has been directly demonstrated by electron cryomicroscopy (EM) studies of bovine, frog, and squid rhodopsins (Schertler et al., 1993, 1995, 1996; Unger and Schertler, 1995; Unger et al., 1995; Schertler and Hargrave, 1995; Davies et al., 1996). The transmembrane helical segments, identified in amino acid sequences of 204 GPCRs, have been tentatively assigned to the density peaks visible in EM maps of rhodopsins (Baldwin, 1993). The assignment, further supported by recent mutagenesis and cross-linking studies (Suryanarayana et al., 1992; Zhang et al., 1994; Pittel and Wess, 1994; Rao et al., 1994; Zhou et al., 1994; Sealfon et al., 1995; Yu et al., 1995; Liu et al., 1995; Elling et al., 1995; Mizobe et al., 1996; Thirstrup et al., 1996), proposes that the seven transmembrane α-helices are sequentially connected in a counterclockwise direction when viewed from the extracellular side, as in bacteriorhodopsin, a bacterial, light-driven proton pump (Henderson et al., 1990), which has no detectable sequence homology with GPCRs (Soppa, 1994) and which differs from rhodopsin in the tilts and positions of some helices (Unger and Schertler, 1995; Schertler and Hargrave, 1995).

Although the extent of structural similarity between bacteriorhodopsin and rhodopsins remains uncertain, it is expected, based on the presence in each transmembrane α-helix of characteristic residues that are conserved throughout the GPCR family (Baldwin, 1993), that all rhodopsin-like GPCRs themselves share a common spatial structure. For remotely related receptors with ~20% sequence identity within the seven transmembrane helices, the expected root mean square deviation (r.m.s.d.) of main-chain atoms within the α-helical core may be estimated as 1.6–2.3 Å by using a calibration curve relating the coordinate r.m.s.d. and sequence identity for proteins with known 3D structures (Chothia and Lesk, 1986).

The resolution currently achieved by EM of rhodopsins (6–7 Å) is insufficient to obtain atomic-level structure. The development of computational methods for docking of the α-helices, using their spatial arrangement identified by EM, provides an alternative approach. Modeling of the transmembrane 7-α-bundle, excluding surface loops, is simplified by identification of residues that are evolutionarily

Received for publication 12 November 1996 and in final form 4 February 1997.

Address reprint requests to Dr. Henry I. Mosberg, College of Pharmacy, University of Michigan, Ann Arbor, MI 48109-1065. Tel.: 313-764-8117; Fax: 313-763-5633; E-mail: him@umich.edu.

0006-3495/97/05/1963/23 $2.00

conserved or hydrophilic, or which are important for folding or ligand binding of GPCRs. These residues should form the protein interior, thus defining a lipid-inaccessible surface of the transmembrane helices and placing a restriction on the rotational orientation of each helix and the depth of its immersion into the α-bundle (Donnelly et al., 1993; Baldwin, 1993; Taylor et al., 1994).

A number of approximate GPCR models have been built from seven rigid α-helices, with arbitrarily chosen side-chain conformers, using the EM maps, restrictions on rotational orientations of helices, and a few experimental constraints derived from mutagenesis and chemical cross-linking data (see, for example, Baldwin, 1993; Donnelly et al., 1994; Lin et al., 1994; Kim et al., 1995; Herzyk and Hubbard, 1995; and the review of Donnelly and Findlay, 1994). These "crude" 7-α-bundle models, although imprecise and different from each other, assist in planning experiments that can be helpful for refining the models further. At the same time, the analysis of nonadditive mutational effects or ligand-binding studies of various receptor mutants does not provide precise constraints on specific interatomic distances, and only qualitatively indicates the proximity of residues or their tentative involvement in the ligand-binding pocket. Moreover, to determine the atomic structure, a set of such constraints must be sufficient to define side-chain conformers and to reproduce the unique geometries of transmembrane helices that are kinked by Pro residues and curved, as is normally observed in proteins (Barlow and Thornton, 1988).

An alternative method of identifying the required constraints, applied here, is based on the presence of numerous polar residues in the transmembrane hydrophobic α-helices of GPCRs. It is known that polar side chains of proteins buried from water have a strong tendency to form H bonds (McDonald and Thornton, 1994). In transmembrane α-helices, backbone peptide groups are already paired, whereas the polar side chains must interact with each other to form intra- or interhelical H bonds. The candidate H-bonding pairs can be identified from the analysis of sequence alignments as polar residues, in intramembrane segments, which appear and disappear simultaneously in various GPCRs, and by using approximate receptor models to exclude all spatially distant residues from the list of possible correlations. H bonds thus identified can be applied as distance constraints for the packing of the transmembrane α-helices, using the distance geometry algorithm. Because the rhodopsin-like GPCRs share a common 3D structure of the transmembrane domain, the side-chain H bonds from many different GPCRs can be combined to increase the number of simultaneously applied constraints and to calculate an "average" 7-α-bundle structure. The distance geometry calculations serve to check the self-consistency of the constraints, to estimate the uncertainty of a model, and to allow some flexibility of α-helix geometry, important for close packing of side chains (Walther et al., 1996).

The computational procedure applied here was organized as an iterative structural refinement with evolving constraints that begins from an initial "crude" rhodopsin model and continues until each buried polar side chain from each of the 410 GPCRs considered can participate in at least one hydrogen bond in the final structure of the 7-α-bundle. This H-bond saturation criterion was very sensitive to mistakes in the initial and hundreds of intermediate 7-α-bundle structures. The transmembrane segments of individual GPCRs are hydrophobic and usually contain no more than 30% polar residues, but when 410 different amino acid sequences are simultaneously considered, all interhelical contacts within the α-bundle are "labeled" by hydrophilic side chains forming intramolecular H bonds, usually in a group of related receptors. Displacement of any α-helix from its correct position breaks some H bonds, producing unpaired polar side chains in tens or hundreds of GPCRs. This approach is somewhat similar to the self-correcting distance geometry algorithm, developed recently for approximate docking of predefined helices (Hänggi and Braun, 1994; Mumenthaler and Braun, 1995), although our application is only a refinement of an already known approximate model, rather than a priori structural calculations.

The "average" receptor model can be tested by using it as a template to calculate 7-α-bundles of various GPCRs whose specific H bonds and close packing of nonpolar side chains must be compatible with the same common structure. Each GPCR must create a binding pocket complementary to its natural and artificial ligands and be consistent with experimental data. Although we have calculated α-bundles for six dissimilar GPCRs (bovine and crayfish rhodopsins, human δ opioid, lutropin/choriogonadotropin hormone, muscarinic acetylcholine, and ATP receptors) and for squid retinochrome (Hara-Nishimura et al., 1993), we report here only the results for bovine rhodopsin, because this receptor has been so extensively studied that any significant structural flaws in its model would be immediately detected as discrepancies with some physicochemical data.

METHODS

The modeling described here consists of the following stages: 1) construction of the initial, "crude" bovine rhodopsin model using EM and a few mutagenesis and cross-linking data; 2) calculation of the common ("average") seven-helix bundle model for rhodopsin-like GPCRs, by using an iterative distance geometry refinement of the initial model with an evolving system of interhelical side-chain H-bonds formed by various GPCRs and collectively applied as distance constraints; and 3) distance geometry calculations of the transmembrane domain for bovine rhodopsin from its own H bonds and using the "average" GPCR model to restrain the relative positions of the helices.

Amino acid sequences of GPCRs

At the initial stage of this study, four sequence alignments of rhodopsin-like GPCRs (opsins and protein, peptide, and cationic amine receptor subfamilies), received via file server TM7@EMBL-Heidelberg.DE, release 7/6/94 (Oliveira et al., 1993), were used to identify conserved positions in transmembrane helices (Fig. 1) and H-bonding constraints. Obvious misalignments in the data base, such as misplaced groups of residues that are conserved throughout the GPCR family, or gaps within transmembrane

```
HELIX I    38-63
OPSIN              ..L...M......G...N..V.....
PROTEIN            ..................N..V.....
PEPTIDE            ......Y......G..GN..V.....
AMINE              ..........L.T..GN.LV..A..
CONSENSUS          ......M......G..GN..V.....

HELIX II   70-95
OPSIN              .P.N..L.NLA..D............
PROTEIN            .........LA.AD............
PEPTIDE            ......I.NLA.AD............
AMINE              ...N....SLA.AD............
CONSENSUS          ...N..L.NLA.AD............

HELIX III  111-136
OPSIN              ...G......G....WS...I..ERY
PROTEIN            ............S...L..L..DRY
PEPTIDE            K...........S...L...S.DRY
AMINE              ..W...DV...TASI..LC.IS.DRY
CONSENSUS          ............S...L..IS.DRY

HELIX IV   151-176
OPSIN              ..A...I.F.W........PP..GWS
PROTEIN            ..........W...............
PEPTIDE            ..........W..A.....P......
AMINE              ......I...W..S...S.P......
CONSENSUS          ......I...W..A.....P......

HELIX V    202-227
OPSIN              ............P...I...Y....
PROTEIN            ..........F.........CY....
PEPTIDE            ..........F..P...I...Y..I.
AMINE              .....S...SFY.P...M...Y..I.
CONSENSUS          ..........F..P...I...Y..I.

HELIX VI   250-275
OPSIN              ..R.V......F...W.PY...A...
PROTEIN            ..........F..C..P.........
PEPTIDE            ........V..F..CW.P........
AMINE              A...L..I.G.F..CW.PFF......
CONSENSUS          ....V......F..CW.PY.......

HELIX VII  286-311
OPSIN              .....PA.FAK....YNP.IY...N.
PROTEIN            ............NS. DP.IY.....
PEPTIDE            ........LA..N.C.NP..Y.....
AMINE              .......WLGY.NS..NP.IY...N.
CONSENSUS          ........LAK.NS..NP.IY...N.
```

FIGURE 1 Conserved residues in four subfamilies (57 opsin, 54 protein, 114 peptide, and 122 amine receptor sequences) of rhodopsin-like GPCRs. The transmembrane segments are shown as proposed by Baldwin (1993). The "Consensus" string indicates positions that are conserved in at least two subfamilies. A position was considered as conserved if the occurrence of one specific residue in a sequence alignment was >66%, or the total occurrence of two residues (for example, Ala and Ser) was >90%. Non-conservative positions are indicated by dots. Numbering of all residues is that for the equivalent positions in bovine rhodopsin.

α-helices, were corrected. All duplicated sequences and sequences with an identity of <22% (for amine GPCRs, <24%) compared with the common sequence pattern for the corresponding subfamily (Oliveira et al., 1993) were excluded from the original alignments. As a result, the opsin, protein, peptide, and amine receptor subfamilies included 57, 54, 114, and 122 unique sequences, respectively. Later, a more recent release (2/25/95) of the data base with 57 opsin, 215 peptide and protein, and 122 amine GPCR sequences, and including 15 purine receptors and squid retinochrome (a total of 410 sequences), was employed to verify receptor models for saturation of H bonding potential using the program ADJUST.

Transmembrane α-helices

In the modeling approach applied here, the structure of the transmembrane α-bundle is established by the interhelical H-bonds. Therefore, the precise identification of the helix ends in the amino acid sequence was not crucial. The locations of helices (Fig. 2) were calculated using a thermodynamic model of α-helix formation, previously verified for peptide-micelle complexes and water-soluble proteins (Lomize and Mosberg, manuscript submitted for publication) and modified for transmembrane proteins. The model combines free energy terms defining α-helix stability in aqueous solution (derived mostly from protein engineering data) and terms describing immersion of every helical or coil fragment in a spherical micelle, a planar bilayer, or a nonpolar droplet created by the rest of the protein to calculate, using the dynamic programming algorithm, the lowest energy partitioning of the peptide chain into helical and coil fragments.

The calculated locations of helices vary slightly in different GPCRs, reflecting the inaccuracy of the computational model, which systematically underestimates lengths of helices in α-bundle proteins (Lomize and Mosberg, manuscript submitted for publication). To obtain more reliable results, the lowest energy helix-coil partitions were calculated independently for 57 rhodopsins, and the fragments of rhodopsin sequence that are calculated to be helical in more than 75% of sequences were considered as transmembrane helices of GPCRs. All α-helices identified in this way, except helix V, are slightly longer than the hydrophobic intramembrane portions identified by Baldwin (1993). Furthermore, the calculations identified two additional amphiphilic helices (226–235 and 311–320) that may be present independently on the intracellular side of the α-bundle, or may continue transmembrane helices V and VII, respectively. (Numbering of residues in this article corresponds to the amino acid sequence of bovine rhodopsin.) These two smaller helices were not included in our model. Recent site-directed spin-labeling studies (Altenbach et al., 1996) indicate that the C-terminal part of helix V, the shortest helix in our model, could be extended up to residue 237, whereas the fragment 240–248 may form an amphiphilic extension of helix VI (in our model, transmembrane helix VI begins at residue 247). On the other hand, NMR studies of the 231–252 peptide demonstrate the presence of helix in fragment 239–242 (Yeagle et

```
HELIX I          ********** ********** *
Rbv     PWQFSML40AAYMFLLIML50GFPINFLTLY60VTVQHK
Avr     ---YN--  S--Y---T--  S---------  --Q---

HELIX II         ********* ********** ***
Rbv           PLNYILLNLA80VADLFMVFGG90FTTTLYTS
Avr           ----Y-K--S  ---HCEC--C-  ------DQ

HELIX III        ********  **********  ****
Rbv           NLEGFFATLG120GEIALWSLVV130LAIERYVVVC140K
Avr           ---THC-CYC  CCSS-C--CC-  -----C----  -

HELIX IV         *********  **********  **
Rbv         E150NHAIMGVAFT160WVMALACAAP170PLVGWSRY
Avr         -  ------C---  ---CT-WCT-  -V--H---

HELIX V          ********  **********  ***
Rbv        N200ESFVIYMFVV210HFIIPLIVIF220FCY
Avr        -  ----YDCC-W  CCY----C--  ---

HELIX VI         *********  **********  **
Rbv      EKEV250TRMVIIMVIA260FLICWLPYAG270VAFYIFT
Avr      ----  --T---T-Q-  DC-SMT--N-  HNW----

HELIX VII        ***  **********  ********
Rbv          PIFMTI290PAFFAKTSAV300YNPVIYIMMN310
Avr          --NKH-  -YC-H-SNSC  H-------N-
```

FIGURE 2 Amino acid sequences used for calculations of the 7-α-bundle structure of bovine rhodopsin (Rbv) and the final "average" GPCR model (Avr). In the Avr sequence, a dash indicates the same residue as in Rbv sequence. The intramembrane segments 21–22 residues in length, which were examined for saturation of "H-bonding potential" of polar side chains, are indicated by an asterisk.

al., 1995a). Additional analysis of this intracellular loop connecting helices V and VI is necessary to clarify the situation. It is also noteworthy that NMR studies of the C-terminal rhodopsin peptide (316–348) revealed an α-helical (i, i + 4) nuclear Overhauser effect (Yeagle et al., 1995b) that corresponds to the last turn of the calculated helix 311–320.

Initial model of bovine rhodopsin

The initial "crude" 7α-bundle model of bovine rhodopsin served only as a starting point for refinement and was very different from the final structure (r.m.s.d. of C^{α} atoms is 4.1 Å). The model was built with QUANTA from standard α-helices ($\varphi = -58°$, $\psi = -48°$) using arbitrary side-chain conformers acceptable in α-helices. The kinks induced by proline residues in helices I (Pro^{53}), IV (Pro^{170}, Pro^{171}), V (Pro^{215}), VI (Pro^{267}), and VII (Pro^{303}) were reproduced using helix 3 of the L subunit of the bacterial photoreaction center (1prc Protein Data Bank (PDB) file) as a structural template. The helices were arranged to qualitatively reproduce published 3D EM maps of bovine and frog rhodopsins (Schertler et al., 1993, 1995), using the proposed assignment of the helices to the maps (Baldwin, 1993), with the rotational orientations of helices determined by their inner, conserved residues (Fig. 1).

Next, the helices were shifted along their longitudinal axes to satisfy a few reliable experimental data and to maximize contacts between conserved polar GPCR residues (Fig. 1), which are expected to form clusters as in other proteins (Altschuh et al., 1987). First, helices II, III, and VII, which are situated in the middle of the α-bundle, were arranged relative to each other. The conserved polar residues, which are in the middle of these helices (Asp^{83}, Ser^{124}, Asn^{298}, Ser^{299}, and Asn^{302}) or close to their intracellular ends (Asn^{73}, Glu^{134}, Arg^{135}, and Tyr^{306}), were merged to form two corresponding polar clusters, simultaneously providing the experimentally demonstrated proximity of Lys^{296} (helix VII) with Gly^{90} (helix II) (Rao et al., 1994), Glu^{113} (helix III) (Sakmar et al., 1989; Zhukovsky and Oprian, 1989; Nathans, 1990b), and Ala^{117} (helix III) (Zhukovsky et al., 1992). Helix I was then linked to helix II by an H bond between the conserved Asn^{55} and Asp^{83} residues, which also provides the proximity of residues 44 and 289, as suggested from mutagenesis studies of muscarinic receptors (Liu et al., 1995). Helix V was linked to helix III by a Glu^{122}-His^{211} H bond, because appearance of these residues is highly correlated in the rhodopsin subfamily. Helix VI was aligned to place Trp^{265} at approximately the same distance from the membrane plane as Lys^{296} of helix VII, because the β-ionone ring of retinal can be cross-linked with Trp^{265} (Nakayama and Khorana, 1990) and retinal, which forms a Schiff base (SB) with Lys^{296}, is approximately parallel to the membrane plane (Liebman, 1962).

The "average" 7-α-bundle model of rhodopsin-like GPCRs

Distance geometry refinement with evolving constraints

The common ("average") 7-α-bundle structure of GPCRs was calculated starting from the initial rhodopsin model via 528 iterations of a refinement procedure. The α-bundle was built primarily from residues of bovine rhodopsin; however, many nonpolar side chains were replaced by polar ones to incorporate H bonds from other GPCRs (Fig. 2, Table 1). Each iteration of refinement included 1) examination of the structures calculated in the previous iteration for incomplete "H-bond saturation" in 410 GPCRs and other structural flaws (violations of constraints, appearance of hindrances or holes produced by incorrectly packed side chains, helices that are multiply curved by contradictory constraints or are loosely packed because of insufficient constraints); 2) modification of distance and angle constraints (H bonds and conformers of side chains) to correct these flaws, and 3) distance geometry calculations with the modified constraints using the distance geometry program DIANA (Güntert et al., 1991). The constraints and the corresponding α-bundle structure simultaneously evolved during the refinement, and as a result, the spatial positions of transmembrane helices were substantially changed: the r.m.s.d. of C^{α} atoms between the initial and final structures is 4.1 Å. The final structure and constraints (Tables 1–4) are described under Results and Discussion.

The structural flaws and lipid-exposed polar side chains, detected in the initial and intermediate models, were eliminated in a stepwise fashion, sequentially improving different miscalculated regions of the α-bundle, starting from the middle of the membrane and from the central four helices, II, III, VI, and VII. To improve the structure by changing the distance and angle constraints at each iteration, it was necessary to identify "unsatisfied" H bonds, correlations, possible alternative H bonds, and the corresponding conformers of all polar side chains throughout all 410 GPCR sequences considered. This was done using the program ADJUST, described below. Then possible H bonds and side-chain conformers were analyzed visually using the molecular modeling software QUANTA (Molecular Simulations) to determine which are sterically allowed and mutually consistent in the current iteration of the calculated structural model. Because the spatial positions of entire helices were imprecise in the intermediate models, the helices were sometimes translated (by 1–2 Å) or rotated (by ~20–30°) to better examine the possible alternative H bonds, especially in the early steps of refinement.

The search for the proper side-chain conformers and H bonds was guided primarily by correlations in multiple sequence alignments. Depending mostly on the χ^1 angle, side-chain conformers can interact with different groups of residues, and the conformer providing the highest correlation index Q (Eq. 1 below) can be chosen. The correlations are evident when a polar residue appears in a position of a sequence that is usually occupied by nonpolar residues, or when a bulky aromatic residue replaces a smaller one. These replacements are adopted in GPCRs by the concomitant appearance in the vicinity of these changed residues of additional polar side chains, or of side chains of smaller volume, respectively. The identification of the proper H bonds was simplified by considering first the most hydrophobic GPCRs with the minimum number of polar side chains in hydrophilic clusters. The H bonds and side-chain conformers that provided better correlations and appeared realistic from visual analysis with QUANTA, were tested further by calculations with DIANA. The examination of many alternative H bonds and side-chain conformers in different regions of the α-bundle required >500 iterations.

Identification of side-chain conformers, included as the corresponding dihedral angle constraints (Table 4), was necessary to calculate a well-defined structure of the α-bundle. The conformers were identified by using 1) sequence correlations, 2) H bonds, and 3) packing requirements. When the side-chain conformers were improperly chosen or undefined, the distance geometry calculations usually produced structures with inner holes. The conformers that fill these spaces were identified by visual analysis of the improper structures with QUANTA and checked again by distance geometry. Side-chain conformers were considered to be identical throughout the GPCR family. Conformers of some aliphatic side chains could then be determined by comparison with polar or aromatic side chains, in the same position in other GPCRs, the orientations of which could be defined more easily. Nevertheless, conformers of some side chains situated at the lipid-exposed surface of the α-bundle were not uniquely defined, and calculations with their different conformers produced nearly identical structures. Some aromatic side chains of this type were oriented to interact with the main chain or with other aromatic side chains from the closest helix in the α-bundle, and the lipid-facing Tyr and Trp side chains near the ends of α-helices were directed toward the water-membrane interface. All Val and Ile side chains were fixed in their only allowed α-helix conformer, with $\chi^1 \approx 180°$ and $-60°$, respectively. Most Ser and Thr residues had $\chi^1 \approx -60°$, which provides formation of H bonds between their $O^{\gamma}H$ groups and the main chain of their own α-helices (Baker and Hubbard, 1984).

ADJUST

The program ADJUST first identifies all residue pairs that are spatially close in a model and calculates a "correlation index", Q, for them (Eq. 1). To identify all of the spatially close residues and their possible H bonds, the

TABLE 1 Final system of side-chain H bonds applied as distance constraints for calculation of the "average" α-bundle model*

Atom 1	Atom 2	$N^{\#}$	Atom 1	Atom 2	$N^{\#}$
Asn^{38} $H^{\delta22}$	Gln^{98} $O^{\epsilon1}$	8	Lys^{141} $H^{\zeta1}$	Glu^{150} $O^{\epsilon1}$	16
Ser^{41} O^{γ}	Gln^{98} $H^{\epsilon22}$	16	Trp^{167} $H^{\epsilon1}$	Asp^{206} $O^{\delta1}$	1
Tyr^{43} H^{η}	Ser^{297} O^{γ}	22	His^{175} $H^{\epsilon2}$	Glu^{201} $O^{\epsilon1}$	6**
Ser^{51} H^{γ}	Asn^{55} $O^{\delta1}$	89	Ser^{202} O^{γ}	Trp^{273} $H^{\epsilon1}$	6
Asn^{55} $H^{\delta21}$	Ser^{80} O^{γ}	31	Tyr^{205} H^{η}	Asn^{272} $O^{\delta1}$	13
Asn^{55} $H^{\delta22}$	Asp^{83} $O^{\delta1}$	376	Asp^{206} $O^{\delta2}$	Trp^{210} $H^{\epsilon1}$	1
Asn^{55} $O^{\delta1}$	His^{84} $H^{\delta1}$	4	Tyr^{213} H^{η}	Thr^{266} $H^{\gamma1}$	27§
Gln^{63} $O^{\epsilon1}$	Lys^{77} $H^{\zeta1}$	10	Tyr^{213} H^{η}	Asn^{269} $O^{\delta1}$	20
His^{65} $H^{\delta1}$	Asn^{310} $O^{\delta1}$	38	Tyr^{223} O^{η}	Arg^{252} H^{ϵ}	120
Asn^{73} $O^{\delta1}$	Asn^{310} $H^{\delta22}$	156	Thr^{253} $H^{\gamma1}$	Asn^{309} $O^{\delta1}$	6
Asn^{78} $O^{\delta1}$	Tyr^{119} H^{η}	18	Thr^{257} $O^{\gamma1}$	Tyr^{306} H^{η}	12
Asn^{78} $H^{\delta21}$	Ser^{127} O^{γ}	72	Gln^{260} $H^{\epsilon21}$	Ser^{264} O^{γ}	3
Asn^{78} $O^{\delta1}$	Trp^{161} $H^{\epsilon1}$	193	Gln^{260} $O^{\epsilon1}$	His^{301} $H^{\delta1}$	3
Asp^{83} $O^{\delta2}$	Cys^{120} H^{γ}	74	Asp^{261} $O^{\delta1}$	His^{295} $H^{\epsilon2}$	1
Asp^{83} $O^{\delta2}$	Ser^{124} O^{γ}	316§	Asp^{261} $O^{\delta2}$	Asn^{298} $H^{\delta21}$	19
Asp^{83} $H^{\delta2}$	Asn^{302} $O^{\delta1}$	338¶	His^{271} $N^{\epsilon2}$	Asn^{287} $H^{\delta22}$	5
Glu^{86} $O^{\epsilon1}$	Ser^{299} H^{γ}	12§	Asn^{272} $H^{\delta21}$	Asn^{287} $O^{\delta1}$	5
Thr^{94} O^{γ}	His^{289} $H^{\delta1}$	13	His^{295} $H^{\delta1}$	Ser^{299} O^{γ}	1
Tyr^{96} H^{η}	Asn^{111} $O^{\delta1}$	15	Asn^{298} $O^{\delta1}$	Asn^{302} $H^{\delta22}$	256
Asp^{97} $O^{\delta1}$	Lys^{288} $H^{\zeta1}$	18§	Asn^{302} $H^{\delta21}$	Tyr^{306} O^{η}	344
Glu^{113} $O^{\epsilon2}$	Lys^{296} $H^{\zeta1}$	37§	Tyr^{44} H^{η}	Cys^{87} O	127
Thr^{114} $O^{\gamma1}$	His^{175} $H^{\delta1}$	11	Thr^{48} $H^{\gamma1}$	Cys^{87} O	17
His^{115} $H^{\epsilon2}$	Thr^{169} $O^{\gamma1}$	4	Tyr^{75} H^{η}	Ala^{153} O	72
Ser^{123} H^{γ}	Thr^{165} $O^{\gamma1}$	26‖	His^{84} $N^{\epsilon2}$	Phe^{52} O	4
Ser^{124} H^{γ}	Asn^{302} $O^{\delta1}$	306	Glu^{86} $H^{\epsilon2}$	Ala^{117} O	22
Glu^{134} $O^{\epsilon1}$	Arg^{135} H^{ϵ}	82	Tyr^{268} H^{η}	Asn^{287} O	50
Glu^{134} $O^{\epsilon2}$	Arg^{135} $H^{\eta12}$	82	Asn^{298} $H^{\delta22}$	Thr^{257} O	292
Arg^{135} $H^{\eta21}$	Glu^{247} $O^{\epsilon1}$	209			

*The upper distance constraints were 1.9 Å for H· · ·O, 2.9 Å for O· · ·O, and N· · ·O and 2.6 Å for H· · ·S bonds. In addition, 10 threonine (58, 92, 93, 94, 160, 165, 169, 251, 253, 266) and 7 serine (41, 80, 124, 176, 264, 297, 299) side chains, with $\chi^1 = -60°$, form H bonds with backbone carbonyls of their own helices ($O^{\gamma}H_i \cdots O{=}C_{i-4}$), and Thr^{257} with $\chi^1 = +60°$ forms an H bond with the backbone carbonyl of residue 254.

#*N* indicates the number of sequences (from a total of 410) in which the indicated pair of residues occurs.

§The corresponding constraint was systematically violated; therefore the H-bond distance was increased by 0.3 Å.

¶The H-bond distance was increased by 0.7 Å.

‖Present as Thr^{123}-Ser^{165} pair in 26 GPCRs.

**Present as His^{175}-Gln^{201} pair in 6 GPCRs.

program then places all residue pairs k and l (k and l denote types of residues, such as Ala, Pro, Asn, and so on), which are present in each pair of *i* and *j* positions in a sequence alignment (*i* and *j* are the numbers of residues), into the corresponding spatial *i* and *j* positions of the current iteration of the model and calculates distances between all donor and acceptor groups (or any atoms for nonpolar side chains) of k and l residues in all possible combinations of their side-chain conformers taken from an α-helix rotamer library (Blaber et al., 1994). The side-chain conformers considered are limited by the current file of dihedral angle constraints, to provide consistency of side-chain conformers throughout all GPCR sequences. The possible H-bond distances were limited by a distance cutoff, which must be greater than the standard H-bond length (1.9 Å), because the helix positions may vary slightly in different GPCRs and the spatial positions of the donor and acceptor groups are only approximated by the standard side-chain conformers (the standard χ angles of side chains can be altered to form the H bonds). During the refinement procedure, as the model became more precise, the distance cutoff was decreased gradually from 6.5 to 4.0 Å. This examination of possible H bonds also identifies all polar side chains from transmembrane α-helical segments that form no H bonds, thus indicating flaws in the model.

The correlation index, $Q_{k,l}$ (i, j), is simply the number of times, $n_{k,l}$ (i, j), residues k and l occupy positions *i* and *j* simultaneously, i.e., in the same sequence, divided by the total number of times residue k occupies position *i* (and any residue occupies position *j*) and residue l occupies position *j* (and any residue occupies position *i*):

$$Q_{k,l}(i,j) = \frac{n_{k,l}(i,j)}{\sum_{m\neq l}^{20} n_{k,m}(i,j) + \sum_{m\neq k}^{20} n_{m,l}(i,j) + n_{k,l}(i,j)}. \quad (1)$$

The index $Q_{k,l}$ (i, j) is close to zero if different residues are possible in position *i* of the sequence alignment when residue l is in position *j*. The index $Q = 1$ (the maximum possible correlation) for residues with identical "conservation patterns," using the definition of Altschuh et al. (1987). If a residue k (for example, Glu) appears only once in position *i* of the sequence alignment, and residue l (for example, Lys) appears only once in position *j*, but simultaneously with residue k, i.e., in the same amino acid sequence, then $Q_{k,l}$ $(i, j) = 1$, which may be due to the formation of a Glu· · ·Lys ion pair in the single representative of the family. If both residues k and l are present in positions *i* and *j* in 100% of the sequences, then the $Q_{k,l}$ $(i, j) = 1$, again.

In practice, the correlations often were not pairwise or had low *Q* indices, because many conserved, highly polar residues (such as Arg, Lys, Asn, Gln, Asp, or Glu) often can form several H bonds with surrounding, less conserved and less polar residues (such as Ser or Thr), which can occupy various spatial positions and substitute for each other as H-bonding participants. For example, the conserved Asn^{55} residue forms its "main" H bond with Asp^{83} in nearly all GPCRs, and a "supplementary" H bond with His^{84} only in four frog rhodopsins. Therefore, the index is low for the pair Asn^{55}· · ·His^{84}. Another example is Glu^{122}, which forms an H bond with

His^{211} in rhodopsins but switches to the adjacent Asn^{212} residue in glycoprotein receptors. As a result, the index for the $Glu^{122}\cdot\cdot\cdot His^{211}$ pair was significantly lower (0.43) for the entire GPCR family than in the smaller subset of 57 opsins (0.86). Such situations were explored by displaying all combinations of amino acid residues present in all sequence alignments for a set of several spatially close positions.

DIANA

In calculations with DIANA, the α-helix geometry was restrained by 167 backbone H bonds (upper limits for $NH_i\cdot\cdot\cdot O{=}C_{i-4}$ distances = 1.9 Å, except those broken by Pro residues) and by 504 dihedral angle constraints ($\varphi = -70°$ to $-50°$, $\psi = -50°$ to $-30°$). Because the program requires a single chain, the loops connecting α-helices were approximated by Gly_n fragments, with the number of Gly residues corresponding to the minimum length of the loop in rhodopsin-like GPCRs (4, 11, 9, 12, 12, and 7 Gly residues following helices I, II, III, IV, V, and VI, respectively). The standard target function minimization strategy (Güntert et al., 1991) was used for calculations. The weighting factors for upper and lower distance limits, van der Waals, and angle constraints initially were 1, 1, 0.6, and 20, respectively, and 1, 1, 2.0, and 5 by the two last iterations. One to three structures with the lowest target function from a total of 50–150 trial structures were used for analysis with QUANTA and ADJUST at each iteration of the refinement.

Distance geometry calculations of the rhodopsin α-bundle

The structure of bovine rhodopsin was calculated using its specific set of H bonds and constraints on $C^\beta\cdot\cdot\cdot C^\beta$ distances taken from the "average" GPCR structure to restrain spatial positions of the helices in the α-bundle. Approximately 50 additional refinement iterations were necessary to examine possible H bonds and conformers of flexible side chains that were not present or had no alternative orientations in the "average" receptor model. During the refinement, χ^1 conformers of some rhodopsin side chains (positions 223, 288, and 294) were changed from *trans* to *gauche*$^-$, and for some others (positions 37, 46, 47, 49, 65, 131, 151, 216, 253, 257, 262, 264, and 265), from *gauche*$^-$ to *trans*. The final calculations were done using 16 interhelical and 20 intrahelical H bonds, 357 $C^\beta\cdot\cdot\cdot C^\beta$ constraints from the "average" model, and 301 side-chain dihedral angle constraints. The upper $C^\beta\cdot\cdot\cdot C^\beta$ limits were calculated as the corresponding average $C^\beta\cdot\cdot\cdot C^\beta$ distances in the 10 best DIANA-generated structures of the "average" model, plus a deviation of 1 Å. A deviation of 0.5 Å was used for C^β distances of the more loosely packed helix I. The buried Asp^{83}, Glu^{122}, and His^{211} residues were considered to be uncharged; the uncharged (protonated) states of Asp^{83} and Glu^{122} in rhodopsin have been observed by FTIR spectroscopy (Fahmy et al., 1993; Sakmar and Fahmy, 1995).

The DIANA-generated structures with the lowest target function were selected for ligand docking and energy minimization. 11-*cis*-Retinal and all-*trans*-retinal were incorporated manually into the binding pockets of bovine rhodopsin using QUANTA to minimize bumps of retinal with surrounding protein residues. After this docking, retinal had only minor (<0.5 Å) hindrances with some surrounding rhodopsin atoms. These hindrances were removed by energy minimization of the α-bundle with the covalently bound retinal (35 iterations of unconstrained minimization with the CHARMm force field (Brooks et al., 1983; Momany and Rone, 1992) using a dielectric constant $\epsilon = 3$ and the adopted-basis Newton-Raphson method). Ten additional minimization iterations were made after incorporation of a water molecule in the model near the protonated SB. The final energy of rhodopsin with 11-*cis*-retinal was −2361 kcal/mol.

RESULTS AND DISCUSSION

"Average" 7-α-bundle structure of rhodopsin-like GPCRs

Final distance constraints

A single, specific GPCR does not provide a sufficient number of H bonds to calculate a convergent set of structures with pairwise r.m.s.d. <1.5 Å. Therefore, H bonds from different GPCRs were combined and verified for consistency by distance geometry calculations. The H bonds can be mutually consistent only if the spatial positions of all transmembrane helices are nearly the same in all different GPCRs. This can be expected because each of the helices is attached to the common core by side chains conserved throughout the GPCR family (Fig. 1). During the refinement the set of constraints was increased gradually, starting with H bonds between conserved polar residues deeply buried within the membrane, and finishing with the least reliable H bonds formed near the water-bilayer interface.

Surprisingly, in the final stages of refinement, 16 pairs of spatially proximate cysteines, which collectively are present in 103 different GPCRs, were detected (Table 2). In all of these pairs, the cysteine side chains had mutual orientations appropriate for disulfide bond formation. As shown by site-directed mutagenesis, all spatially close cysteines (even pairs inappropriately arranged, geometrically) form disulfide bonds in proteins (Matsumura and Matthews, 1991; Clarke et al., 1995). Thus all spatially close cysteines in individual GPCRs can be expected to form disulfide bonds. To check whether the disulfide bonds can actually be formed in the model, all of them were simultaneously in-

TABLE 2 Disulfide bonds of GPCRs, used in calculation of "average" 7α-bundle model*

Pairs of residues	Types and number of receptors
51–300#	2 opsins
85–116	3 opsins
87–300	3 opsins, 3 cholecystokinin, 3 C5a anaphylatoxin receptors
90–293	1 angiotensin II, 1 galanin receptor
118–207	7 melanocortin receptors
121–208	9 α_2-adrenergic receptors
122–167#	2 opsins
122–168	4 opsins
126–163#	5 bradikinin, 3 melanocortin receptors
126–164	11 melanocortin receptors
129–218	5 endothelin receptors
130–157	17 angiothensin, 8 high affinity interleukin-8 receptors
130–160#	1 high affinity interleukin-8 receptor
136–222	2 oxytocin, 6 FMLP-related receptors
168–211#	4 cone opsins
212–262	10 cone opsins

*The upper distance constraints applied for disulfide geometry were 4.20 Å for C^β-C^β, 3.05 Å for C^β-S^γ, and 2.04 Å for S^γ-S^γ (Momany et al., 1975).

#S-S bonds 51–300, 122–167, 126–163, 130–160, and 168–211 were approximated by C^β-C^β distances.

cluded in the calculations as supplementary constraints (11 disulfide bonds were included directly, and five others were approximated by $C^{\beta}\cdots C^{\beta}$ constraints; Table 2). The incorporation of these disulfide bonds increased the structural compactness of the α-bundle without causing violations of other constraints. Only disulfide bonds Cys^{85}-Cys^{116} and Cys^{90}-Cys^{293} have distorted S-S dihedral angles (with deviation >40° from the standard ±90° value). It must be emphasized that our model is not based on the assumed presence of these disulfide bonds, because they were incorporated only in the final stages of refinement and did not substantially change the model. On the contrary, the spatial proximity of the corresponding cysteines follows from the model, which was independently derived from hydrogen bonding constraints.

Only a fraction of all H bonds detected by ADJUST that could be formed in the final model by various GPCRs were applied directly as constraints on the distances between the corresponding donor and acceptor groups (Table 1) or were approximated by $C^{\beta}\cdots C^{\beta}$ or other constraints (Table 3). The most restrictive constraints were always chosen. For example, Glu^{122}, which forms an H bond with His^{211} in the bovine rhodopsin model, was replaced by Cys to incorporate a more restrictive Cys^{122}-Cys^{168} disulfide bond from the crayfish rhodopsin model. The alternative $Glu^{122}\cdots His^{211}$ and $Glu^{122}\cdots Asn^{212}$ H bonds, which could be formed in vertebrate rhodopsins and glycoprotein hormone receptors, respectively, and the disulfide bond Cys^{168}-Cys^{211} expected in four cone opsins could not be included directly after incorporation of the Cys^{122}-Cys^{168} bond; however, they were approximated by the corresponding distances between C^{β} and other atoms (Table 3).

In addition to the H bonds between side chains, some weakly polar side chains (Ser, Thr, Tyr, and Trp) form H bonds with backbone carbonyls (Table 1). For example, Tyr^{44} (helix I) and Tyr^{268} (helix VI), which appear simultaneously with Pro^{91} (helix II) and Pro^{291} (helix VII), respectively, in GPCR sequences, form H bonds with C=O groups of residues 87 and 287, which are otherwise unpaired because their NH counterparts are replaced by the prolines.

Final model

The final distance geometry calculations of the "average" model made use of 47 interhelical and 25 intrahelical H bonds (Table 1), 11 disulfide bonds (Table 2), 27 other distance constraints approximating hydrogen and disulfide bonds (Tables 2 and 3), 280 constraints for dihedral side-chain angles (Table 4), and restraints on the geometry of α-helices (Methods). The calculations yielded a well-defined set of structures with pairwise r.m.s.d of C^{α} atoms of 0.5–0.7 Å for the 10 structures with the lowest target functions (Fig. 3). The model is in agreement with the spatial proximity of residues suggested from mutagenesis and cross-linking studies. For example, the $Glu^{113}\cdots Lys^{296}$

TABLE 3 H bonds approximated by constraints, d_{upp}, between C^{β} and other atoms for the "average" model

Distance constraints			H bonds in GPCRs	
Atom 1	Atom 2	d_{upp}, Å*	Partners	Number of GPCRs
44 C^{β}	90 S^{γ}	5.40	Ser-Asp	8
72 C^{β}	134 C^{β}	7.90	Lys-Glu	1
73 C^{β}	306 C^{β}	8.30	Asn-Tyr	247
75 H^{η}	157 S^{γ}	2.80	Tyr-Asn	23
86 C^{β}	296 C^{β}	6.60	His-Ser	4
90 S^{γ}	113 $O^{\varepsilon 2}$	3.50	Asn-Tyr	8
90 S^{γ}	296 N^{ζ}	4.00	Asp-Lys	8
117 C^{β}	296 C^{β}	9.00	Asp-Tyr	146
118 S^{γ}	208 C^{β}	5.60	Asn-Ser	15
118 C^{β}	292 C^{β}	9.00	Tyr-Tyr	20
120 S^{γ}	299 O^{γ}	3.40	Cys-Thr	12
122 C^{β}	211 C^{β}	4.50	Glu-His	17
122 C^{β}	212 C^{β}	8.70	Glu-Asn	19
126 C^{β}	211 C^{β}	5.90	Glu-Ser	15
134 C^{β}	152 C^{β}	7.00	Glu-His	60
168 C^{β}	207 C^{β}	6.40	Thr-His	4
201 C^{β}	288 C^{β}	8.20	Arg-Thr	15
212 C^{β}	261 C^{β}	5.40	Asn-Asp	19
265 C^{β}	295 C^{β}	7.80	Tyr-Ser	2
268 C^{β}	291 C^{β}	5.80	His-Thr	20
222 C^{β}	247 O	7.60	Tyr-O#	6
301 C^{β}	253 O	8.00	Tyr-O#	66

* At each refinement iteration, the upper distance constraints, d_{upp}, were estimated as $d_{upp} = d^{C^{\beta}\cdots C^{\beta}} - (d_{O\cdots H} - 1.9)$, where $d^{C^{\beta}\cdots C^{\beta}}$ and $d_{O\cdots H}$ are the $C^{\beta}\cdots C^{\beta}$ and the minimum donor-acceptor distance, respectively, in the previous iteration of the model. The $d_{O\cdots H}$ distance was estimated (using ADJUST and QUANTA) by replacing the residues in the model with the residues forming the H bond in another GPCR (last column in the Table) and choosing their sterically allowed conformers with the closest donor-acceptor distance.

H-bond between Tyr side chain and main chain carbonyl.

TABLE 4 Side-chain conformers (χ^1, χ^2, χ^3, χ^4, degrees) in the "average" α-bundle model fixed by dihedral angle constraints*

Helix I residues $\chi^1, \chi^2, \chi^3, \chi^4$	Helix II residues $\chi^1, \chi^2, \chi^3, \chi^4$	Helix III residues $\chi^1, \chi^2, \chi^3, \chi^4$	Helix IV residues $\chi^1, \chi^2, \chi^3, \chi^4$	Helix V residues $\chi^1, \chi^2, \chi^3, \chi^4$	Helix VI residues $\chi^1, \chi^2, \chi^3, \chi^4$	Helix VII residues $\chi^1, \chi^2, \chi^3, \chi^4$
P34	P71	N111 60	E150 −60, 60	N210 180	E247 −60, 60	P285
W35 180, 90	L72 −60	L112 −60	N151 −60	E201 −60, −60	K248 −60	I286 −60, 180
Q36 −60, 180	N73 −60, 0	E113 180, 180, 90	H152 180, −90	S202 −60	E249 −60	N287 180, −90
Y37 −60, 90	Y74 −60, 90	T114 −60	A153	F203 −60, 90	V250 180	K288 180, −60
N38 −60, 90	Y75 −60, 90	H115 180, −90	I154 −60	V204 180	T251 −60	H289 −60, −90
M39 180	L76 180, 180	C116 −60	M155 180	Y205 −60, 90	R252 −60, 180, 180, 90	I290 −60
L40 −60, 180	K77 −60, 60, 180, −60	A117	G156	D206 180	T253 −60	P291
S41 −60	N78 180, 90	C118 180	C157 −60	C207 180	V254 180	Y292 −60, 90
A42	L79 −60, 180	Y119 180, 90	A158	C208 180	I255 −60	C293 −60
Y43 180, 90	S80 −60	C120 180	F159 −60, 90	V209 180	I256 −60	F294 180, 90
Y44 180, 90	V81 180	C121 180	T160 −60	W200 −60, 90	T257 60	H295 180, −90
F45 180, 90	A82	C122 −60	W161 180, 90	C211 −60	V258 180	K296 −60, 180, 180, 60
L46 −60	D83 −60, 140	S123 −60	V162 180	C212 180	I259 −60	S297 −60
L47 −60, 180	H84 −60, 90	S124 −60	M163 180	Y213 −60, 90	Q260 180, −60, −90	N298 180, 0
T48 −60	C85 180	L125 −60	C164 −60	I214 −60	D261 180, −90	S299 −60
M49 −60, −60	Q86 180, 180, 60	C126 −60	T165 −60	P215	C262 −60	C300 −60
L50 180	C87 −60	S127 −60	A166	L216 −60	I263 −60	H301 180, 90
S51 180	F88 180, 90	L128 −60	W167 180, −90	I217 −60	S264 −60	A302 −60, 90
F52 −60, 90	G89	C129 180	C168 −60	C218 180	M265 −60, −60	P303
P53	C90 180	C130 180	T169 −60	I219 −60, 180	T266 −60	V304 180
I54 −60	F91 180, 90	L131 −60	P170	F220 −60, 90	P267	I305 −60
N55 −60, 90	T92 −60	A132	P171	F221 180, 90	Y268 180, 90	Y306 −60, 90
F56 −60, 90	T93 −60	I133 −60	V172 180	C222 −60	N269 −60	I307 −60, −60
L57 180	T94 60	E134 180, −60	V173 180	Y223 180, 90	G270	M308 180, 180
T58 −60	L95 180	R135 180, 60, −60, −90	G174		H271 180, 90	N309 −60, −90
L59 −60, 180	Y96 180, 90	C136 −60	H175 180, −90		N272 180, −90	N310 180, −90
Y60 180, 90	D97 60, −90	V137 180	S176 −60		W273 180, 90	
V61 180	Q98 180, 180	V138 180	R177 180		Y274 180, 90	
T62 −60		V139 180	Y178 180, 90		I275 −60	
Q63 −60, −60, 90		C140 −60			A276	
Q64 180, 180		K141 180, 180, 60, 180			T277 −60	
H65 −60, −90						
K66 180, −60						

*The dihedral angle constraints were defined as the "standard" angles from the table ±15°.

and $Asp^{97}\cdots Lys^{288}$ ion pairs (the corresponding H bonds are present in Table 1) have previously been proposed from mutagenesis studies of rhodopsin (Zhukovsky and Oprian, 1989; Sakmar et al., 1989; Nathans, 1990b) and lutropin/choriogonadotropin hormone receptor (Fernandez and Puett, 1996a), respectively. The conserved Asp^{83} and Asn^{302} residues, which are suggested to be close from compensatory mutagenesis data (Zhou et al., 1994; Sealfon et al., 1995), were far from each other (>8 Å) in the initial model; however, the $O^{\delta}H\cdots O^{\delta}$ distance is 2.0 Å in the final "average" structure. Furthermore, although constraints between residues 200, 204, and 276 were not explicitly included, these residues, when replaced by histidines with properly chosen χ^1 conformers, can adopt a geometry (with $N^{\epsilon 2}\cdots N^{\epsilon 2}$ distances of 3.5–4.5 Å) appropriate for formation of a Zn^{2+} binding center, as has been artificially designed in tachykinin NK-1 and κ opioid receptors (Elling et al., 1995; Thirstrup et al., 1996). In the model, residues 44 (helix I) and 289 (helix VII), suggested to be close in muscarinic receptors (Liu et al., 1995), do not contact directly; however, the side chain of Trp, conserved in amine receptors in position 293, is tightly packed between residues 44 and 289 and cannot tolerate β-branched side chains in both positions. This may explain the incorrect folding of chimeric m2/m5 muscarinic receptors containing both Thr^{44} and Thr^{289} (Liu et al., 1995).

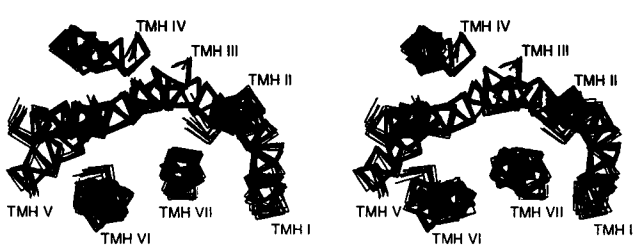

FIGURE 3 Superposition of 10 structures, representing the "average" GPCR model (viewed from the extracellular side).

The evolutionarily conserved domain of GPCRs

In the "average" 7α-bundle model, 40 of 44 conserved intramembrane GPCR residues (the "consensus" string in

Fig. 1) form a single continuous domain. This domain consists of an aliphatic "minicore" in the intracellular half of bilayer (Fig. 4 *A*); four polar clusters (Fig. 4 *B*), which are probably important for transduction, as discussed below; and two clusters (Fig. 4 *C*), which interact with retinal and other ligands. The extracellular part of the α-bundle contains fewer conserved residues, because it serves to adopt ligands of various structures and sizes.

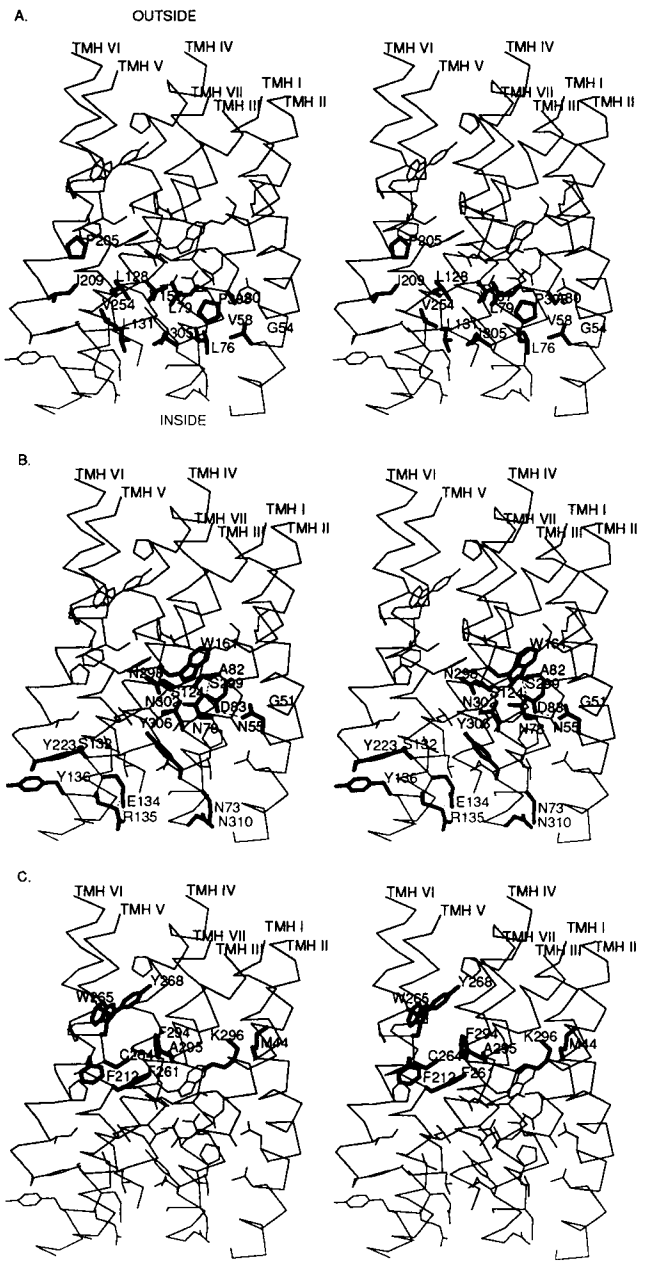

FIGURE 4 The evolutionarily conserved domain of GPCRs (depicted as the bovine rhodopsin sequence with six residues replaced by more conserved ones: I54G, T58V, A124S, A132S, S298N, A299S; the side chains shown are from the "consensus" string in Fig. 1). (*A*) Cluster of conserved aliphatic residues. (*B*) Four clusters of conserved polar side chains. (*C*) Two clusters in the vicinity of the ligand-binding site.

The evolutionarily conserved GPCR residues are segregated into seven clusters of different polarities. Cluster 1 (Fig. 4 *A*) is an extensive layer of 13 aliphatic side chains from all seven helices: transmembrane helix (TMH) I (Gly^{54} and Val^{58}), TMH II (Leu^{76}, Leu^{79}, and Ala^{80}), TMH III (Leu^{128} and Leu^{131}), TMH IV (Val^{157}), TMH V (Pro^{215}, Ile^{219}), TMH VI (Val^{254}), and TMH VII (Pro^{303} and Ile^{305}). This aliphatic layer is in the middle of the conserved domain and is surrounded by two polar clusters near the intracellular surface (cluster 2: Asn^{73} and Asn^{310}; cluster 3: Ser^{132}, Glu^{134}, Arg^{135}, Tyr^{136}, and Tyr^{223}) and by a large group of polar side chains in the middle of the α-bundle (cluster 4: Gly^{51}, Asn^{55}, Asp^{83}, Ser^{124}, Asn^{298}, Ser^{299}, and Asn^{302}) (Fig. 4 *B*). Cluster 4, the largest hydrophilic cluster, contains an empty space, which can be filled by water or by sodium coordinated with oxygens of the Asp^{83}, Ser^{124}, and Asn^{302} side chains. It has been shown that Asp^{83} is important for allosteric regulation of some GPCRs by sodium (Horstman et al., 1990; Quintana et al., 1993; Kong et al., 1993a,b; Ceresa and Limbird, 1994). Bovine rhodopsin lacks Na^+ regulation, and its Ser^{124} is replaced by Ala. Gly^{51} of cluster 4 is replaced by more polar Ser or Thr residues in many GPCRs. In some GPCRs, this cluster is surrounded by other polar but less conserved side chains (Ser/Thr/Glu^{54}, Ser/Thr^{80}, Thr/His/Tyr/Glu^{84}, Ser/Thr/His/Tyr/Glu^{86}, Ser/Thr^{87}, and Ser/Thr/Asn^{120}), which, together with the conserved polar side chains, create H-bond networks of various sizes. For example, in human red cone opsin, the side chains of Thr^{54}, Glu^{86}, Thr^{87}, and Cys^{131} supplement the central polar cluster. Similarly, the smaller, hydrophilic cluster 5 (Asn^{78}, Ala^{82}, and Trp^{161}, Fig. 4) is supplemented in various GPCRs by Tyr/His^{119} or Ser/Thr/Asn/His^{127}, which form H bonds with Asn/Ser^{78}, or by a Ser/Thr^{123}···Ser/Thr^{165} H-bond pair. Depending on its χ^1 angle (180° or −60°), the Tyr^{306} side chain can participate in clusters 2 or 4, respectively.

Cluster 6 (Fig. 4 *C*) is closer to the extracellular side of the α-bundle and consists mostly of aromatic and small side chains from helices V (Phe^{212}), VI (Phe^{261}, Cys^{264}, Trp^{265}, and Tyr^{268}) and VII (Phe^{294} and Ala^{295}). This common binding site for the retinal β-ionone ring in opsins, the tyramine fragment in opioid receptors, adenine in ATP receptors, and catecholamines in cationic amine receptors probably plays a key role in triggering signal transduction, as discussed below for rhodopsin. In glycoprotein receptors, this aromatic cluster becomes more hydrophilic (Asn^{212}, Asp^{261}, and His/Tyr^{295}), includes polar residues from helix III (Ser^{121}, Glu^{122}), and joins a network of H bonds of the large polar cluster 4 through the Asn^{298} side chain.

Cluster 7 consists of only two subfamily-specific side chains, which are present as a Met^{44}, Lys^{296} pair in vertebrate opsins (both residues participate in the binding site of the retinal Schiff base NH group in the model), and as a Tyr^{44}, Tyr^{296} pair in many peptide receptors. In cationic amine receptors Tyr^{296} forms a cluster with the conserved Trp^{293}, which spatially substitutes for Tyr^{44} (Fig. 1).

Saturation of H-bonding potential

The final α-bundle model was examined by ADJUST for saturation of H-bonding potential of all polar side chains in transmembrane segments 21–22 residues long (Fig. 2), using χ^1 conformers from Table 4. Some "unsatisfied" long, hydrophilic side chains (Gln, Glu, Lys, or Arg) appear in the first and last turns of the transmembrane α-helices and can be solvated at the membrane interface or can form H bonds with loops. The polar groups of side chains in positions 48, 113, 114, 118, 122, 204, 208, 268, and 269, situated in the area of the binding pocket near the extracellular surface, usually form intramolecular H bonds, detected by ADJUST; however, in some receptors they cannot, and interact only with ligands, water molecules, or extracellular loops. The importance of residues in most of these positions for ligand binding in various GPCRs has been demonstrated by mutagenesis (see review of Baldwin, 1994, for example). All of the remaining numerous, highly polar residues (Arg, Asn, Asp, Glu, Gln, His, Lys) form at least one (but usually more) H bond with other side chains in each of the 410 GPCRs considered, with four exceptions. "Unsatisfied" Asn^{55} and Asn^{302} residues appear in two green-sensitive opsins of goldfish (Swiss-Prot amino acid sequence identifiers, P32311 and P32312). These residues are conserved and situated within the transmembrane α-bundle and form H bonds with Asp^{83} from helix II in most GPCRs. However, this conserved Asp^{83} residue is replaced by Gly in goldfish opsins, creating a cavity that can be filled by bound water molecules solvating the polar Asn^{55} and Asn^{302} side chains. In the remaining two cases, Asn^{166} or Glu^{166} side chains from helix IV of the human high-affinity interleukin-8 receptor (P25024) or the m4-muscarinic receptor (P30544), respectively, are immersed in the lipid phase. These polar side chains may participate in dimerization or trimerization of these receptors, as seen, for example, in 2D crystals of squid rhodopsin (Davies et al., 1996), where symmetrical molecules are linked through helix IV. In this case the polar side chains in positions 166 from two or three identical molecules would form H bonds with each other.

All less polar Ser, Thr, and Tyr side chains of GPCRs also satisfy their H-bonding potential. Most Ser and Thr side chains have a χ^1 angle of $\sim -60°$, which allows formation of an H bond between their $O^{\gamma}H$ group and the corresponding $(i - 4)$ backbone C═O group in the same α-helix (Baker and Hubbard, 1984), although they are often also involved in other H bonds. All $O^{\eta}H$ groups of Tyr side chains near the middle of the membrane (positions 44, 75, 86, 115, 157, 212, 213, 261, 268, and 301 in different GPCRs) also form H bonds, sometimes with main-chain carbonyls of an adjacent helix. The side chains of Tyr and Trp residues situated near the ends of the transmembrane α-helices (within two turns) at the lipid-exposed surface of the α-bundle sometimes lack H bonds, but can reach the water-lipid interface.

Side-chain packing

The "average" 7-α-bundle model was also tested for compatibility with close packing of side chains in six dissimilar GPCRs (bovine and crayfish rhodopsins, human δ opioid, lutropin/choriogonadotropin hormone, muscarinic acetylcholine, and ATP receptors) and squid retinochrome. The α-bundles of all these proteins were calculated with DIANA, using their individual systems of H bonds and the supplementary $C^{\beta}\cdots C^{\beta}$ constraints taken from the common "average" model, as for bovine rhodopsin (Methods). The compatibility test is particularly demanding when a bulky aromatic side chain appears within the transmembrane α-bundle, because this usually requires a concomitant decrease in volume of several surrounding side chains. For instance, the appearance of the Phe^{62} side chain in crayfish and other invertebrate opsins, instead of Thr^{62} present in bovine and other vertebrate opsins, is correlated with $Leu^{76}\rightarrow$Val and $Ile^{307}\rightarrow$Ala replacements (Fig. 5).

3-D structure of bovine rhodopsin

The α-bundle of bovine rhodopsin was calculated using its own set of H bonds (Table 5) and $C^{\beta}\cdots C^{\beta}$ constraints taken from the "average" model (Methods). An extra deviation of 1 Å, applied to the $C^{\beta}\cdots C^{\beta}$ distances, allowed the rhodopsin structure to relax compared to the "average" structure, while still maintaining structural compactness of the α-bundle. The rhodopsin structure is well defined: the r.m.s.d. for 202 C^{α} atoms is 0.5–0.6 Å for the five structures with the lowest target function. The dispersion between structures is higher near the ends of the helices (Fig. 6). All backbone angles of the model are within the α-helical region of the Ramachandran map, and all side chains have allowed χ^1-χ^4 conformers, as is automatically provided by the dihedral angle constraints (violations of the individual angle constraints were $<10°$). A few violations of van der Waals constraints of ~0.5 Å were present near Pro residues in α-helices; no violations of H-bond distances greater than 0.6 Å were found.

The rhodopsin model thus obtained differs from other published GPCR models, the coordinates of which are available from the GPCRDB data base (http://swift.embl-heidelberg.de/7tm/models/models.html). Differences be-

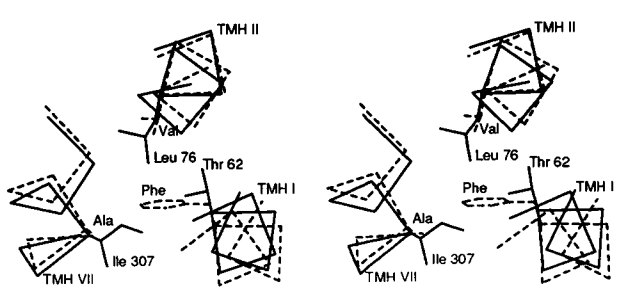

FIGURE 5 An example of correlated replacements of nonpolar residues in bovine (——) and crayfish (– – –) rhodopsins.

TABLE 5 H bonds of side chains applied as distance constraints for calculation of bovine rhodopsin model*

Tyr^{43} H^{η}	Thr^{297} $O^{\gamma 1}$	*Only in inactive conformation:*	
Asn^{55} $H^{\delta 22}$	Asp^{83} $O^{\delta 1}$	Glu^{113} $O^{\varepsilon 2}$	Lys^{296} $H^{\zeta 1\#}$
His^{65} $H^{\delta 1}$	Asn^{73} $O^{\delta 1}$	Glu^{134} $O^{\varepsilon 1}$	Arg^{135} H^{ε}
Asn^{73} $O^{\delta 1}$	Asn^{310} $H^{\delta 22}$	Glu^{134} $O^{\varepsilon 2}$	Arg^{135} $H^{\eta 12}$
Asn^{78} $H^{\delta 21}$	Ser^{127} O^{γ}	Lys^{141} $H^{\zeta 1}$	Glu^{150} $O^{\varepsilon 1}$
Asn^{78} $O^{\delta 1}$	Trp^{161} $H^{\varepsilon 1}$	Tyr^{223} H^{η}	Thr^{251} O
Tyr^{96} H^{η}	Asn^{111} $O^{\delta 1}$	Asn^{302} $H^{\delta 21}$	Tyr^{306} O^{η}
Glu^{122} $O^{\varepsilon 1}$	His^{211} $H^{\delta 1}$	*Only in active conformation:*	
Arg^{135} $H^{\eta 21}$	Glu^{247} $O^{\varepsilon 1}$	Asp^{83} $H^{\delta 2}$	Asn^{302} $O^{\delta 1}$
Tyr^{206} H^{η}	Cys^{167} S^{γ}	Glu^{134} $O^{\varepsilon 2}$	Lys^{141} $H^{\zeta 1}$
Trp^{175} $H^{\varepsilon 1}$	Glu^{201} $O^{\varepsilon 1}$	Glu^{134} $O^{\varepsilon 1}$	His^{152} $H^{\delta 1}$
Asn^{200} $N^{\delta 2}$	Glu^{201} $O^{\varepsilon 1}$	Tyr^{223} O^{η}	Arg^{252} H^{ε}
Glu^{122} $H^{\varepsilon 2}$	Thr^{118} O	Asn^{73} $O^{\delta 1}$	Tyr^{306} H^{η}
Tyr^{268} H^{η}	Phe^{287} O		
Tyr^{301} H^{η}	Ile^{253} O		

*H-bond distance constraints were as in Table 1. In addition, H bonds of 12 threonine (58, 62, 92, 93, 94, 97, 118, 160, 251, 277, 289, 297) and 3 serine (38, 127, 176) side chains with *i*-4 backbone carbonyls were included.

#This H bond was systemically violated in calculations with DIANA, and was replaced by two H bonds through a bound molecule of water in the final model (Glu^{113} $O^{\varepsilon 2}$···H-O···H^{ζ} Lys^{296}; see Fig. 9).

tween our model and the models of Oliveira et al. (1993), Donnelly et al. (1994), and Lin et al. (1994) are especially pronounced (r.m.s.d. of 5.7–6.3 Å for corresponding C^{α} atoms in the transmembrane domain) because of differences in tilt angles (all of these models have nearly parallel or antiparallel orientations of all helices, in contradiction of EM data), and because some helices in these models are shifted relative to our model by ~1–2 turns in the direction perpendicular to the membrane plane. The model of Herzyk and Hubbard (1995) is much closer to our structure, because these authors adjusted helix tilts to agree with the EM data and vertically "aligned" transmembrane helices, identified by Baldwin (1993). Nevertheless, helices II and IV in the model of Herzyk and Hubbard are shifted relative to our structure by 2–3 Å in the middle of the α-bundle, and up to 5 Å near the helix ends, producing a r.m.s.d of 3.9 Å.

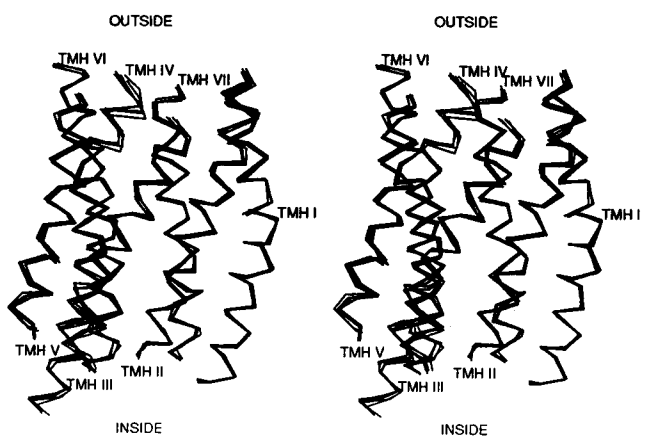

FIGURE 6 Superposition of five structures of bovine rhodopsin with the lowest target function.

In all of the previous models, side-chain conformers were arbitrarily chosen without analysis of the packing. All of these previously reported models are also less compact than our model (all contain large empty inner spaces), and none of them satisfy the H-bonding saturation test. Examination of the models using our program ADJUST reveals many polar side chains (Asn, Asp, Glu, Gln, and others) in hundreds of GPCRs that have no H bonds and often are immersed in the lipid. The sets of H bonds proposed in other publications are smaller (for example, eight interhelical H bonds in the rhodopsin model proposed by Fanelli et al., 1995a, compared with 16 H bonds here (Table 5)), and only a few of them (Asn^{55}···Asp^{83} and Glu^{113}···Lys^{296} (Fanelli et al., 1995a,b); Asn^{55}···Asp^{83}, Asp^{83}···Asn^{302}, Asn^{302}···Tyr^{306} (Scheer et al., 1996); Asp^{83}···Asn^{302} (Zhou et al., 1994; Sealfon et al., 1995), and Glu^{122}···His^{211} (Herzyk and Hubbard, 1995)) are the same as in our model.

Arrangement of helices

The calculated structure of rhodopsin (Fig. 6) qualitatively reproduces all features of 6–9-Å resolution EM maps (Schertler et al., 1995; Schertler and Hargrave, 1995; Unger and Schertler, 1995). Helices IV, VI, and VII are approximately parallel to each other, whereas the remaining four helices are tilted (see also Fig. 3). The directions of these tilts are the same as in the published low-resolution 3D EM structures. The strongly tilted helix III is in the center of the α-bundle and makes contacts simultaneously with helix II (close to the extracellular region), VI, and VII (in the middle of the membrane), and IV and V (near the intracellular surface). The N-terminal part of helix V interacts with helix IV, whereas its C-terminus is inserted between the intracellular ends of helices III and VI. In agreement with EM data, the intracellular ends of helices II and IV are closer in the rhodopsin structure than in bacteriorhodopsin.

The directions of the kinks induced by Pro residues in helices I, IV, V, VI, and VII, visible in the 7-Å 3D map of frog rhodopsin (Unger et al., 1995), are also reproduced in the model. These kinks shift the N-terminus of helix V and the C-terminus of helix VII slightly outside the α-bundle, bend the C-terminal ends of helices IV and VI toward helices III and V, respectively, and produce an additional tilt of the N-terminal end of helix I, which is shifted out of the α-bundle, in front of helix VII (Fig. 6). Helix III is slightly curved in the middle, near Ala^{124}. This curvature was induced mostly by the H bond between the Ser^{124} and Asn^{302} side chains in the "average" GPCR model. In gonadotropin-releasing hormone receptors, Ser^{124} is replaced by Pro, which probably produces a kink at this site.

In agreement with the EM maps, all interhelical angles are −150° to −165° or 10° to 20° for antiparallel or parallel helices, respectively, as required for "knobs-into-holes" side-chain packing, with the exception of helix IV, which forms a positive (~+160°) angle with helix III, exactly as in bacteriorhodopsin (PDB file, 2brd), and helices VI and

VII, which are parallel to each other, as in the EM projection maps.

Buried and accessible residues

The arrangement of the buried and lipid-accessible residues in the model is consistent with chemical modifications of ovine rhodopsin by the photosensitive hydrophobic probe 1-azido-4-iodobenzene, which labels Cys, Tyr, Trp, His, and Lys side chains exposed to a nonpolar environment (Davison and Findlay, 1986a,b) (in ovine and bovine rhodopsins, all of these residues are identical within the transmembrane segments). This hydrophobic probe modifies all 14 Cys, Tyr, Trp, and His side chains, the reactive groups of which are exposed to lipid in the model and four side chains (Trp^{175}, Trp^{265}, Tyr^{268}, and Lys^{296}) located in the nonpolar retinal-binding pocket (Fig. 7). By contrast, all residues that are completely buried in the model (Trp^{126} and His^{211}) or are located at the ends of helices in the hydrophilic environment (Tyr^{96}, Tyr^{178} and Tyr^{274}, Tyr^{136}, Cys^{140}, Lys^{141}, and Lys^{248}) were not modified by the hydrophobic probe (Fig. 7). All polar residues that are accessible at the intracellular surface in the model (His^{65}, Lys^{66}, Tyr^{74}, Tyr^{136}, Cys^{140}, Lys^{141}, and Lys^{248}) were modified by hydrophilic reagents applied from the cytoplasmic membrane side (Barclay and Findlay, 1984; Ridge et al., 1995). The polar Glu^{122} and Glu^{134} residues, completely buried in the model, were not modified by any reagents (Davison and Findlay, 1986b).

The rhodopsin model is also consistent with accessibilities and mobilities of residues near the intracellular ends of helices III and IV, studied by site-directed spin labeling (Farahbakhsh et al., 1995). The spin labels are less flexible in positions 136, 138, 139, 140, and 153, where the corresponding side chains are closely packed with each other and with proximal residues from adjacent helices V and VI. The boundaries of the intramembrane hydrophobic parts of helices III and IV, identified as being located near Val^{139} and His^{152} residues, respectively (Farahbakhsh et al., 1995), correspond to the border between a layer of buried aliphatic side chains and the polar intracellular surface in the model. A similar, recent study of nonpolar boundaries of helices V and VI at the cytoplasmic side (Altenbach et al., 1996) is in agreement with the end of helix VI in our model, but suggests that the C-terminus of helix V may be extended by seven additional residues.

Clusters of polar and aromatic side chains

Polar and aromatic side chains of the bovine rhodopsin model are segregated in clusters, some of which are common for all GPCRs (Fig. 4), whereas others can be found only in related vertebrate opsins. For example, an opsin-specific cluster is formed by the polar Glu^{122} and His^{211} side chains, which are shielded from a highly nonpolar aliphatic environment in the model by a "shell" of aromatic and sulfur-containing groups of Trp^{126}, Met^{163}, Cys^{167}, Tyr^{206}, Met^{207}, and Phe^{208} residues with intermediate polarity (Fig. 8). The spatial arrangement of side chains in this region looks like a polarity or solubility gradient, which is energetically favorable, because chemically similar compounds are more soluble in each other. Clustering of side chains with similar solubilities (i.e., the hydrophilic, aromatic, sulfur-containing, and aliphatic groups) is probably the major driving force of protein folding, and the tendency of nonpolar residues in proteins to be shielded from water, which is usually treated as the main principle of protein structure, is only a specific case of this more general rule.

The "solubility forces" are equally important for protein-lipid interactions. The rhodopsin model has a long, contin-

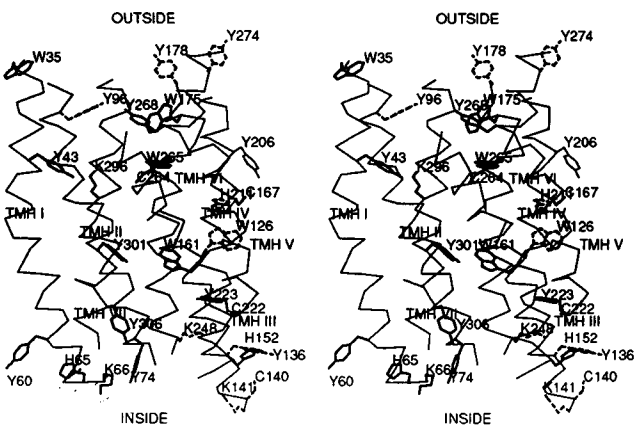

FIGURE 7 Cys, His, Tyr, Trp, and Lys residues accessible to modification by hydrophobic probe 1-azido-4-iodobenzene in ovine rhodopsin (*thick solid line*) (Davison and Findlay, 1986a,b). The dashed line indicates residues unmodified by the probe. The thin line indicates Tyr^{206}, for which experimental data are unavailable.

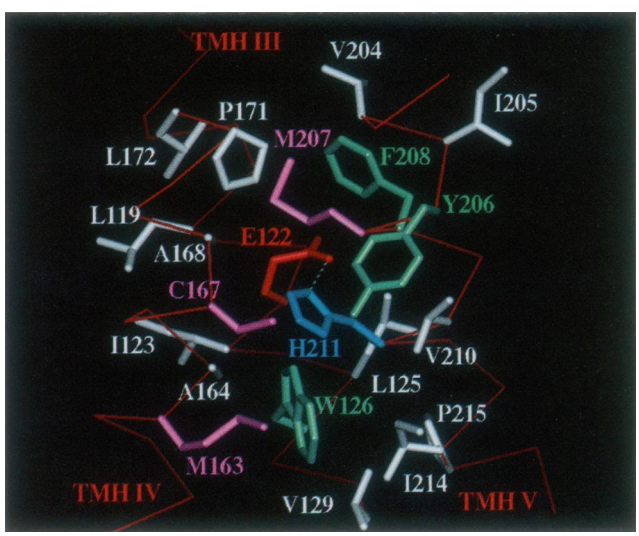

FIGURE 8 Glu^{122}-His^{211} pair and its surrounding in bovine rhodopsin. Color indicates residues: white, aliphatic (Ala, Pro, Val, Leu, Ile); green, aromatic (Phe, Tyr, Trp); magenta, sulfur-containing (Cys, Met), yellow, polar (Gly, Ser, Thr, Asn, Gln); blue, His, Lys, Arg; red, Asp, Glu.

uous belt of aromatic rings at the outer, lipid-facing surface of the α-bundle, created by Trp^{161} and by Phe in positions 45, 52, 56, 85, 88, 91, 115, 116, and 159 from helices II, III, and IV, whereas the more polar aromatic side chains of Trp^{35}, Tyr^{60}, His^{65}, Tyr^{74}, Tyr^{96}, Tyr^{136}, Tyr^{178}, Tyr^{223}, Tyr^{274}, and Tyr^{306} are situated closer to the ends of the transmembrane segments and are embedded in the interfacial area of the lipid bilayer, similar to layers of Trp residues in the bacterial photoreaction center (Deisenhofer and Michel, 1991), bacteriorhodopsin (Henderson et al., 1990), and gramicidin A dimer (Lomize et al., 1992), or to regular rows of Phe and Tyr side chains in porins (Schultz, 1992). The lipid-facing Phe side chains of rhodopsin may be important for solubility of the protein in photoreceptor membranes, which have an unusually high content of lipids with unsaturated fatty acids (Lamba et al., 1994).

Inner cavities

The interior of the rhodopsin α-bundle is tightly packed, yet contains a large retinal-binding pocket and four smaller cavities, which are present in all structures with low target function. All four cavities are formed by polar side chains and can be filled by bound water molecules, which have been detected by FTIR difference spectroscopy of rhodopsin (Maeda et al., 1993; Kandori and Maeda, 1995). The presence of such cavities is typical for proteins undergoing conformational transitions, such as citrate synthase or the α subunit of a Gi-protein, for example (2cts and 1gfi PDB files, respectively).

The largest cavity in the rhodopsin structure appears near the retinal-binding pocket, between the Ala^{82}, Asp^{83}, Met^{86}, Ala^{124} Met^{257}, Phe^{261}, Ser^{298}, Ala^{299}, and Asn^{302} side chains from helices II, III, and VII. This cavity, which probably forms an allosteric Na^+ binding site in other GPCRs (see above), resembles a narrow transmembrane channel that is blocked from the extracellular side by retinal and from the intracellular side by the conserved Tyr^{306} side chain.

The second cavity is situated between the polar side chains of Thr^{93}, Thr^{94}, Thr^{97}, Glu^{113}, Thr^{289}, and Lys^{296} and can accept several bound water molecules. The presence of this funnel-like cavity explains the fast deuterium exchange of the Schiff base (SB) imine proton (Deng et al., 1994), as discussed below. The third cavity appears between the Asp^{83}, Gly^{51}, and Asn^{55} side chains; this space is occupied by larger Ser^{51} or Thr^{51} side chains in most GPCRs. The fourth cavity is situated near the Tyr^{223} side chain, which is probably involved in the photoactivation of rhodopsin, as described below.

Genetic polymorphism of rhodopsins

The rhodopsin model is consistent with the effects of naturally occurring mutations that cause visual defects, such as autosomal dominant retinitis pigmentosa or congenital night blindness (Kaushal and Khorana, 1994). These effects are expected to be deleterious when a charged residue appears in a nonpolar intramembrane site; however, some such replacements (T58R, G90E, and L125R) have a very small effect on the properties of rhodopsin, such as absorption spectra, retinal binding, and activation of transducin. All of these results are consistent with the model: the guanidinium group of the Arg^{125} side chain (with χ angles mimicking the wild-type Leu^{125} residue) occupies the major hydrophilic cavity of the α-bundle; the COO group of Glu^{90} can occupy another polar cavity near Glu^{113}; and the Arg^{58} guanidinium group can access the water-membrane interface between helices I and VII. On the other hand, the appearance of charged residues at the lipid-exposed surface of the model (Arg^{53}, Asp^{89}, Arg^{167}, Arg^{211}, Arg^{310}) or in nonpolar sites inside the α-bundle (Arg^{51}, Asp^{87}) produces mutants that are defective in protein folding and retinal binding (Kaushal and Khorana, 1994).

Retinal-binding pocket

Arrangement of chromophore group. The calculated structure of rhodopsin has an elongated, narrow, deep cavity between helices II, III, V, VI, and VII, formed by 36 residues, which can accommodate retinal. The cavity is partially covered from the extracellular side by Glu^{113}, Trp^{175}, Glu^{201}, Phe^{276}, Phe^{287}, and Met^{288} side chains. Conformers of all side chains within the binding pocket are uniquely defined in the model, except for Trp^{265} and Lys^{296}. Two possible conformers of the Trp^{265} side chain ($\chi^1 = -60°$, $\chi^2 = 10°$, or $\chi^1 = 180°$, $\chi^2 = 60°$) create two corresponding alternative binding sites for the retinal β-ionone ring. The *trans* ($\chi^1 \approx 180°$) Trp^{265} conformer (Figs. 9 and 10 *A*) places the β-ionone ring deeper in the membrane,

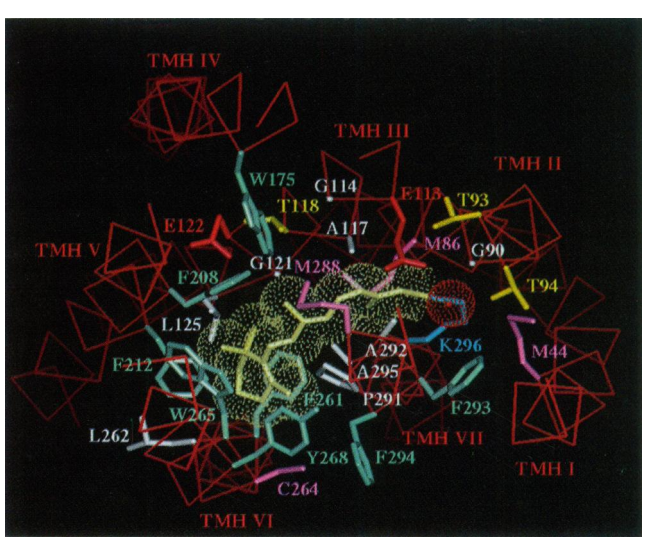

FIGURE 9 Binding pocket of 11-*cis*-retinal in bovine rhodopsin (view from the extracellular side). Colors are as in Fig. 8. The possible bound water molecule is shown by a red dot sphere; the retinal van der Waals surface is shown by dots.

as is consistent with experimental data for rhodopsin with bound 11-*cis*-retinal, described below.

The bottom of the binding cavity, with the properly chosen conformers of Trp^{265} and Lys^{296} side chains, is complementary to 11-*cis*-retinal in its 6-*s*-*cis*, 12-s-*trans* conformation and with the *anti* configuration of the C=N bond of the protonated SB (Fig. 9). This conformation of rhodopsin-bound retinal and the protonation of the SB have been demonstrated by Raman (Oseroff and Callender, 1974; Callender et al., 1976; Palings et al., 1987), FTIR (Bagley et al., 1985), and solid-state ^{13}C NMR (Smith et al., 1987, 1990; Mollevanger et al., 1987) spectroscopies. The retinal fits tightly in the bottom of the binding cavity, creates no new holes, and causes only minor (<0.5 Å) hindrances with some surrounding rhodopsin atoms. These hindrances were readily removed by the subsequent energy minimization of rhodopsin with the covalently bound retinal, which did not significantly alter the structure: r.m.s.d. between rhodopsin structures before and after energy minimization was 0.06 Å for 244 nonhydrogen atoms of 27 residues surrounding the 11-*cis*-retinal.

The complementarity of the retinal SB and its binding pocket is evident from their geometrical fit and matching of polarity characteristics. The nitrogen of the SB is in a relatively polar site, which is formed by the Glu^{113} COO^- group, the aromatic ring of Phe^{293}, the backbone of helix II near Gly^{90}, the side chain of Thr^{94}, and the δ sulfur of Met^{44} (Fig. 9). The polyene chain of retinal passes through a narrow "gate" (~4-Å width) between helices III and VII created by residues with small side chains (Ala^{117}, Thr^{118}, Gly^{121}, Ala^{292}, and Ala^{295}). The β-ionone ring occupies a wider (~7.5 Å) nonpolar cavity formed by Leu^{125}, Phe^{208}, Phe^{212}, Phe^{261}, Leu^{262}, Cys^{264}, Trp^{265}, and Phe^{294} residues (Fig. 9).

The retinal-specific structure of the binding pocket is uniquely defined by residues that are conserved in all rhodopsins but disappear in GPCRs activated by other ligands: Gly^{114}, Thr/Ser^{118}, Gly^{121}, Trp^{175}, Tyr^{268}, Pro^{291}, Ala^{292}, Phe^{294}, Ala/Ser^{295}, and Lys^{296}. The Met^{44}, Glu^{113}, Ala/Thr^{269}, Ala^{272} residues are common only for vertebrate rhodopsins, whereas Phe^{261} and Trp^{265} are conserved throughout the entire GPCR family, consistent with their hypothesized participation in the common transduction pathway, as described below.

The long axis of the retinal chromophore is approximately perpendicular to the α-bundle, and its β-ionone ring is deeper than the SB NH group, consistent with the ~16° angle between the transition dipole moment of retinal and the bilayer plane, which has been estimated from linear dichroism studies (Liebman, 1962).

The retinal is situated 26–28 Å from the intracellular side, but only 10–12 Å from the extracellular surface of the calculated α-bundle. However, the extracellular side of rhodopsin is covered by loops, which are clearly visible in the 7-Å resolution structure of frog rhodopsin (Schertler et al., 1996). Thus the model is consistent with fluorescence studies that estimate the distance from retinal to the intra- and extracellular surfaces as ~28 and 22 Å, respectively (Thomas and Stryer, 1982), if the thickness of the loop area is ~10 Å.

Environment of the Schiff base imine group. In the rhodopsin model, the protonated SB and Glu^{113} form an ion pair, in agreement with mutagenesis data (Sakmar et al., 1989; Zhukovsky and Oprian, 1989; Nathans, 1990b); however, the $N^+\cdots O^{\epsilon 2}$=C distance is 3.6 Å after energy minimization of the rhodopsin model, i.e., ~0.6 Å longer than the standard H-bond length. However, a water molecule that can simultaneously form H bonds with the SB imine hydrogen, O^γ of the Thr^{94}, and $O^{\epsilon 2}$ of Glu^{113} can be readily incorporated at the bottom of a small funnel-like cavity in the model between Thr^{93}, Thr^{94}, Glu^{113}, and Lys^{296} side chains (Fig. 9). Such an H bond between the protonated SB and a bound water has previously been suggested to explain the fast deuterium exchange of the imine proton (Deng et al., 1994). In the model with the incorporated water molecule, the $N^+\cdots O^{\epsilon 2}$ distance is increased to 4.1 Å. The other ($O^{\epsilon 1}$) carboxylate oxygen of Glu^{113} is 3.3 Å from the C_{12} retinal carbon, consistent with ^{13}C solid-state NMR data (Mollevanger et al., 1987; Han et al., 1993; Han and Smith, 1995a,b). The relative arrangement of the retinal SB, the $O^{\epsilon 1}$ oxygen of Glu^{113}, and the water molecule in the model is very similar to that previously proposed by Han et al. (1993), and the $O^{\epsilon 1}$-C_{12}-H angle (25°) is within the limits (from −45° to 90°) estimated by Han and Smith (1995b).

Despite the fast deuterium exchange, the protonated retinal SB has an extremely high pK_a (Steinberg et al., 1993) and is inaccessible to hydroxylamine (Sakmar et al., 1991), because it is shielded by the Glu^{113} side chain (Fig. 9). Replacement of Glu^{113} by the smaller Ala or Asp residues deshields the imine group in the model, which correlates with decreased pK_a and increased accessibility to hydroxylamine of the SB in the corresponding mutants (Sakmar et al., 1991). Free space above Glu^{113} in the model allows different χ^2 conformers of its side chain; however, the ionic interaction with the SB imine fixes χ^2 at 180°, blocking the C=N bond from the bulk solution by the Glu^{113} side chain. In a E113Q rhodopsin mutant, lacking this ionic interaction, the Gln^{113} side chain is expected to be flexible, conferring increased accessibility of the C=N bond to water and to hydroxylamine (Sakmar et al., 1991).

It has been observed that Asp residues in positions 90 (Rao et al., 1994) and 117 (Zhukovsky et al., 1992; Zvyaga et al., 1993) can substitute for the Glu^{113} counterion in rhodopsin mutants. The corresponding ionic interactions between these Asp side chains and the protonated SB are readily reproduced in our model. Conformers of Asp^{90} and Asp^{117} side chains in the model are uniquely defined by packing requirements as $\chi^1 \approx 180°$ and $\chi^1 = -60°$, respectively. The closest $N^+\cdots O^\delta$ distances then are ~3.5 Å for both Asp residues.

Environment of retinal polyene chain. The retinal chain between the C_{11} and C_{14} atoms is tightly packed between residues Met^{86}, Glu^{113}, Gly^{114}, Ala^{117}, and Ala^{292}. Replacement of Ala^{117} by Asp or Phe produces hindrances between

the larger side chains of these residues and 11-*cis*-retinal in the model, which is consistent with the impaired binding of retinal with the corresponding mutant rhodopsins (Nakayama and Khorana, 1991; Zhukovsky et al., 1992). Incorporation of the 13-*cis*-retinal isomer into the binding pocket produces overlap of the 13-methyl group with the Glu^{113} side chain, which is consistent with the inability of 13-*cis*-retinal to be reconstituted with rhodopsin (Liu and Mirzadegan, 1988). However, 7-*cis*, 9-*cis*, and 11-*cis* isomers can be arranged in the pocket because their 13-methyl group has an opposite orientation, toward the main polar cavity of GPCRs between helices II, III, and VII. These 7-*cis* and 9-*cis*-retinal isomers form photoactive pigments and can be detected in irradiated opsins (Maeda et al., 1979). Substituents in the C_{12} position overlap with the Glu^{113} side chain, consistent with the low yield of photopigments reconstituted with 12-methyl or 12-chloro, 11-*cis*-retinal (Liu et al., 1984). On the other hand, rhodopsin can tolerate incorporation of substituents in positions C_5, C_9, C_{10}, and C_{13} of 11-*cis*-retinal (Asato et al., 1989; Liu and Mirzadegan, 1988; Liu and Asato, 1989). This matches the distribution of small empty spaces near the C_5, below the C_{10} and C_{13}, and above the C_9 and C_{11} retinal atoms in our model. The empty space above the retinal in the model may facilitate its *cis-trans* isomerization and dissociation from opsin.

In the rhodopsin model, bulky side chains incorporated in place of Gly^{121} are overlapped with the retinal C_7 atom, β-ionone ring, and rhodopsin Phe^{261} side chain, thus explaining the changes in spectral, chromophore-binding, and transducin-activating features in the G121L mutant (Han et al., 1996b). The properties of the mutant protein can be partially reversed by replacement of Phe^{261}, which interacts with Gly^{121} in our model, by the smaller Ala (Han et al., 1996a). In the G121L/F261A double mutant, the F261A substitution provides extra space for Leu^{121} and retinal.

Environment of the β-ionone ring. The spatial position of the β-ionone ring in the binding pocket is in agreement with chemical cross-linking of a nonisomerizable 11-*cis*-retinal analog with Trp^{265} and Leu^{266} residues (Zhang et al., 1994): the distances between the C_3 retinal atom, the attachment site of the photoreactive diazo group, and backbone nitrogens of Trp^{265} and Leu^{266} are 4.0 and 5.5 Å, respectively, in the model. Replacements of Trp^{265} and Tyr^{268} residues, the side chains of which are in direct contact with the β-ionone ring in the model, have been shown to affect retinal binding, spectroscopic properties of photopigments, and transducin activation (Nakayama and Khorana, 1991). The observed proximity of Glu^{122} to the β-ionone ring (the distance between the C_{17} retinal carbon and $O^{\epsilon 2}$ of Glu^{122} is 4.7 Å) has been suggested from the blue absorption spectra shift in E122Q mutant (Nathans, 1990a; Weitz and Nathans, 1993), from studies of a E122Q mutant by Raman spectroscopy (Lin et al., 1992), and from FTIR studies of artificial pigments with retinal analogs (Jäger et al., 1994b).

In some invertebrate opsins, retinal is substituted by 3-hydroxy-retinal (Vogt, 1987). In the model, a 3-OH substituent in 11-*cis*-retinal can form H bonds with Ser^{208}, commonly found in invertebrate opsins, as has previously been suggested (Oprian, 1992), and a 3-OH substituent in all-*trans*-retinal, after a shift of the β-ionone ring (see below), can interact with Asn, Asp, Ser, or Cys, usually present in position 272 in insect opsins. In the bovine rhodopsin model (Fig. 9), there is an empty space above the β-ionone ring that can adopt small substituents in the C_3 position, which is in agreement with experimental studies of the corresponding artificial pigments (Gärtner et al., 1991).

Spectral tuning of rhodopsin. The significant bathochromic shift of the retinal SB absorption spectrum in rhodopsin (from 380 to 500 nm) originates mostly from its protonation (Honig et al., 1976). Additional spectral tuning of rhodopsin can be provided by distortion of retinal geometry (rotations around single and double bonds) or by electrostatic interactions of the protonated SB with the protein environment (Honig et al., 1975, 1976; Irving et al., 1969; Kliger et al., 1977). In the excited state, the electric dipole moment of the protonated retinal SB changes direction, and the β-ionone ring becomes positively charged (Mathies and Stryer, 1976). Therefore, the presence of a negatively charged protein group near the β-ionone ring, which electrostatically stabilizes the dipole moment of the excited state, would produce a red shift, whereas a negatively charged group near the SB nitrogen would cause the opposite effect. Thus the association of the Glu^{113} counterion with the protonated SB is expected to produce a blue shift; however, this shift is diminished because of the relatively long distance (4.1 Å) between counterion ($O^{\epsilon 2}$) and imine nitrogen in the model. Similarly, a blue shift (10 nm) is produced by an additional negatively charged Asp^{292} near the SB (3.1 Å from the imine nitrogen in the model) in a mutant rhodopsin (Nakayama and Khorana, 1991).

It has been shown experimentally that, unlike bacteriorhodopsin and red cone opsins, bovine rhodopsin has no negatively charged groups near the β-ionone ring (Chen et al., 1989), exactly as in our model, taking into account that Glu^{122} is not charged (Fahmy et al., 1993). However, in the model, the β-ionone ring is surrounded by rhodopsin groups with electric dipole moments that stabilize the retinal SB excited state and therefore cause a red shift: the indole ring of Trp^{265}, which is parallel to the retinal polyene chain and has its N^{ϵ}-H bond pointed toward the SB imine group, and an electronegative Glu^{122} carboxyl oxygen, which contacts the β-ionone ring. Replacements of Trp^{265} by Phe and Glu^{122} by Gln cause a 18–22-nm blue absorption shift (Nakayama and Khorana, 1991; Nathans, 1990a,b). Furthermore, F261Y and A269T replacements, which incorporate additional OH groups with oxygens pointed toward the β-ionone ring, lead to 5–10-nm red absorption shifts (Chan et al., 1992); such replacements are important for spectral tuning of red-sensitive cone opsins (Asenjo et al., 1994).

The "deformation mechanism" explanation of the bathochromic shift is less probable, judging from our model, because the 11-*cis*-retinal has no significant distortions compared with its crystal structure: its nonplanar torsion

angles for the C_6-C_7, C_{11}-C_{12}, C_{12}-C_{13}, and C_{15}-N bonds are 57°, 8°, −160°, and −174°, respectively, after energy minimization of bovine rhodopsin with 11-*cis*-retinal. In the crystal structure of 11-*cis*-retinal (Gilardi et al., 1972) the C_6-C_7 and C_{11}-C_{12} angles are 41° and 2°, respectively, but the C_{12}-C_{13} angle (~40°) differs, because the 12-*s-cis* conformer is present in the crystal.

Photoactivation of rhodopsin

Photoisomerization of retinal and flexibility of functionally important side chains

The conformational transition of rhodopsin to its active (metarhodopsin II, Meta II) state is driven by *cis-trans* photoisomerization of 11-*cis*-retinal. 11-*cis*-Retinal is a natural antagonist that locks the inactive conformation of rhodopsin, whereas all-*trans*-retinal acts as an agonist stabilizing the active state, which binds to and activates transducin (Rao and Oprian, 1996). The structure of Meta II with all-*trans*-retinal is especially important, because this is a prototype of the active conformations in all other GPCRs that recognize their natural, agonist ligands, whereas 11-*cis*-retinal is a unique case of a naturally designed antagonist in the family. The straight and longer all-*trans*-retinal isomer can also be incorporated in the binding pocket of the model, but this requires that space be provided for its strongly shifted β-ionone ring by rotating the Trp^{265} side chain and that the side-chain conformer of Lys^{296}, which is covalently linked to the retinal, be changed (Figs. 10 *B* and 11).

Some other changes during photoactivation can be suggested from the possibility of alternative conformations for several side chains that are conserved and important for transduction. The conformations of evolutionarily conserved side chains are usually fixed by packing requirements within the transmembrane α-bundle, with only Trp^{265} and Lys^{296} (within the binding pocket) and Glu^{134}, Tyr^{223}, and Tyr^{306} (near the intracellular side of the α-bundle) able to change their conformers in the model. The Tyr^{223} and Tyr^{306} side chains can be within the α-bundle (when χ^1 is ~−60°) or can be exposed to the cytoplasmic surface (when χ^1 is ~180°), in each case participating in alternative systems of H bonds (Table 5). These residues are important for transduction (Hunyady et al., 1995; Fernandez and Puett, 1996b) and sequestration (Barak et al., 1994; Hunyady et al., 1995) processes. In the Meta II state, both tyrosines probably adopt the more exposed conformer (χ^1 ~180°), because they are accessible to chemical modification in bleached opsin (Davison and Findlay, 1986b). However, calculations with χ^1 angles of ~180° yield holes within the transmembrane α-bundle, which are filled by the side chains in their alternative conformations with $\chi^1 \approx -60°$. Therefore we suggest that these tyrosines rotate from $\chi^1 = -60°$ to $\chi^1 = 180°$ during activation.

Depending on its χ^2 angle, the Glu^{134} side chain can form H bonds with Arg^{135} or with His^{152} in the rhodopsin model. It has previously been suggested that an ionic interaction between Arg^{135} and Glu^{134} locks the receptor in its inactive conformation, because replacement of the negatively charged Glu^{134} by neutral Gln, eliminating this ionic interaction, leads to the constitutive activation of rhodopsin (Sakmar et al., 1989; Cohen et al., 1993); the corresponding D134A mutant of the α_{1b} adrenergic receptor is also constitutively active (Scheer et al., 1996). Thus the Glu^{134} conformer that forms an H bond with Arg^{135} can be assigned to the inactive state. The formation, in the active state, of the alternative H bond between the negatively charged Glu^{134} side chain and His^{152}, which electrostatically stabilizes the positively charged state of His^{152}, is consistent with the observed uptake of at least one proton by a histidine side chain in the Meta II state (Mathews et al.,

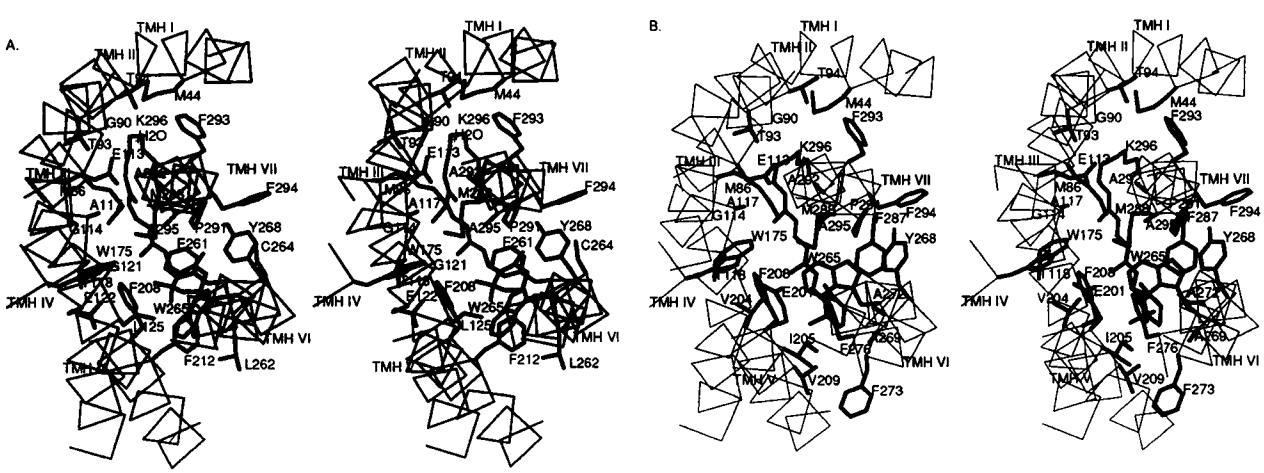

FIGURE 10 The binding pockets of 11-*cis*-retinal in rhodopsin (*A*) and all-*trans*-retinal in the metarhodopsin II (*B*) models (stereoviews). Twenty-seven residues within 4.5 Å of 11-*cis*-retinal, and 29 residues within 4.5 Å of all-*trans*-retinal are shown. All-*trans*-retinal shown with 6s-*cis* and C=N *syn* conformations (dihedral angles C_6-C_7 and C_{15}=N are 50° and 29°, respectively), although the alternative 6s-*trans* and C=N *anti* conformations are also possible in the model. The orientation of the retinal SB may vary slightly, because its NH group can be shifted by ~1 Å without causing hindrances in the Meta II model.

1963). The proton uptake disappears in a Glu/Gln^{134} rhodopsin mutant (Arnis et al., 1994), presumably because the uncharged Gln^{134} residue does not stabilize the protonated state of His^{152}. The possibility that Glu^{134} itself absorbs a proton was not supported by FTIR spectroscopy studies (Sakmar and Fahmy, 1995).

Metarhodopsin II model

The Meta II model was calculated with DIANA, using the alternative conformers and H bonds of Glu^{134}, Tyr^{223}, Trp^{265}, Lys^{296}, and Tyr^{306} side chains suggested above. The resulting model can accommodate all-*trans*-retinal (Fig. 10 *B*), yet is very close to the rhodopsin structure (r.m.s.d of C^{α} atoms is 0.6 Å).

Rearrangements around retinal SB

In the Meta II model the retinal SB is shifted by ~1.5 Å toward helix V, which may be associated with the deprotonation of the SB and protonation of Glu^{133} (Jäger et al., 1994a; Smith et al., 1992) (Fig. 10 *B*). The center of the β-ionone ring is shifted by 7.1 Å (Figure 11) from its initial position in rhodopsin toward the extracellular side and occupies an alternative binding site formed by aliphatic Val^{204}, Ile^{205}, Val^{209}, Ala^{269}, Ala^{272}, and Pro^{291} and aromatic Phe^{208}, Trp^{265} Tyr^{268}, Phe^{273}, Phe^{276}, and Phe^{287} side chains (Fig. 10). The shift of the β-ionone ring makes retinal more parallel to the membrane plane, in agreement with linear dichroism data (Chabre and Breton, 1979b). The displacement of the β-ionone ring is coupled in the model with the reorientation of the Trp^{265} side-chain conformer from *trans* to *gauche*$^-$ (Fig. 11). The ^1L_b transition dipole moment of the Trp^{265} indole ring is closer to the normal of the membrane plane in the *trans* conformer of the Trp^{265} side chain (inactive state), but becomes nearly parallel to the membrane plane in the *gauche*$^-$ conformer. This is consistent with the changes in the linear dichroism spectra of tryptophans during photoactivation of bovine and frog rhodopsins (Chabre and Breton, 1979a). The strongly shifted β-ionone ring and the 9-methyl group of all-*trans*-retinal force the Trp^{265} indole ring deeper into the binding pocket (Fig. 11). This may be crucially important for transduction, judging from observations that incorporation of retinal analogs lacking the 9-methyl group or β-ionone ring (Ganter et al., 1989; Jäger et al., 1994b), or W265F and W265A mutations (Nakayama and Khorana, 1991) impair photoactivation of rhodopsin. The rotation of the Trp^{265} side chain is also in agreement with changes in UV absorption spectra of Trp^{126} and Trp^{265} (Lin and Sakmar, 1996) and in the NH stretching frequency of an undefined indole ring (Kandori and Maeda, 1995) during photoactivation of rhodopsin. The Glu^{122} COOH group, situated near the β-ionone ring in the model, is also sensitive to its movement: the photoinduced change of its protonation pattern was observed in native rhodopsin (Fahmy et al., 1993), but not in artificial pigments reconstituted with retinal analogs lacking the β-ionone ring (Jäger et al., 1994b).

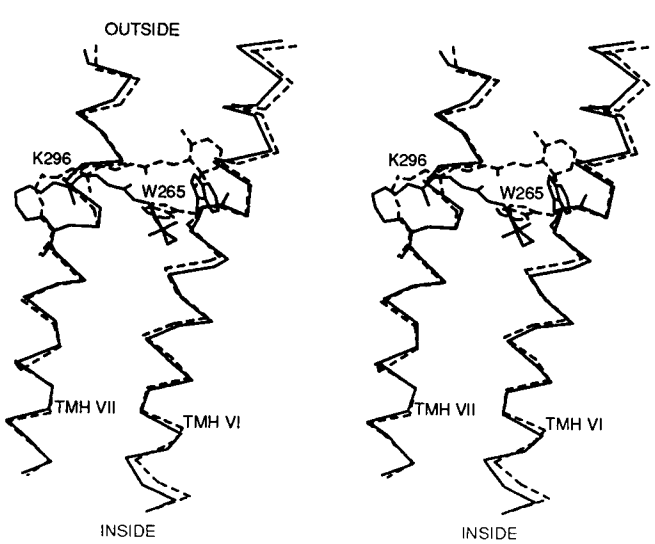

FIGURE 11 Superposition of rhodopsin (——) and Metarhodopsin II (– – –) models with 11-*cis* and all-*trans*-retinal, respectively. Only fragments of helices VI and VII are shown.

Further conformational changes

The proposed rotation of the Trp^{265} indole ring slightly reorients the Phe^{261} aromatic ring, and through the interaction between Phe^{261} and Ser^{301} can influence the side-chain packing in the hydrophilic cluster formed by Asn^{55}, Asp^{83}, Asn^{302}, and Tyr^{306}. The rearrangement of the polar side chains is facilitated by the presence of the central, water-filled cavity in the rhodopsin structure. This rearrangement likely causes rotation of the Tyr^{306} side chain, as described above, and formation of a Tyr^{306} O$^{\eta}$H···O Asn^{73} H bond instead of the Tyr^{306} O$^{\eta}$H···O Asn^{302} H bond, present in the inactive conformation of rhodopsin (Table 5). As a result, the Asn^{302} side chain, losing its H-bonding partner, Tyr^{306}, forms a compensatory H bond with Asp^{83} in Meta II (Table 5). This is consistent with FTIR spectroscopy data, which indicate that the protonated COOH group of Asp^{83} forms an H bond in Meta II (Rath et al., 1993). All residues in the hypothetical transduction pathway, Asp^{83}, Ser^{124} (replaced by Ala in bovine rhodopsin), Phe^{261}, Trp^{265}, Asn^{302}, and Tyr^{306}, are highly conserved in the GPCR family, and their importance for activation has been shown by site-directed mutagenesis (Nakayama and Khorana, 1991; Chanda et al., 1993; Ceresa and Limbird, 1994; Laue et al., 1995; Monnot et al., 1996; Fernandez and Puett, 1996b).

The conformational changes must finally be transmitted to the intracellular surface of the α-bundle, where conserved Asn^{73}, Glu^{134}, Arg^{135}, Tyr^{136}, Val^{139}, Tyr^{223}, Lys^{248}, Arg^{252}, and Tyr^{306} residues form an almost continuous surface and may interact with G-proteins: it is known that replacement of Glu^{134}, Arg^{135}, Tyr^{223}, and Tyr^{306} residues affects G-protein coupling (Sakmar et al., 1989; Cohen et al., 1993; Fahmy et al., 1993; Hunyady et al., 1995; Fer-

nandez and Puett, 1996b), and that residues 136, 139, 248, and 252 are probably not obscured by loops because they can be chemically modified in ovine rhodopsin (Barclay and Findlay, 1984; Ridge et al., 1995). The proposed rotation of Tyr^{223} and Tyr^{306} side chains during activation creates cavities near the intracellular surface.

The changes of side-chain conformers and H bonds for Glu^{134}, Tyr^{223}, Trp^{265}, Lys^{296}, and Tyr^{306} residues produce small systematic shifts of some helices in the Meta II model, which can be identified after superposition of the sets of structures with low target function representing rhodopsin and Meta II. The shifts are ~1 Å for helix VI and ~0.6 Å for helices III and VII. Although these shifts are within the error of the modeling procedure, experimental data suggest that helices VI and III do change their spatial positions during photoactivation. Replacements of several residues in helix VI and in helices III and V (at their interfaces with helix VI) have produced a number of constitutively active mutants in various GPCRs (see the review of Rao and Oprian, 1996), thus indicating a possible movement of helix VI during activation, which is consistent with spin-labeling data (Altenbach et al., 1996). The shift of helix III in Meta II was experimentally detected as an increased flexibility of the Cys^{140} side chain, observed by electron paramagnetic resonance spectroscopy (Pogozheva et al., 1985; Farahbakhsh et al., 1993; Resek et al., 1993). Trp^{126}, at the interface of helices III and IV, is also structurally active during photoactivation (Lin and Sakmar, 1996). It is possible that activation of some GPCRs involves the thiol-disulfide exchange of cysteines from helices III and IV. Two opsins, three melanocortin receptors, and one high-affinity interleukin-8 receptor (Table 2) have one cysteine in helix III (in position 122, 126, or 130) and two proximal cysteines in helix IV (Cys^{167} and Cys^{168}, Cys^{163} and Cys^{164}, Cys^{157} and Cys^{160}, respectively). In these cases the Cys from helix III can form a disulfide bond with either of the two corresponding cysteines from helix IV. Cys^{122}-Cys^{168}, Cys^{126}-Cys^{164}, and Cys^{130}-Cys^{157} disulfide bonds are more consistent with other constraints, incorporated in the "average" inactive conformation. The alternative disulfide bonds may be formed by thiol-disulfide exchange, if the position of helix III relative to helix IV changes slightly during activation of these receptors. This hypothesis can be experimentally tested by mutagenesis.

Rhodopsin and bacteriorhodopsin

The rhodopsin model and the published refined bacteriorhodopsin structure (2brd PDB file; Grigorieff et al., 1996) differ in the positions of helices IV and V and in the tilts of helices II and III. Nonetheless, superposition of the structures produces an r.m.s.d. of 2.9 Å for 140 common C^{α} atoms from all seven helices. (The helical segments 12–33, 41–61, 84–104, 108–123, 139–156, 167–187, and 207–227 of bacteriorhodopsin were superimposed with segments 36–57, 78–98, 112–132, 161–176, 202–219, 250–270, and 287–307, respectively, of rhodopsin.) Thus the final model of rhodopsin is closer to the bacteriorhodopsin structure than to the initial "crude" model designed here for rhodopsin itself (r.m.s.d. = 4.1 Å) or to any previously proposed GPCR model (see above). Remarkably, the same superposition, but using only 67 C^{α} atoms of the central four helices, II, III, VI, and VII (r.m.s.d. is 2.0 Å), yields close spatial positions of all-*trans*-retinals and seven functionally important surrounding residues in the Meta II and bacteriorhodopsin models (Fig. 12): Lys^{296} and Lys^{216}, which form the retinal SB; Glu^{113} and Asp^{85} counterions; Met^{44} and Met^{20} (their S^{δ} atoms surround the imine group of the SB); Trp^{265} and Trp^{182}; $Thr^{118}$and Thr^{90}; Tyr^{268} and Tyr^{185}; and Ala^{295} and Ala^{215} (r.m.s.d. = 2.7 Å for 46 nonhydrogen atoms of side chains and retinals, excluding the β-ionone ring, which is shifted closer to the extracellular surface in bacteriorhodopsin). All of these residues from the binding pocket are conserved in eukaryotic and bacterial photopigments, except Met^{44}, Glu^{113}, and Thr^{118} of bovine rhodopsin, which are replaced in invertebrate rhodopsins, and Met^{20} and Asp^{85} of bacteriorhodopsin, which are replaced in bacterial halorhodopsins. Asp^{212}, the second counterion of the protonated SB in bacteriorhodopsin, is replaced by Ala^{292} in rhodopsin (Fig. 12), although the corresponding A292D (or A292E) replacement has been found in patients with congenital night blindness (Rao and Oprian, 1996).

Other similar side chains also spatially substitute for each other in the retinal-binding pocket, even though they come from different positions in the amino acid sequences. The side chains of Phe^{293} (helix VII) and Tyr^{57} (helix II), of Phe^{276} (helix VI) and Trp^{189} (helix VI, one turn apart), and of Phe^{287} (helix VII) and Phe^{206} (helix VII, adjacent position) from rhodopsin and bacteriorhodopsin, respectively, are spatially overlapped and interact with the same groups of retinal. It is also noteworthy that Asp^{115} (helix IV), which

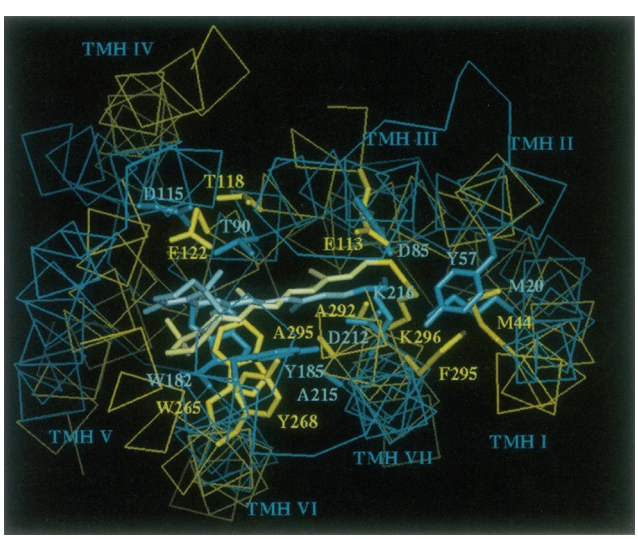

FIGURE 12 Superposition of bovine Metarhodopsin II model (*yellow*) and bacteriorhodopsin structure (*blue*, structure from PDB file 2brd) (Grigorieff et al., 1996). Only several selected side chains are shown.

is conserved in the bacteriorhodopsin family, spatially substitutes for Glu^{122} (helix III), which is in contact with the β-ionone ring of 11-*cis*-retinal in rhodopsin.

The cytoplasmic part of the bacteriorhodopsin proton-transfer pathway, near Asp^{96}, Phe^{219}, and Thr^{46}, spatially coincides with the largest water-filled cavity in the rhodopsin model, which probably forms a regulatory Na^+-binding site in other GPCRs (Asp^{83}, Ser^{124}, and Asn^{302} residues; see above). The identical spatial positions of the binding sites of rhodopsin and bacteriorhodopsin are to be expected, because this is typical for evolutionarily remote, related proteins that have lost the similarity of their amino acid sequences but still maintain the original 3D fold (Farber and Petsko, 1990; Murzin, 1993, 1996).

CONCLUSIONS

We have described the development of an "average" seven helix bundle model that has been calculated using H bonds of intramembrane polar side chains from various rhodopsin-like GPCRs and can be used to facilitate modeling of the transmembrane domain of any protein from this family simply by homology. The "average" model allows H bonding by almost every buried polar side chain simultaneously in 410 various receptors, although only some of the possible H bonds have been applied in the distance geometry calculations. The model also indicates a close relatedness of all considered receptors: 40 of their conserved residues form a single continuous domain consisting of an inner "minicore" of 13 aliphatic side chains and six smaller peripheral clusters of polar and aromatic residues. The corresponding model of bovine rhodopsin has many interesting structural features, such as layers of lipid-facing aromatic side chains, shielding of buried highly polar groups from the aliphatic surrounding by aromatic and sulfur-containing groups with intermediate polarity, four small polar cavities that are probably filled by several water molecules, and a narrow and deep binding "cleft," with the bottom complementary to 11-*cis*-retinal and an empty space above, which allows isomerization and dissociation of the chromophore. Both 11-*cis* and all-*trans* isomers of retinal can be inserted in the binding pocket, but their β-ionone rings must occupy two alternative sites arising from two different orientations of the Trp^{265} side chain. It is proposed that the β-ionone ring and Trp^{265} side chain change their spatial positions after photoactivation of rhodopsin, triggered by retinal isomerization, consistent with published linear dichroism and other data. Two corresponding alternative binding sites for agonists and antagonists may also be present in other GPCRs.

The model is in agreement with a vast sample of published experimental data, which were not considered during calculations and which thus serve as an independent control. These data include the arrangement of α-helices in the low-resolution 3D EM structures; mapping of water- and lipid-accessible rhodopsin residues by chemical probes; identification of residues surrounding retinal by site-directed mutagenesis and cross-linking; determination of retinal orientation relative to the membrane plane and its distances to intra- and extracellular surfaces; reconstitution studies of opsin with synthetic retinal analogs; the conformation and surrounding of the protonated retinal SB; the compensatory replacement of the Glu^{113} counterion by Asp^{90} or Asp^{117} residues; and others.

The model of bovine rhodopsin is also supported by its comparison with the experimentally determined structure of bacteriorhodopsin. The independently developed models of eukaryotic and bacterial photoreceptors are different, but have a striking similarity in the vicinity of the binding pockets, as is always observed for proteins with common 3D folds and chemically identical ligands.

The models of rhodopsin and metarhodopsin II have been deposited in the Protein Data Bank (1bok and 1boj files, respectively) to make them available for further verification or refinement, for example, by direct fit with 3D EM maps (unavailable at this time) or by incorporating artificial disulfide bonds between spatially close residues (for example, from Table 2) using site-directed mutagenesis.

This work was supported by grants DA03910, DA09989, and DA00118 from the National Institutes of Health and by an Upjohn Research Award from the College of Pharmacy of the University of Michigan. We are grateful to Molecular Simulations, Inc. for making their molecular modeling software available to us.

REFERENCES

Altenbach, C., K. Yang, D. L. Farrens, Z. T. Farahbakhsh, H. G. Khorana, and W. L. Hubbell. 1996. Structural features and light-dependent changes in the cytoplasmic interhelical E-F loop region of rhodopsin: a site-directed spin-labeling study. *Biochemistry.* 35:12470–12478.

Altschuh, D., A. M. Lesk, A. C. Bloomer, and A. Klug. 1987. Correlation of co-ordinated amino acid substitutions with function in viruses related to tobacco mosaic virus. *J. Mol. Biol.* 193:693–707.

Arnis, S., K. Fahmy, K. P. Hofmann, and T. P. Sakmar. 1994. A conservative carboxylic acid group mediates light-dependent proton uptake and signaling by rhodopsin, *J. Biol. Chem.* 269:23879–23881.

Asato, A. E., B.-W. Zhang, M. Denny, T. Mirzadegan, K. Seff, and R. S. H. Liu. 1989. A study of the binding site requirements of rhodopsin using isomers of alpha-retinal and 5-substituted alpha-retinal analogs. *Bioorg. Khim.* 17:410–421.

Asenjo, A. B., J. Rim, and D. D. Oprian. 1994. Molecular determinants of human red/green color discrimination. *Neuron.* 12:1131–1138.

Bagley, K. A., V. Balogh-Nair, A. A. Croteau, G. Dollinger, T. G. Ebrey, L. Eisenstein, M. K. Hong, K. Nakanishi, and J. Vittitow. 1985. Fourier transform infrared difference spectroscopy of rhodopsin and its photoproducts at low temperature. *Biochemistry.* 24:6055–6071.

Baker, E. N. and R. E. Hubbard. 1984. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* 44:97–179.

Baldwin, J. M. 1993. The probable arrangement of the helices in G protein-coupled receptors. *EMBO J.* 12:1693–1703.

Baldwin, J. M. 1994. Structure and function of receptors coupled to G proteins. *Curr. Opin. Cell Biol.* 6:180–190.

Barak, L. S., M. Tiberi, N. J. Freedman, M. M. Kwatra, R. J. Lefkowitz, and M. G. Caron. 1994. A highly conserved tyrosine residue in G-protein-coupled receptors is required for agonist-mediated β_2-adrenergic receptor sequestration. *J. Biol. Chem.* 269:2790–2795.

Barclay, P. L. and J. B. C. Findlay. 1984. Labelling of the cytoplasmic domains of ovine rhodopsin with hydrophilic chemical probes. *Biochem. J.* 220:75–84.

Barlow, D. J. and J. M. Thornton. 1988. Helix geometry in proteins. *J. Mol. Biol.* 201:601–619.

Bernstein, F. C., T. F. Koetzle, G. J. Williams, E. E. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Blaber, M., X.-J. Zhang, J. D. Lindstrom, S. D. Pepiot, W. A. Baase, and B. W. Matthews. 1994. Determination of α-helix propensity within the context of a folded protein. Sites 44 and 131 in bacteriophage T4 lysozyme. *J. Mol. Biol.* 235:600–624.

Borjigin, J. and J. Nathans. 1994. Insertional mutagenesis as a probe of rhodopsin's topography, stability, and activity. *J. Biol. Chem.* 269: 14715–14722.

Brooks, B. R., E. R. Bruccoleri, E. R. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* 4:187–217.

Callender, R. H., A. Doukas, R. Crouch, and K. Nakanishi. 1976. Molecular flow resonance Raman effect from retinal and rhodopsin. *Biochemistry.* 15:1621–1629.

Ceresa, B. P. and L. E. Limbird. 1994. Mutation of an aspartate residue highly conserved among G-protein-coupled receptors results in nonreciprocal disruption of alpha-2-adrenergic receptor-G-protein interactions. A negative charge at amino acid residue. 79 forecasts alpha-2A-adrenergic receptor sensitivity to allosteric modulation by monovalent cations and fully effective receptor/G-protein coupling. *J. Biol. Chem.* 269:29557–29564.

Chabre, M., and J. Breton. 1979a. Orientation of aromatic residues in rhodopsin. Rotation of one tryptophan upon the Meta I to Meta II transition after illumination. *Photochem. Photobiol.* 30:295–299.

Chabre, M., and J. Breton. 1979b. The orientation of the chromophore of vertebrate rhodopsin in the "Meta" intermediate states, and the reversibility of the Meta II–Meta III transition. *Vision Res.* 19:1005–1018.

Chan, T., M. Lee, and T. P. Sakmar. 1992. Introduction of hydroxyl-bearing amino acids causes bathochromic spectral shifts in rhodopsin. Amino acid substitutions responsible for red-green color pigment spectral tuning. *J. Biol. Chem.* 267:9478–9480.

Chanda, P. K., M. C. Minchin, A. R. Davis, L. Greenberg, Y. Reilly, W. H. McGregor, R. Bhat, M. D. Lubeck, S. Mizutani, and P. P. Hung. 1993. Identification of residues important for ligand binding to the human 5-hydroxytryptamine 1A serotonin receptor. *Mol. Pharmacol.* 43:516–520.

Chen, J. G., T. Nakamura, T. G. Ebrey, H. Ok, K. Konno, F. Derguini, K. Nakanishi, and B. Honig. 1989. Wavelength regulation in iodopsin, a cone pigment. *Biophys. J.* 55:725–729.

Chothia, C., and A. M. Lesk. 1986. The relation between the divergence of sequence and structure in proteins. *EMBO J.* 5:823–826.

Clarke, J., K. Henrick, and A. R. Fersht. 1995. Disulfide mutants of barnase I: changes in stability and structure assessed by biophysical methods and X-ray crystallography. *J. Mol. Biol.* 253:493–504.

Cohen, G. B., T. Yang, P. R. Robinson, and D. D. Oprian. 1993. Constitutive activation of opsin: influence of charge at position 134 and size at position 296. *Biochemistry.* 32:6111–6115.

Davies, A., G. F. X. Schertler, B. E. Gowen, and H. R. Saibil. 1996. Projection structure of an invertebrate rhodopsin. *J. Struct. Biol.* 117: 36–44.

Davison, M. D., and J. B. C. Findlay. 1986a. Modification of ovine opsin with the photosensitive hydrophobic probe 1-azido-4-[^{125}I]iodobenzene. Labelling of the chromophore-attachment domain. *Biochem. J.* 234: 413–420.

Davison, M. D., and J. B. C. Findlay. 1986b. Identification of the sites in opsin modified by photoactivated azido[^{125}I]iodobenzene. *Biochem. J.* 236:389–395.

Deisenhofer, J., and H. Michel. 1991. Structures of bacterial photosynthetic reaction centers. *Annu. Rev. Cell Biol.* 7:1–23.

Deng, H., L. Huang, R. Callender, and T. Ebrey. 1994. Evidence for a bound water molecule next to the retinal Schiff base in bacteriorhodopsin and rhodopsin: a resonance Raman study of the Schiff base hydrogen/deuterium exchange. *Biophys. J.* 66:1129–1136.

Dohlman, H. G., M. Bouvier, J. L. Benovic, M. G. Caron, and R. J. Lefkowitz. 1987. The multiple membrane spanning topography of the beta 2-adrenergic receptor. Localization of the sites of binding, glycosylation, and regulatory phosphorylation by limited proteolysis. *J. Biol. Chem.* 262:14282–14288.

Donnelly, D., and J. B. C. Findlay. 1994. Seven-helix receptors: structure and modelling. *Curr. Opin. Struct. Biol.* 4:582–589.

Donnelly, D., J. B. C. Findlay, and T. L. Blundell. 1994. The evolution and structure of aminergic G protein-coupled receptors. *Receptors Channels.* 2:61–78.

Donnelly, D., J. P. Overington, S. V. Ruffle, J. H. A. Nugent, and T. L. Blundell. 1993. Modeling α-helical transmembrane domains: the calculation and use of subsitution tables for lipid-facing residues. *Protein Sci.* 2:55–70.

Elling, C. E., S. M. Nielsen, and T. W. Schwartz. 1995. Conversion of antagonist-binding site to metal-ion site in the tachykinin NK-1 receptor. *Nature.* 374:74–77.

Fahmy, K., F. Jäger, M. Beck, T. A. Zvyaga, T. P. Sakmar, and F. Seibert. 1993. Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin. II. A Fourier-transform infrared spectroscopy study of site-directed mutants. *Proc. Natl. Acad. Sci. USA.* 90:10206–10210.

Fanelli, F., M. C. Menziani, M. Cocchi, and P. G. De Benedetti. 1995a. Comparative molecular dynamics study of the seven-helix bundle arrangment of G-protein coupled receptors. *J. Mol. Struct. (Theochem.).* 333:49–69.

Fanelli, F., M. C. Menziani, and P. G. De Benedetti. 1995b. Computer simulations of signal transduction mechanism in α_{1B}-adrenergic and m3-muscarinic receptors. *Protein Eng.* 8:557–564.

Farahbakhsh, Z. T., K. Hideg, and W. L. Hubbell. 1993. Photoactivated conformational changes in rhodopsin: a time resolved spin label study. *Science.* 262:1416–1419.

Farahbakhsh, Z. T., K. D. Ridge, H. G. Khorana, and W. L. Hubbell. 1995. Mapping light-dependent structural changes in the cytoplasmic loop connecting helices C and D in rhodopsin: a site-directed spin labeling study. *Biochemistry.* 34:8812–8819.

Farber, G. K., and G. A. Petsko. 1990. The evolution of α/β barrel enzymes. *Trends Biochem. Sci.* 15:228–234.

Fernandez, L. M., and D. Puett. 1996a. Lys^{583} in the third extracellular loop of the lutropin/choriogonadotropin receptor is critical for signaling. *J. Biol. Chem.* 271:925–930.

Fernandez, L. M., and D. Puett. 1996b. Identification of amino acid residues in transmembrane helices VI and VII of the lutropin/choriogonadotropin receptor involved in signaling. *Biochemistry.* 35: 3986–3993.

Ganter, U. M., E. D. Schmid, D. Perez-Sala, R. R. Rando, and F. Siebert. 1989. Removal of the 9-methyl group of retinal inhibits signal transduction in the visual process. A Fourier transform infrared and biochemical investigation. *Biochemistry.* 28:5954–5962.

Gärtner, W., D. Ullrich, and K. Vogt. 1991. Quantum yield of CHAPSO-solubilized rhodopsin and 3-hydroxy retinal containing bovine opsin. *Photochem. Photobiol.* 54:1047–1055.

Gilardi, R. D., I. L. Karle, and J. Karle. 1972. The crystal and molecular structure of 11-cis-retinal. *Acta Crystallogr.* B28:2605–2612.

Goldsmith, T. H. 1990. Optimization, constraint, and history in the evolution of eyes. *Q. Rev. Biol.* 65: 281–322.

Grigorieff, N., T. A. Ceska, K. H. Downing, J. M. Baldwin, and R. Henderson. 1996. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J. Mol. Biol.* 259:393–421.

Güntert, P., W. Brawn, and K. Wüthrich. 1991. Efficient computation of three-dimensional protein structure in solution from NMR data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* 217:517–530.

Han, M., B. S. DeDecker, and S. O. Smith. 1993. Localization of the retinal protonated Schiff base counterion in rhodopsin. *Biophys. J.* 65:899–906.

Han, M., S. W. Lin, M. Minkova, S. O. Smith, and T. P. Sakmar. 1996a. Functional interaction of transmembrane helices 3 and 6 in rhodopsin. Replacement of phenylalanine 261 by alanine causes reversion of phenotype of a glycine 121 replacement mutant. *J. Biol. Chem.* 271: 32337–32342.

Han, M., S. W. Lin, S. O. Smith, and T. P. Sakmar. 1996b. The effects of amino acid replacements of glycine 121 on transmembrane helix 3 of rhodopsin. *J. Biol. Chem.* 271:32330–32336.

Han, M., and S. O. Smith. 1995a. High-resolution structural studies of the retinal-Glu113 intheraction in rhodopsin. *Biophys. Chem.* 56:23–29.

Han, M., and S. O. Smith. 1995b. NMR constraints on the location of the retinal chromophore in rhodopsin and bathorodopsin. *Biochemistry.* 34: 1425–1432.

Hänggi, G., and W. Braun. 1994. Pattern recognition and self-correcting distance geometry calculations applied to myohemerythrin. *FEBS Lett.* 334:147–153.

Hara-Nishimura, I., M. Kondo, M. Nishimura, R. Hara, and T. Hara. 1993. Amino acid sequence surronding the retinal-binding site in retinochrome of the squid, *Todarodes pacificus. FEBS Lett.* 335:94–98.

Hargrave, P. A., J. H. McDowell, R. J. Feldmann, P. H. Atkinson, J. K. M. Rao, and P. Argos. 1984. Rhodopsin's protein and carbohydrate structure: selected aspects. *Vision Res.* 24:1487–1499.

Henderson, R., J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing. 1990. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* 213: 899–929.

Herzyk, P., and R. E. Hubbard, 1995. Automated method for modeling seven-helix transmembrane receptors from experimental data. *Biophys. J.* 69:2419–2442.

Honig, B., A. D. Greenberg, U. Dinur, and T. G. Ebrey. 1976. Visual-pigment spectra: implications of the protonation of the retinal Schiff base. *Biochemistry.* 15:4593–4599.

Honig, B., A. Warsel, and M. Karplus. 1975. Theoretical studies of the visual chromophore. *Accounts Chem. Res.* 8:92–100.

Horstman, D. A., S. Brandon, A. L. Wilson, C. A. Guyer, E. J. Cragoe, Jr., and L. E. Limbird. 1990. An aspartate conserved among G-protein receptors confers allosteric regulation of alpha 2-adrenergic receptors by sodium. *J. Biol. Chem.* 265:21590–21595.

Hunyady, L., M. Bor, T. Balla, and K. J. Catt. 1995. Critical role of a conserved intramembrane tyrosine residue in angiotensin II receptor activation. *J. Biol. Chem.* 270:9702–9705.

Irving, C. S., G. W. Byers, and P. A. Leermakers. 1969. Effect of solvent polarizability on the absorption spectrum of all-trans-retinylpyrrolidiniminium perchlorate. *J. Am. Chem. Soc.* 91:2141–2143.

Jäger, F., K. Fahmy, T. P. Sakmar, and F. Siebert. 1994a. Identification of glutamic acid 113 as the Schiff base proton acceptor in the metarhodopsin II photointermediate of rhodopsin. *Biochemistry.* 33:10878–10882.

Jäger, F., S. Jäger, O. Krutle, N. Friedman, M. Sheves, K. P. Hofmann, and F. Siebert. 1994b. Interactions of the β-ionone ring with the protein in the visual pigment rhodopsin control the activation mechanism. An FTIR and fluorescence study on artificial vertebrate rhodopsins. *Biochemistry.* 33:7389–7397.

Jones, D. T., W. R. Taylor, and J. M. Thornton. 1994. A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry.* 33:3038–3049.

Kandori, H., and A. Maeda. 1995. A. FTIR spectroscopy reveals microscopic structural changes of the protein around the rhodopsin chromophore upon photoisomerization. *Biochemistry.* 34:14220–14229.

Kaushal, S., and H. G. Khorana. 1994. Structure and function in rhodopsin. 7. Point mutations associated with autosomal dominant retinitis pigmentosa. *Biochemistry.* 33:6121–6128.

Kim, J., J. Wess, A. M. vanRhee, T. Schöneberg, and K. A. Jacobson. 1995. Site-directed mutagenesis identifies residues involved in ligand recognition in the human A_{2a} adenosine receptor. *J. Biol. Chem.* 270: 13987–13997.

Kliger, D. S., S. J. Milder, and E. A. Dratz. 1977. Solvent effects on the spectra of retinal Schiff bases. I. Models for the bathochromic shift of the chromophore spectrum in visual pigments. *Photochem. Photobiol.* 25:277–286.

Kong, H., K. Raynor, K. Yasuda, G. I. Bell, and T. Reisine. 1993a. Mutation of an aspartate at residue 89 in somatostatin receptor subtype 2 prevents Na^+ regulation of agonist binding but does not alter receptor-G protein association. *Mol. Pharmacol.* 44:380–384.

Kong, H., K. Raynor, K. Yasuda, S. T. Moe, P. S. Portoghese, G. I. Bell, and T. Reisine. 1993b. A single residue, aspartic acid 95, in the delta opioid receptor specifies selective high affinity agonist binding. *J. Biol. Chem.* 268:23055–23058.

Lamba, O. P., D. Borchman, and P. J. O'Brien. 1994. Fourier transform infrared study of the rod outer segment disk and plasma membranes of vertebrate retina. *Biochemistry.* 33:1704–1712.

Laue, L., W. Y. Chan, A. J. Hsueh, M. Kudo, S. Y. Hsu, S. M. Wu, L. Blomberg, and G. B. Cutler, Jr. 1995. Genetic heterogeneity of constitutively activating mutations of the human luteinizing hormone receptor in familial male-limited precocious puberty. *Proc. Natl. Acad. Sci. USA.* 92:1906–1910.

Liebman, P. A. 1962. In situ microspectrophotometric studies on the pigment of single retinal rods. *Biophys. J.* 2:161–178.

Lin, S. W., Y. Imamoto, Y. Fukada, Y. Shichida, T. Yoshizawa, and R. A. Mathies. 1994. What makes red visual pigments red? A resonance Raman microprobe study of retinal chromophore structure in iodopsin. *Biochemistry.* 33:2151–2160.

Lin, S. W., and T. P. Sakmar. 1996. Specific tryptophan UV-absorbance changes are probes of the transition of rhodopsin to its active state. *Biochemistry.* 35:11149–11159.

Lin, S. W., T. P. Sakmar, R. R. Franke, H. G. Khorana, and R. A. Mathies. 1992. Resonance Raman microprobe spectroscopy of rhodopsin mutants: effect of substitutions in the third transmembrane helix. *Biochemistry.* 31:5105–5111.

Liu, R. S. H., and A. Asato. 1989. The binding site of opsin based on analogue studies with isomeric, fluorinated, alkylated, and other modified retinals. *In* Chemistry and Biology of Synthetic Retinoids. M. I. Dawson and W. H. Okamura, editors. CRC Press, Boca Raton, FL. 52–75.

Liu, R. S. H., A. Asato, M. Denny, and D. Mead. 1984. The nature of restrictions in the binding site of rhodopsin. A model study. *J. Am. Chem. Soc.* 106:8298–8300.

Liu, R. S. H., and T. Mirzadegan. 1988. The shape of a three dimensional binding site of rhodopsin based on molecular modeling analysis of isomeric and other visual pigment analogues. Bioorganic studies of visual pigments. *J. Am. Chem. Soc.* 110:8617–8623.

Liu, J., T. Schöneberg, M. van Rhee, and J. Wess. 1995. Mutational analysis of the relative orientation of transmembrane helices I and VII in G-protein-coupled receptors. *J. Biol. Chem.* 270:19532–19539.

Lomize, A. L., V. Yu. Orekhov, and A. S. Arseniev. 1992. Refinement of the spatial structure of the gramicidin A transmembrane ion-channel. *Bioorg. Khim.* 18:182–200.

Maeda, A., Y. J. Ohkita, J. Sasaki, Y. Shichida, and T. Yoshizawa. 1993. Water structural changes in lumirhodopsin, metarhodopsin I and metarhodopsin II upon photolysis of bovine rhodopsin: analysis by Fourier transform infrared spectroscopy. *Biochemistry.* 32:12033–12038.

Maeda, A., Y. Shichida, and T. Yoshizawa. 1979. Formation of 7-cis and 13-cis retinal pigments by irradiating squid rhodopsin. *Biochemistry.* 18:1449–1453.

Martynov, V. I., M. B. Kostina, M. Y. Feigina, and A. Miroshnikov. 1983. I. Study of the molecular organization of visual rhodopsin in photoreceptor membranes by limited proteolysis. *Bioorg. Khim.* 9:735–745.

Mathews, R. G., R. Hubbard, P. K. Brown, and G. Wald. 1963. Tautomeric forms of metarhodopsin. *J. Gen. Physiol.* 47:215–240.

Mathies, R., and L. Stryer. 1976. Retinal has highly dipolar vertically excited singlet state: implication for vision. *Proc. Natl. Acad. Sci. USA.* 73:2169–2173.

Matsumura, M., and B. W. Matthews. 1991. Stabilization of functional proteins by introduction of multiple disulfide bonds. *Methods Enzymol.* 202:336–356.

McDonald, I. K., and J. M. Thornton. 1994. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* 238:777–793.

Mizobe, T., M. Maze, V. Lam, S. Suryanarayana, and B. K. Kobilka. 1996. Arrangement of transmembrane domains in adrenergic receptors. Similarity to bacteriorhodopsin. *J. Biol. Chem.* 271:2387–2389.

Mollevanger, L. C. P. J., A. P. M. Kentgens, J. A. Pardoen, J. M. L. Courtin, W. S. Veeman, J. Lugtenburg, and W. J. De Grip. 1987. High-resolution solid-state ^{13}C NMR study of carbon C-5 and C-12 of the chromophore of bovine rhodopsin. Evidence for a 6-*s*-*cis* conformation with negative-charge perturbation near C-12. *Eur. J. Biochem.* 163:9–14.

Momany, F. A., R. F. McGuire, A. W. Burgess, and H. A. Scheraga. 1975. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions,

and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.* 79:2361–2181.

Momany, F. A., and R. Rone. 1992. Validation of the general purpose QUANTA3.2/CHARMm force field. *J. Comput. Chem.* 13:888–900.

Monnot, C., C. Bihoreau, S. Conchon, K. M. Curnow, P. Corvol, and E. Clauser. 1996. Polar residues in the transmembrane domains of the type 1 angiotensin II receptor are required for binding and coupling. Reconstitution of the binding site by co-expression of two deficient mutants. *J. Biol. Chem.* 271:1507–1513.

Mumenthaler, C., and W. Braun. 1995. Predicting the helix packing of globular proteins by self-correcting distance geometry. *Protein Sci.* 4:863–871.

Murzin, A. G. 1993. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* 12:861–867.

Murzin, A. G. 1996. Structural classification of proteins: new superfamilies. *Curr. Opin. Struct. Biol.* 6:386–394.

Nakayama, T. A., and H. G. Khorana. 1990. Orientation of retinal in bovine rhodopsin determined by cross-linking using a photoactivatable analog of 11-*cis*-retinal. *J. Biol. Chem.* 265:15762–15769.

Nakayama, T. A., and H. G. Khorana. 1991. Mapping of the amino acids in membrane-embedded helices that interact with the retinal chromophore in bovine rhodopsin. *J. Biol. Chem.* 266:4269–4275.

Nathans, J. 1990a. Determinants of visual pigment absorbance: role of charged amino acid in the putative transmembrane segments. *Biochemistry.* 29:937–942.

Nathans, J. 1990b. Determinants of visual pigment absorbance: identification of the retinylidene Schiff's base counterion in bovine rhodopsin. *Biochemistry.* 29:9746–9752.

Oliveira, L., A. C. M. Paiva, and G. Vriend. 1993. A common motif in G-protein-coupled seven transmembrane helix receptors. *J. Comput. Aided Mol. Des.* 7:649–658.

Oprian, D. D. 1992. The ligand-binding domain of rhodopsin and other G-protein-linked receptors. *J. Bioenerg. Biomembr.* 24:211–217.

Oseroff, A. R., and R. H. Callender. 1974. Resonance Raman spectroscopy of rhodopsin in retinal disk membranes. *Biochemistry.* 13:4243–4248.

Palings, I., J. A. Pardoen, E. van der Berg, C. Winkel, J. Lugtenburg, and R. A. Mathies. 1987. Assignment of fingerprint vibrations in the resonance Raman spectra of rhodopsin, isorhodopsin, and bathorhodopsin: implication for chromophore structure and environment. *Biochemistry.* 26:2544–2656.

Persson, B., and P. Argos. 1994. Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *J. Mol. Biol.* 237: 182–192.

Pittel, Z., and J. Wess. 1994. Intramolecular interactions in muscarinic acetylcholine receptors studied with chimeric m2/m5 receptors. *Mol. Pharmacol.* 45:61–64.

Pogozheva, I. D., V. A. Kuznetsov, V. A. Livshits, I. B. Fedorovich, and M. A. Ostrovsky. 1985. ESR saturation transfer study of photoinduced changes in the hydrophilic regions of rhodopsin. *Biol. Membr. (USSR).* 2:880–896.

Quintana, J., H. Wang, and M. Ascoli. 1993. The regulation of the binding affinity of the luteinizing hormone/choriogonadotropin receptor by sodium ions is mediated by a highly conserved aspartate located in the second transmembrane domain of G-protein-coupled receptors. *Mol. Endocrinol.* 7:767–775.

Rao, V. R., G. B. Cohen, and D. D. Oprian. 1994. Rhodopsin mutation G90D and a molecular mechanism for congenital night blindness. *Nature.* 367:639–642.

Rao, V. R., and D. D. Oprian. 1996. Activating mutations of rhodopsin and other G-protein-coupled receptors. *Annu. Rev. Biophys. Biomol. Struct.* 25:287–314.

Rath, P., L. L. J. DeCaluwé, P. H. M. Bovee-Geurts, W. J. DeGrip, and K. J. Rothschild. 1993. Fourier transform infrared difference spectroscopy of rhodopsin mutants: light activation of rhodopsin causes hydrogen-bonding changes in residue aspartic acid-83 during Meta II formation. *Biochemistry.* 32:10277–10282.

Resek, J. F., Z. T. Farahbakhsh, W. L. Hubbell, and H. G. Khorana. 1993. Formation of the Meta II photointermediate is accompanied by conformational changes in the cytoplasmic surface of rhodopsin. *Biochemistry.* 32:12025–12032.

Ridge, K. D., C. Zhang, and H. G. Khorana. 1995. Mapping of the amino acids in the cytoplasmic loop connecting helices C and D in rhodopsin. Chemical reactivity in the dark state following single cysteine replacements. *Biochemistry.* 34:8804–8811.

Rost, B., R. Casadio, P. Fariselli, and C. Sander. 1995. Transmembrane helices predicted at 95% accuracy. *Protein Sci.* 4:521–533.

Sakmar, T. P., and K. Fahmy. 1995. Properties and photoactivity of rhodopsin mutants. *Isr. J. Chem.* 35:325–337.

Sakmar, T. P., R. R. Franke, and H. G. Khorana. 1989. Glutamic acid-113 serves as a retinylidene Schiff base counterion in bovine rhodopsin. *Proc. Natl. Acad. Sci. USA.* 86:8309–8313.

Sakmar, T. P., R. R. Franke, and H. G. Khorana. 1991. The role of the retinylidene Schiff base counterion in rhodopsin in determining wavelength absorbance and Schiff base pKa. *Proc. Natl. Acad. Sci. USA.* 88:3079–3083.

Sealfon, S. C., L. Chi, B. J. Ebersole, V. Rodic, D. Zhang, J. A. Ballesteros, and H. Weinstein. 1995. Related contribution of specific helix 2 and 7 residues to conformational activation of the serotonin 5-HT_{2A} receptor, *J. Biol. Chem.* 270:16683–16688.

Scheer, A., F. Fanelli, T. Costa, P. G. De Benedetti, and S. Cotecchia. 1996. Constitutively active mutants of the α_{1B}-adrenergic receptor: role of highly conserved polar amino acids in receptor activation. *EMBO J.* 15:3566–3578.

Schertler, G. F. X., and P. A. Hargrave. 1995. Projection structure of frog rhodopsin in two crystal forms. *Proc. Natl. Acad. Sci. USA.* 92: 11578–11582.

Schertler, G. F. X., P. A. Hargrave, and V. M. Unger. 1996. Three dimentional structure of rhodopsin obtained by electron cryomicroscopy. *Invest. Ophthalmol. Visual Sci.* 37:S805.

Schertler, G. F. X., V. M. Unger, and P. A. Hargrave. 1995. The structure of rhodopsin obtained by cryo-electron microscopy to 7 Å resolution. *Biophys. J.* 68:A21.

Schertler, G. F. X., C. Villa, and R. Henderson. 1993. Projection structure of rhodopsin. *Nature.* 362:770–772.

Schultz, G. E. 1992. Structure-function relationships in the membrane channel porin as based on 1.8. Å resolution crystal structure. *In* Membrane Proteins: Structures, Interactions and Models. A. Pullman, J. Jortner, B. Pullman, editors. Kluwer Academic Publishers, Netherlands. 403–412.

Smith, S. O., H. De Groot, R. Gebhard, and J. Lugtenburg. 1992. Magic angle spinning NMR studies on the metarhodopsin II intermediate of bovine rhodopsin: evidence for an unprotonated Schiff base. *Photochem. Photobiol.* 56:1035–1039.

Smith, S. O., I. Palings, V. Copié, D. P. Raleigh, J. Courtin, J. A. Pardoen, J. Lugtenburg, R. A. Mathies, and R. G. Griffin. 1987. Low-temperature solid-state ^{13}C NMR studies of the retinal chromophore in rhodopsin. *Biochemistry.* 26:1606–1611.

Smith, S. O., I. Palings, M. E. Miley, J. Courtin, H. de Groot, J. Lugtenburg, R. A. Mathies, and R. G. Griffin. 1990. Solid-state NMR studies of the mechanism of the opsin shift in the visual pigment rhodopsin. *Biochemistry.* 29:8158–8164.

Soppa, J. 1994. Two hypotheses—one answer. Sequence comparison does not support an evolutionary link between halobacterial retinal proteins including bacteriorhodopsin and eukaryotic G-protein-coupled receptors. *FEBS Lett.* 342:7–11.

Steinberg, G., M. Ottolenghi, and M. Sheves. 1993. pKa of the protonated Schiff base of bovine rhodopsin. A study with artificial pigments. *Biophys. J.* 64:1499–1502.

Suryanarayana, S., M. von Zastrow, and B. K. Kobilka. 1992. Identification of intramolecular interactions in adrenergic receptors. *J. Biol. Chem.* 267:21991–21994.

Taylor, W. R., D. T. Jones, and N. M. Green. 1994. A method for α-helical integral membrane protein fold prediction. *Proteins Struct. Funct. Genet.* 18:281–294.

Thirstrup, K., C. E. Elling, S. A. Hjorth, and T. W. Schwartz. 1996. Construction of a high affinity zinc switch in the κ-opioid receptor. *J. Biol. Chem.* 271:7875–7878.

Thomas, D. D., and L. Stryer. 1982. Transverse location of the retinal chromophore of rhodopsin in rod outer segment disc membranes. *J. Mol. Biol.* 154:145–157.

Unger, V. M., P. A. Hargrave, and G. F. X. Schertler. 1995. Localization of the transmembrane helices in the three-dimentional structure of frog rhodopsin. *Biophys. J.* 68:A330.

Unger, V. M., and G. F. X. Schertler. 1995. Low resolution structure of bovine rhodopsin determined by electron cryo-microscopy. *Biophys. J.* 68:1776–1786.

Vogt, K. 1987. Chromophores of insect visual pigments. *Photobiochem. Photobiophys.* 2(Suppl.):273–296.

Walther, D., F. Eisenhaber, and P. Argos. 1996. Principles of helix-helix packing in proteins: the helical lattice superposition model. *J. Mol. Biol.* 255:536–553.

Watson, S., and S. Arkinstall. 1994. The G-Protein Linked Receptor Facts Book. Academic Press, San Diego, CA.

Weitz, C. J., and J. Nathans. 1993. Rhodopsin activation: effect on the Metarhodopsin I-Metarhodopsin II equilibrium of neutralization or introduction of charged amino acids within putative transmembrane segments. *Biochemistry.* 32:14176–14182.

Yarden, Y., H. Rodriguez, S. K.-F. Wong, D. R. Brandt, D. C. May, J. Burnier, R. N. Harkins, E. Y. Chen, J. Ramachandran, A. Ullrich, and E. M. Ross. 1986. The avian beta-adrenergic receptor:primary structure and membrane topology. *Proc. Natl. Acad. Sci. USA.* 83:6795–6799.

Yeagle, P. L., J. L. Alderfer, and A. D. Albert. 1995a. Structure of the third cytoplasmic loop of bovine rhodopsin. *Biochemistry.* 34:14621–14625.

Yeagle, P. L., J. L. Alderfer, and A. D. Albert. 1995b. Structure of the carboxy-terminal domain of bovine rhodopsin. *Nature Struct. Biol.* 2:832–834.

Yu, H., M. Kono, T. D. McKee, and D. D. Oprian. 1995. General method for mapping tertiary contacts between amino acid residues in membrane-embedded proteins. *Biochemistry.* 34:14963–14969.

Zhang, H., K. A. Lerro, T. Yamamoto, T. H. Lien, L. Sastry, M. A. Gawinowicz, and K. Nakanishi. 1994. The location of the chromophore in rhodopsin: a photoaffinity study. *J. Am. Chem. Soc.* 116: 10165–10173.

Zhou, W., C. Flanagan, J. A. Ballesteros, K. Konvicka, J. S. Davidson, H. Weinstein, R. P. Millar, and S. C. Sealfon. 1994. A reciprocal mutation supports helix 2 and helix 7 proximity in the gonadotropin-releasing hormone receptor. *Mol. Pharmacol.* 45:165–170.

Zhukovsky, E. A., and D. D. Oprian. 1989. Effect of carboxylic acid side-chain on the absorption maximum of visual pigments. *Science.* 246:928–930.

Zhukovsky, E. A., P. R. Robinson, and D. D. Oprian. 1992. Changing the location of the Schiff base counterion in rhodopsin. *Biochemistry.* 31: 10400–10405.

Zvyaga, T., K. C. Min, M. Beck, and T. P. Sakmar. 1993. Movement of the retinylidene Schiff base counterion in rhodopsin by one helix turn reverses the pH dependence of the metarhodopsin I to metarhodopsin II transition. *J. Biol. Chem.* 268:4661–4667.