Modeling of G-Protein-Coupled Receptors: Application to Dopamine, Adrenaline, Serotonin, Acetylcholine, and Mammalian Opsin Receptors

Susanne Trumpp-Kallmeyer,* Jan Hoflack, Anne Bruinvels, and Marcel Hibert

Marion Merrell Dow Research Institute (MMDRI), BP 447 R/9, 16 Rue d'Ankara, 67009 Strasbourg Cedex, France. Received May 15, 1992

Hydropathicity analysis of 39 G-protein-coupled receptors (GPCR) reveals seven hydrophobic stretches corresponding to membrane spanning α-helices. The alignment of the primary sequences shows a high degree of homology in the GPCR transmembrane regions. 3D models of 39 GPCRs were generated using the refined model of bacteriorhodopsin as a template. Five cationic neurotransmitter receptors (serotonergic 5-HT$_2$, dopaminergic D$_2$, muscarinic m$_2$, adrenergic α_2 and β_2 receptors) were taken as prototypes and studied in detail. The 3D models of the cationic neurotransmitter receptors, together with their primary structure comparison, indicate that the agonist binding site is located near the extracellular face of the receptor and involves residues of the membrane-spanning helices 3, 4, 5, 6, and 7. The binding site consists of a negatively-charged Asp located at the middle of transmembrane helix 3 and a hydrophobic pocket containing conserved aromatic residues on helices 4, 5, 6, and 7. To define the precise receptor–ligand interactions, the natural neurotransmitters were docked into the binding sites. Residues responsible for the affinity, selectivity, and eventually stereospecificity of dopamine, adrenaline, noradrenaline, serotonin, and acetylcholine for their receptors were identified. The ligands are involved in electrostatic interactions as well as hydrogen bonds and specific hydrophobic aromatic interactions. All the GPCRs possess invariant hinge residues, which might be responsible for a conformational change during agonist binding and therefore influence dissociation and association of G-proteins to the receptors. The role of hydrophobic interactions and hydrogen bonds in the conformational change of the receptors, modulating the coupling to the G-protein, is discussed with regard to these residues. The models are in agreement with published data obtained from mutagenesis and labeling studies and represent important working hypotheses to direct future mutagenesis studies. They also enable structure–activity relationship studies and more rational drug design. The 3D models of other G-protein-coupled receptors have been generated in a similar way.

Introduction

Hormonal receptors are of major importance in regulatory physiological processes, and major efforts have been dedicated to their pharmacological characterization. The discovery of receptor classes was generally associated with the discovery of new endogenous ligands although major exceptions exist (e.g. benzodiazepines and opioid receptors).[1] More recently, pharmacological, physiological, and structure–activity relationship studies have demonstrated the existence of receptor subtypes activated by the same neurotransmitter. In the particular case of 5-HT receptors, models of the recognition sites were proposed on the basis of ligand conformational analysis.[2,3] These 3D models highlighted similarities and differences between receptor subtypes. Definitive evidence to resolve this controversial question for the existence of receptor subtypes has come from molecular biology since numerous receptor classes and subtypes have now been cloned, identified with human functional receptors, and found expressed in human tissues.

Site-directed mutagenesis and the study of chimeric receptors have contributed considerably to the understanding of the main functional characteristics of GPCRs; a number of residues and domains which are likely to be involved in the binding of agonists and antagonists, in the coupling with G-protein, and in the desensitization process have been identified.[4] Despite this rapid and decisive progress, the major step of moving to a three-dimensional understanding of the mechanisms which control the potency and selectivity of the ligand binding and the efficacy of the signal transduction at the atomic level remained. Unfortunately, the nondegenerative purification of membrane receptors and their crystallization still remain difficult problems.

A number of observations suggests that all GPCRs evolved from a common ancestor. In earlier studies, it was shown that the G-protein-coupled neurotransmitter and the mammalian opsin receptors share significant similarities in amino acids composition.[5] All GPCRs contain two possible N-glycosylation sites near their amino-terminal region.[6] The carboxy-terminal region contains several serine residues, which represent possible phosphorylation sites.[7] Additionally, the neurotransmitter receptors, as well as the mammalian opsins, have similar hydropathicity profiles with seven hydrophobic stretches most probably corresponding to α-helical regions.[8] From this similarity, it was proposed that adrenaline and retinal bind in a similar manner to the hydrophobic transmembrane helices of their receptors, although the specific binding modes are different: retinal forms a Schiff base with a conserved Lys

(1) Hansch, C., Ed. *Comprehensive Medicinal Chemistry*; Pergamon Press: Oxford, U.K., 1990; Vol. 3.

(2) Hibert, M. F.; Gittos, M. W.; Middlemiss, D. N.; Mir, A. K.; Fozard, J. R. Graphic computer-aided receptor mapping as a predictive tool for drug design: development of potent, selective and stereospecific ligands for the 5-HT$_{1A}$ receptor. *J. Med. Chem.* 1988, *31*, 1087–1093.

(3) Hibert, M. F. Modélisation des sites de reconnaissance sérotoninergique: de la rationalisation à la prédiction. *Actual. Chim. Ther.* 1989, *16*, 37–46.

(4) Strader, C. D.; Sigal, I. S.; Register, R. B.; Candelore, M. R.; Rands, E.; Dixon, R. A. F. Identification of residues required for ligand binding to the β-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* 1987, *84*, 4384–4388.

(5) Cotecchia, S.; Schwinn, D. A.; Randall, R. R.; Lefkowitz, R. J.; Caron, M. G.; Kobilka, B. K. Molecular cloning and expression of the C-DNA for the hamster α-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* 1988, *85*, 7159–7163.

(6) Benovic, J. L.; Mayor, F.; Sumers, R. L.; Caron, M. G.; Lefkowitz, R. J. Light dependent phosphorylation of rhodopsin by β-adrenergic receptor-kinase. *Nature* 1986, *321*, 869–872.

(7) Hargrave, P. A. Rhodopsin chemistry, structure and topography. *Prog. Retinal Res.* 1982, *1*, 1–51.

(8) Kubo, T.; Fukuda, K.; Mikami, A.; Maeda, A.; Takahashi, H.; Mishina, M.; Haga, T.; Ichiyama, A.; Kangawa, K.; Kojima, M.; Matsuo, H.; Hirose, T.; Numa, T. Cloning, sequencing and expression of complementary DNA encoding the muscarinic acetylcholine receptor. *Nature* 1986, *323*, 411–416.

*Send reprint requests to Susanne Trumpp-Kallmeyer, Marion Merrell Dow Research Institute, 16 rue d'Ankara, BP 447 R/9, 67009 Strasbourg Cedex, France.

residue in the seventh transmembrane region, while the positively charged nitrogen of the natural neurotransmitter ligands seems to be involved in electrostatic interactions with the side chain of a negatively charged Asp on the second and/or third transmembrane region.[9] Given this similarity in structure, Pappin and Findlay modeled bovine rhodopsin from a low-resolution electron density map of bacteriorhodopsin.[10]

Recently, a high quality 3D model for bacteriorhodopsin was obtained based on cryomicroscopy experiments.[11] The bacteriorhodopsin receptor consists of seven membrane-spanning helices connected by hydrophilic loops. Despite the lack of sequence homology with GPCRs, the parallel between the overall three-dimensional structural patterns is striking. In addition, the retinal attachment Lys residue is located in approximately the same position on transmembrane helix 7 (TM7) in bacteriorhodopsin as in the mammalian opsins. The generally accepted pattern of seven antiparallel transmembrane helices for the GPCRs thus begs for direct comparison with bacteriorhodopsin. Although they do not belong to the same functional class of receptors, it seems very likely that they belong to the same structural class.

There is now enough evidence to generate reasonable 3D models of GPCRs using molecular modeling techniques. The advent of a refined structural model and the availability of a large number of different GPCR classes allow a more meaningful analysis of the regions directly responsible for GPCR function. We report here in detail our investigation of GPCR primary sequence homology and alignment, the construction of three-dimensional models for all GPCRs, and a more detailed study of the muscarinic m_2, serotonergic 5-HT$_2$, adrenergic α_2 and β_2, and dopaminergic D_2 receptors with their endogeneous ligands using bacteriorhodopsin as a template. (*Note: Preliminary results were published as an accelerated communication.*)[12]

Methods

As a first step in the construction of the GPCR 3D models, exhaustive primary sequence comparison and hydropathicity analysis were required.

The following GPCR sequences were analyzed: human 5-HT$_{1A}$;[13] rat 5-HT$_{1C}$[14] and rat 5-HT$_2$;[15] canine 5-HT$_{1D}$[16]

serotonergic; human D_1[17] and D_2,[18,19] rat D3,[19] human D4[20] and D5[21] dopaminergic; rat α_1,[22] human α_2,[23] β_1,[24] β_2,[25] and β_3[26] adrenergic; human m_1, m_2, m_4,[27] m_5,[28] and pig m_3[29] muscarinic; canine H$_2$[30] histaminic; human r and d opsins;[31,32] bovine NK2;[33] rat NK1[34] and rat NK3;[35] bovine

(9) Hulme, E. C.; Curtis, C. A.; Wheatley, M.; Harris, A. C.; Aitken, A. Localisation and structure of the muscarinic receptor ligand binding site. *Trends Pharmacol. Sci.* **1989**, *10*, 22–25.

(10) Findlay, J. B. C.; Pappin, D. J. C. The opsin family of proteins. *Biochem. J.* **1986**, *238*, 625–642.

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

(12) Hibert, M. F.; Trumpp-Kallmeyer, S.; Bruinvels, A. T.; Hoflack, J. 3D models of cationic neurotransmitter G-protein coupled receptors. *Mol. Pharmacol.* **1991**, *40*, 8–15.

(13) Kobilka, B. K.; Frielle, T.; Collins, S.; Yang-Feng, T. L.; Kobilka, T. S.; Francke, U.; Lefkowitz, R. J.; Caron, M. G. An intronless gene encoding a potential member of the family of receptors coupled to guanine nucleotide regulatory proteins. *Nature* **1987**, *329*, 75–79.

(14) Julius, D.; McDermott, A. B.; Axel, R.; Jessell, T. M. Molecular characterization of a functional cDNA encoding the serotonin 1c receptor. *Science* **1988**, *241*, 558–564.

(15) Pritchett, D. B.; Bach, A. W. J.; Wozny, M.; Taleb, O.; Dal Toso, R.; Shih, J. C.; Seeburg, P. H. Structure and functional expression of cloned rat serotonin 5-HT$_2$ receptor. *EMBO J.* **1988**, *7*, 4135–4140.

(16) Voigt, M. M.; Kispert, J.; Chin, H. Sequence of a rat brain cDNA encoding an α-1B adrenergic receptor. *Nucleic Acid Res.* **1990**, *18*, 1053–1053.

(17) Zhou, Q. Y.; Grandy, D. K.; Thambi, L.; Kushner, J. A.; Van Tol, H. H.; Cone, R.; Pridnow, D.; Salon, J.; Bunzow, J. R.; Civelli, O. Cloning and expression of human and rat D$_1$ dopamine receptors. *Nature* **1990**, *347*, 76–83.

(18) Dal Toso, R.; Sommer, B.; Ewert, M.; Herb, A.; Pritchett, D. B.; Bach, A.; Shivers, B. D.; Seeburg, P. H. The dopamine D$_2$ receptor: two molecular forms generated as a target for neuropeptides. *EMBO J.* **1989**, *8*, 4025–4034.

(19) Sokoloff, P.; Giros, B.; Martres, M. P.; Bouthenet, M. L.; Schwarz, J. C. Molecular-cloning and characterization of a novel dopamine receptor as a target for neuropeptides. *Nature* **1990**, *347*, 146–151.

(20) Hubert, M. H.; Van Tol, H. H.; Bumzoco, J. R.; Guan, H. C.; Sunahara, R. K.; Seeman, P.; Niznik, H. B.; Civelli, O. Cloning of the gene for a human dopamine D4 receptor with high affinity for the antipsychotic clozapine. *Nature* **1991**, *350*, 610–614.

(21) Sumakara, R. K.; Guan, H. C.; O'Dowd, B. F.; Seeman, P.; Laurier, G. L.; Ngo, G.; George, S. R.; Torchia, J.; Hubert, H. M.; Van Tol, H. H.; Nizuik, H. B. Cloning of the gene for a human dopamine D$_5$ receptor with higher affinity for dopamine than D$_1$. *Nature* **1991**, *350*, 614–619.

(22) Libert, F.; Parmentier, M.; Lefort, A.; Dinsart, C.; Van Sande, J.; Maenhaut, C.; Simons, M. J.; Dumont, J. E.; Vassart, G. Selective amplification and cloning of four new members of the G-protein coupled receptor family. *Science* **1989**, *244*, 569–572.

(23) Kobilka, B. K.; Matsui, H.; Kobilka, T. S.; Yang-Feng, T. L.; Francke, U.; Caron, M. G.; Lefkowitz, R. J.; Regan, J. W. Cloning, sequencing and expression of the gene coding for the human platelet $α_2$-adrenergic receptor. *Science* **1987**, *238*, 650–656.

(24) Frielle, T.; Collins, S.; Daniel, K. W.; Caron, M. G.; Lefkowitz, R. J.; Kobilka, B. K. Cloning of the cDNA for the human $β$1-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7920–7924.

(25) Kobilka, B. K.; Dixon, R. A. F.; Frielle, T.; Dohlman, H. G.; Bolanowski, M. A.; Sigal, I. S.; Yang-Feng, T. L.; Francke, U.; Caron, M. G.; Lefkowitz, R. J. cDNA for the human $β$2-adrenergic receptor: a protein with multiple membrane spanning domains and encoded by a gene whose chromosomal location is shared with that of the receptor for platelet-derived growth factor. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 46–50.

(26) Emorine, L. J.; Marullo, S.; Briend-Sutren, M. M.; Patey, G.; Take, K.; Delavier-Klutchko, C.; Strosberg, A. D. Molecular characterization of the human $β$3-adrenergic receptor. *Science* **1989**, *245*, 1118–1121.

(27) Kurtenbach, E.; Curtis, C. A. M.; Pedder, E. K.; Aithen, A.; Harris, A. C. M.; Hulme, E. C. Muscarinic acetylcholine receptors. *J. Biol. Chem.* **1990**, *265*, 13702–13708.

(28) (a) Bonner, T. I.; Buckley, N. J.; Young, A. C.; Brann, M. R. Identification of a family of muscarinic acetylcholine receptor genes. *Science* **1987**, *237*, 527–532. (b) Bonner, T. I.; Young, A. C.; Brann, M. R.; Buckley, N. J. Cloning and expression of the human and rat m$_5$ muscarinic acetylcholine receptor. *Neuron* **1988**, *1*, 403–410.

(29) Akiba, I.; Kubo, T.; Maeda, A.; Bujo, H.; Nakai, J.; Mishina, M.; Numa, S. Primary structure of porcine muscarinic acetylcholine receptor III and antagonist binding studies. *FEBS Lett.* **1988**, *235*, 257–261.

(30) Gantz, I.; Schaffer, M.; Delvalle, J.; Logsdon, G.; Campbell, V.; Uhler, M.; Yamada, T. Molecular cloning of a gene encoding the histamine H$_2$ receptor. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 429–433.

(31) Nathans, J.; Hogness, D. S. Isolation and nucleotide sequence of the gene encoding human rhodopsin. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4851–4855.

cannabinol;[36] canine adenosine 1 and 2;[16] murine bombesin;[37] bovine endothelin 1 and 2;[38,39] canine Rdc1;[22] mouse thyrotropin;[40] human thromboxane A_2 (Tx2);[41] guinea pig platelet-activating factor binding receptor;[42] human LSH;[43] human TSH;[44] and the metabotropic glutamate[45] receptors. The sequences of three bacteriorhodopsin genes were also included in this study.[46] The alignment was performed with the method of Needleman and Wunsch using the Dayhoff similarity table for amino acids,[47,48] as implemented in the GCG software.[49] To obtain an optimal

alignment, the comparison was refined manually.

For the identification of the hydrophobic helical regions, the parameter sets of Kyte–Doolittle[50] and Goldman, Engelman, and Steitz[51] were used. For the validation of the prediction, bacteriorhodopsin was included in this study.

The refined model of bacteriorhodopsin was kindly provided by Richard Henderson.[11] Modeling was achieved with the molecular modeling package SYBYL 5.32.[52] The interactive modeling and display were performed on a Silicon Graphics 4D/280 computer. The receptors and their complexes were optimized in SYBYL using molecular mechanics calculations with the Kollman all atoms parameter set[53] in the following way. First, the single helices were minimized for 1000 steps using the conjugate gradient minimizer. The transmembrane part of the receptor models was constructed and again minimized for 2000 steps. The respective ligands were docked into the active site, and the whole complex was minimized for another 2000 steps. A cutoff of 8 Å was used. To account to some extent for the membrane environment, a dielectric constant of 5 was chosen. The geometries of serotonin, adrenaline, dopamine, and acetylcholine were taken from previous active analogue approach studies.[2,3,54]

The charge distributions of the neurotransmitter structures were obtained with the semiempirical molecular orbital package MOPAC using the MNDO approach.[55]

The seven helical stretches were generated using the BIOPOLYMER module of SYBYL, geometry optimized and fitted onto the corresponding backbone of bacteriorhodopsin to obtain the best possible interactions between the helices. To remove bad steric interactions, the orientation of the side chains was refined by hand and again geometry optimized. 3D models for all 39 GPCRs were generated in this way.

Five receptors (dopaminergic D_2, adrenergic α_2 and β_2, serotonergic 5-HT$_2$, and muscarinic m$_2$) were taken as prototypes and studied in detail. Additionally, the mammalian opsins were included in the study, since they represent a crucial link between bacteriorhodopsin and GPCR because they possess a number of common features. The mammalian opsins bind the same ligand (retinal) as bacteriorhodopsin. Retinal forms a Schiff base with a conserved Lys located on the same helix (TM7) as the Lys in bacteriorhodopsin. They become activated through light, and additionally they are able to pump protons (Chabre, personal communication). The natural ligands were manually docked into their putative binding sites. The complexes were optimized by molecular mechanics calcu-

(32) Nathans, J.; Thomas, D.; Hogness, D. S. Molecular genetics of human color vision the genes encoding blue, green and red pigments. *Science* 1986, *232*, 193–202.

(33) Masu, Y.; Nakayama, K.; Tamaki, H.; Haroda, Y.; Kuno, M.; Nakanishi, S. cDNA cloning of bovine substance k-receptor through oocyte expression system. *Nature* 1987, *329*, 836–838.

(34) Yokota, Y.; Sasai, Y.; Tanaka, K.; Shigemoto, R.; Kakizuka, A.; Ohkubo, H.; Nakanishi, S. Molecular characterization of a functional cDNA for rat substance P receptor. *J. Biol. Chem.* 1989, *264*, 17649–17652.

(35) Shigemoto, R.; Yokota, Y.; Tsuchida, K.; Nakanushi, S. Cloning and expression of a rat neuromedin k-receptor cDNA. *J. Biol. Chem.* 1990, *265*, 623–628.

(36) Matsuda, L. A.; Lolait, S. J.; Browstein, M. J.; Young, A. C.; Bonner, T. I. Structure of a cannabinoid receptor and functional expression of the cloned DNA. *Nature* 1990, *346*, 561–564.

(37) Battey, J. F.; Way, J. M.; Corjay, M. H.; Shapirce, H.; Kusano, K.; Harkino, R.; Wu, J. M.; Slattery, T.; Mann, E.; Feldman, R. I. Molecular cloning of the bombesin gastrin-releasing peptide receptor from Swiss 3T3 cells. *Proc. Natl. Acad. Sci. U.S.A.* 1991, *88*, 395–399.

(38) Arai, H.; Hori, S.; Aramori, I.; Ohkubo, H.; Nakunishi, S. Cloning and expression of a cDNA encoding an endothelin receptor. *Nature* 1990, *348*, 730–732.

(39) Sakurai, T.; Masashi, Y.; Takuwa, Y.; Miyazaki, H.; Kimura, S.; Goto, K.; Masaki, T. Cloning of a cDNA encoding a non isopeptide-selective subtype of the endothelin receptor. *Nature* 1990, *348*, 732–735.

(40) Straub, R. E.; Frech, G. C.; Joho, R. H.; Gerohengorn, M. C. Expression cloning of a cDNA encoding the mouse pituitary thyrotropin-releasing hormone receptor. *Proc. Natl. Acad. Sci. U.S.A.* 1990, *87*, 9514–9518.

(41) Hirata, M.; Hayashi, Y.; Ushikubi, F.; Yokoto, Y.; Kajeyama, R.; Nakanishi, S.; Narumiya, S. Cloning and expression of cDNA for a human thromboxane A_2 receptor. *Nature* 1991, *349*, 617–620.

(42) Honda, Z.; Nakamura, M.; Miki, I.; Ninami, M.; Watanabe, T.; Seyma, Y.; Okado, H.; Toh, H.; Ito, R.; Miyamoto, T.; Shimizu, T. Cloning by functional expression of platelet-activating factor receptor from guinea-pig lung. *Nature* 1991, *349*, 342–346.

(43) Fraser, A. L.; Robbins, L. S.; Stork, P. J.; Sprengel, R.; Segaloff, D. L. Isolation of TSH and LH/CG receptor cDNAS from human thyroid: regulation by tissue specific splicing. *Mol. Endocrinol.* 1990, *4*, 1264–1276.

(44) Sprengel, R.; Braun, T.; Nicolics, K.; Segaloff, D. L.; Seeburg, P. H. The testicular receptor for follicle stimulating hormone: structure and functional expression of cloned cDNA. *Mol. Endocrinol.* 1990, *4*, 525–530.

(45) Masu, M.; Tanabe, Y.; Tsuchida, K.; Shigemodo, R.; Nakanishi, S. Sequence and expression of a metabotropic glutamate receptor. *Nature* 1991, *349*, 760–765.

(46) Dunn, R. J.; McCoy, J.; Simsek, J. M.; Majumdar, A.; Chang, S. H.; Rajbhandary, U. L.; Khorana, H. G. The bacteriorhodopsin gene. *Proc. Natl. Acad. Sci. U.S.A.* 1981, *78*, 6744–6748.

(47) Needleman, S. B.; Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 1970, *48*, 443–453.

(48) Dayhoff, H. O.; Schwarz, R. M.; Orcutt, B. C. In *Atlas of Protein Sequence and Structure*; Dayhoff, M. O., Ed.; NBFR: Washington, DC, 1978; 5, Suppl. 3.

(49) Devereux, J.; Haeberli, P.; Smithies, O. A comprehensive set of sequence analysis programs for the Vax. *Nucleic Acid Res.* 1984, *12*, 387–395, as implemented in the GCG software at University of Wisconsin, Biotechnology Center, 1710 University Ave., Madison, Wisconsin 53705.

(50) Kyte, J.; Doolittle, R. F. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 1982, *157*, 105–132.

(51) Hulme, E. C.; Birdsall, N. J. M.; Buckley, N. J. Muscarinic receptor subtypes. *Annu. Rev. Pharmacol. Toxicol.* 1990, *30*, 633–673.

(52) Tripos E & S, 6548 Clayton Road, St. Louis, MO 62117. SYBYL molecular modeling software.

(53) Weiner, S. J.; Kollman, P. A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* 1986, *7*, 230–252.

(54) Takemura, S. Geometry of muscarinic agonists. *J. Pharmacobio-Dyn.* 1984, *7*, 436–444.

(55) Dewar, M. J. S.; Thiel, W. Ground states of molecules 38. The MNDO method, approximations and parameters. *J. Am. Chem. Soc.* 1977, *99*, 4899.
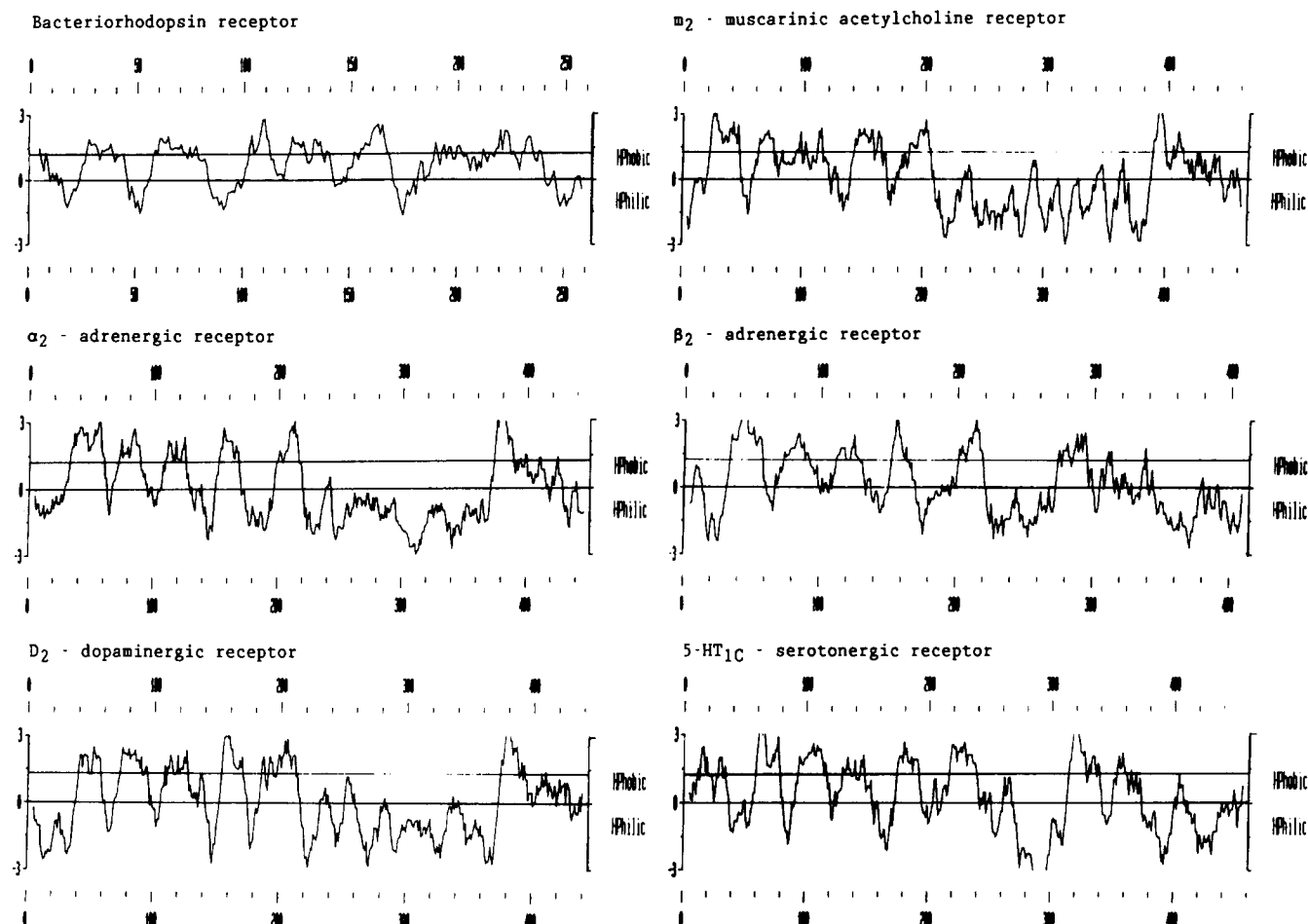
Figure 1. Hydropathicity profiles of five representative GPCRs in comparison to bacteriorhodopsin (bacteriorhodopsin, m_2 muscarinic acetylcholine receptor, α_2- and β_2-adrenergic receptor, D_2-dopaminergic and $5HT_{1c}$-serotonergic receptor). The profiles are obtained using a window region of 7 and the Kyte–Doolittle parameters. The similarity in the hydropathicity profiles indicates seven transmembrane regions for the G-protein-coupled receptors as well as for bacteriorhodopsin.

lations. The docking procedure was repeated several times with different initial orientations of the side chains and of the ligand in order to obtain the best possible interaction complexes. The interaction energies obtained cannot be used to calculate exact affinities between ligand–receptor since changes in entropy and solvation were not taken into account at this stage of the study. Nevertheless, energies obtained for different conformations of the same complex could be compared to each other in terms of more or less favorable states.

Results and Discussion

i. Sequence Similarity and Hydropathicity Analysis. Primary structure alignment clearly defines seven highly conserved hydrophobic sequences corresponding to transmembrane regions connected by hydrophilic sequences with only very low conservation. The hydropathicity analysis of the 42 primary sequences also indicates seven transmembrane regions (Figure 1). Nevertheless, it is not possible to localize precisely the starting and ending amino acids of the transmembrane regions from these analyses. Furthermore, the helices cannot always be well-defined from the hydropathicity profiles. In particular, transmembrane regions 3 and 7 contain a high number of polar residues which induce a marked fluctuation in the hydropathicity plot. The seventh transmembrane region which is not clearly defined from the hydropathicity analysis displays the highest amino acid conservation among the GPCRs and thus can be considered to correspond to a membrane-embedded region.

The application of biophysical techniques such as circular dichroism to transmembrane receptors and theoretical analysis of protein structure have led to the proposal that the transmembrane regions are largely in an α-helical conformation, with the α-helices spanning the membrane. The loops connecting these transmembrane helices are exposed to a polar environment and are normally located at the surface of the phospholipid bilayer. The three-dimensional structures of the photosynthetic reaction center, human annexin V, and bacteriorhodopsin, together with topographic data on the GPCRs, support these generalizations.[56,11,57]

The drawing of the sequences of the putative transmembrane regions of the GPCRs on helical wheels shows that they are all seven strongly amphiphilic with hydrogen bonding residues (such as Tyr, Ser, Thr, Asn, Cys, Asp) concentrated on one face. Additionally, all conserved residues and most aromatic residues are also located on this face. These arrangements correspond to a helical conformation for the GPCRs in which the conserved, charged, and hydrogen bonding amino acids could point to the inside of the membrane receptor. The face directed

(56) Deisenberg, J. O.; Miki, K.; Huber, R.; Michel, H. The structure of the photosynthetic reaction center. *Nature* **1985**, *881*, 618–624.

(57) Huber, R.; Rômisch, J.; Paques, E. P. The crystal and molecular structure of human annexin V, an anticoagulant protein, that binds to calcium and membranes. *EMBO J.* **1990**, *9*, 3867–3874.

TM1   TM2   TM3   TM4   TM5   TM6   TM7
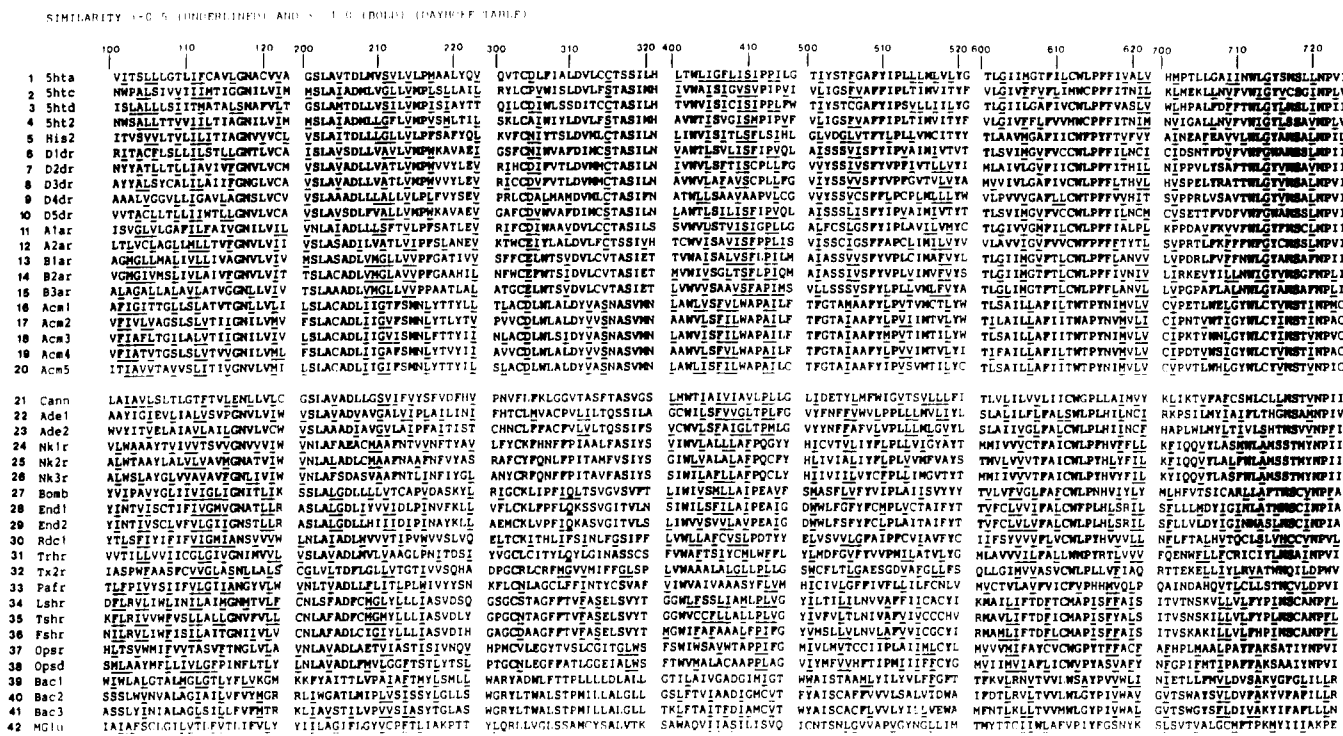
Figure 2. Alignment of the seven selected regions putatively included in transmembrane domains (labeled TM1 to TM7). The considered receptors are as follows: 1/ 5-HT$_{1A}$; 2/ 5-HT$_{1C}$; 3/ 5-HT$_{1D}$; 4/ 5-HT$_2$; 5/ His$_2$; 6/ D$_1$; 7/ D$_2$; 8/ D$_3$; 9/ D$_4$; 10/ D$_5$; 11/ α_1; 12/ α_2; 13/ β_1; 14/ β_2; 15/ β_3; 16/ Acm$_1$; 17/ Acm$_2$; 18/ Acm$_3$; 19/ Acm$_4$; 20/ Acm$_5$; 21/ cannabinol; 22/ adenosin 1; 23/ adenosin 2; 24/ NK1r; 25/ NK2r; 26/ NK3r; 27/ bombesin; 28/ endothelin A (1); 29/ endothelin B (2); 30/ Rdc 1; 31/ TRH; 32/ thromboxane A$_2$ Tx2; 33/ PAF; 34/LSH; 35/ TSH; 36/FSH; 37/ opsin r; 38/ opsin d; 39/ bacteriorhodopsin from *Halobacterium halobium* (Bac 1); 40/ halorhodopsin from *Halobacterium* SP (Bac 2); 41/ halorhodopsin from *Natronobacterium pharaonis* (Bac 3); 42/ glutamate receptor. For convenience, the residues are numbered 101-102-etc. in helix TM1, 201-202-etc. in helix TM2, etc. Invariant residues in all GPCR or in subclasses are indicated in bold. The alignment of bacteriorhodopsin results from a posteriori comparison of three-dimensional structures.

toward the lipid bilayer consists mainly of nonaromatic hydrophobic residues.

Figure 1 shows the hydropathicity plots of representative GPCRs in comparison to bacteriorhodopsin. The hydropathic patterns of GPCRs and bacteriorhodopsin are sufficiently similar to assume that one can reasonably extrapolate from a 2D to a 3D structure in a similar manner for all of them.

ii. Comparison of Sequences. The nature and putative function of some conserved residues have already been extensively analyzed and discussed, but usually within a given receptor class.[58] We have extended this study to the 39 GPCRs mentioned above including the recently published GPCRs (Figure 2).

Global Homology in the Transmembrane Regions of the Cationic Neurotransmitter Receptors. A detailed comparison of the transmembrane regions shows that the highest percentage of homology is found within the muscarinic receptor subtypes with 71-86% identity (Figure 3). The 5-HT$_2$ and 5-HT$_{1C}$ receptors also display a 76% identity. The β-adrenergic receptors show a similar high percentage identity to each other (70%). However, the homology within the α-adrenergic receptors is rather low (43%). Interestingly, but not surprisingly, the 5-HT$_{1A}$ and 5-HT$_{1D}$ receptors possess a very low percentage identity with the 5-HT$_2$ or 5-HT$_{1C}$ receptor (35%), but share a higher identity with the dopaminergic and adrenergic receptors ($\sim$45%). It is also clear that the neurokinin receptors, the mammalian opsins, the cannabinol

receptor, and the other GPCRs mentioned above represent distinct subclasses of GPCRs.

These data correlate very well with structure–activity relationships. Thus, medicinal chemists have failed until now to design highly selective muscarinic receptor ligands. This reflects the extremely high homology of the transmembrane regions among the muscarinic receptor subtypes. Similarly, it appears difficult to design compounds which can differentiate between the 5-HT$_{1C}$ and the 5-HT$_2$ receptors. In contrast, many compounds can discriminate 5-HT$_{1A}$ receptors from the other 5-HT receptor subtypes, but they usually also bind with a significant potency to adrenergic and dopaminergic receptors. Thus, for instance, closely related aminotetralin derivatives are very potent D$_2$ or 5-HT$_{1A}$ ligands while compounds such as MDL 72832 or WB4101 are potent at both 5-HT$_{1A}$ and α_1 receptors.[59] More generally, the sequence similarities observed between dopamine, adrenaline, and serotonin receptors account for the existence of many potent but nonselective compounds for any pair of the catechol and indolamine receptors.

Conserved Residues. Despite substantial differences in the activation processes of the GPCR, we found residues conserved within almost all of them. In the mammalian opsins, retinal is covalently bound to a Lys residue in transmembrane helix 7 forming a Schiff base whereas in the cationic neurotransmitter receptors the ligands are involved only in weaker electrostatic interactions, presumably with an aspartate on transmembrane region 3. Similarly, the cannabinol and peptide receptors can couple

(58) Goldman, A.; Engelman, D. M.; Steitz, T. A. Identifying non polar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Chem.* 1986, *15*, 321–353.

(59) Van Wijngaarden, I.; Tulp, M. T. M.; Soudijn, W. The concept of selectivity in 5-HT receptor research. *Eur. J. Pharmacol.* 1990, *188*, 301–312.

```
   | 1  2  3  4  5  6  7  8  9  10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42
---------------------------------------------------------------------------------------------------------------------------------
 1 |// 35 53 35 42 43 44 43 41 46 48 41 45 41 43 36 36 35 35 35 31 32 36 23 24 27 26 27 23 25 28 16 19 20 19 21 26 24 14  9 11  8
 2 |63 // 39 76 39 35 42 39 36 38 39 39 41 41 39 30 29 33 29 33 21 30 25 21 21 23 22 20 21 22 25 13 16 19 15 20 15 20 12 11 11 14
 3 |75 64 // 39 39 42 45 43 36 41 46 42 43 38 38 33 32 35 33 35 33 25 27 22 21 22 25 23 23 23 25 17 21 17 17 18 21 21 11 11 13 15
 4 |61 92 65 // 39 37 42 37 35 38 40 41 41 37 38 29 29 34 29 35 21 28 26 21 20 23 23 18 21 20 26 12 17 20 16 19 17 19 14 13 14  9
 5 |69 71 66 71 // 41 43 39 40 42 47 39 42 39 43 33 33 34 33 39 24 26 27 25 27 25 27 25 25 23 27 16 23 18 21 20 19 19 11 12 11  9
 6 |68 67 63 67 69 // 45 41 37 83 43 43 44 41 39 29 31 31 31 30 27 29 28 23 22 25 25 19 23 21 21 14 19 20 21 17 21 20 14 11 14 11
 7 |67 73 65 72 69 65 // 76 53 47 49 49 43 41 43 32 32 33 34 33 29 35 36 28 27 29 23 25 25 19 26 11 21 19 17 22 20 25 10 11 15  7
 8 |65 67 65 66 63 63 85 // 55 45 46 45 43 39 43 31 31 29 30 30 25 31 33 27 29 27 22 24 23 29 26 17 26 16 17 20 25 27  8 14 17  9
 9 |60 66 59 65 63 60 73 74 // 41 41 45 41 34 43 35 32 30 32 31 25 31 31 25 25 23 24 24 23 24 24 18 19 18 15 21 22 27 11 11 11 10
10 |69 69 63 69 71 93 67 64 61 // 43 47 47 42 42 33 31 33 32 33 28 27 29 25 23 25 27 19 21 21 23 15 17 23 20 20 23 21 12 12 14 12
11 |72 65 69 64 71 67 68 66 61 70 // 43 48 42 47 30 31 29 33 32 28 29 34 28 31 29 30 24 26 27 26 15 21 20 19 21 22 21  9 14 13 11
12 |69 67 67 67 65 67 74 67 66 70 73 // 47 39 43 30 29 29 28 29 27 31 32 24 23 25 21 18 23 21 23 14 21 21 21 18 25 23 11  7 11  7
13 |65 71 67 68 71 67 69 64 63 70 75 71 // 70 70 33 29 33 30 33 27 34 34 26 25 26 22 25 21 28 17 18 23 19 20 23 22  9 11 13 10
14 |69 65 72 63 70 69 67 64 61 69 72 70 85 // 62 33 29 34 33 31 22 29 31 26 28 29 29 21 24 19 24 15 18 23 23 24 22 21 11 11 12  9
15 |67 67 66 65 67 64 68 65 64 65 70 66 81 77 // 37 29 32 30 31 25 33 32 28 33 27 27 21 26 21 25 17 16 20 18 24 20 23 11 13 13  9
16 |62 63 61 62 61 58 57 57 57 59 61 62 63 59 57 // 71 76 75 80 28 26 25 19 24 23 23 24 26 21 23 14 15 18 20 22 19 18 12 11 10 12
17 |64 61 63 60 61 57 57 56 59 59 61 64 59 57 53 86 // 77 86 71 27 24 24 21 23 25 25 25 23 21 27 15 19 17 19 21 17 20  9  9 11 14
18 |63 64 64 63 62 60 58 55 55 61 62 64 63 63 57 92 88 // 75 79 31 25 27 20 21 22 25 24 23 23 27 13 18 18 19 21 19 19  9 10 11 14
19 |63 61 61 62 59 57 59 55 57 59 60 63 59 57 54 87 95 89 // 74 27 26 24 23 25 25 23 26 25 21 27 14 16 16 20 21 18 18 10 10 10 13
20 |64 61 62 62 58 57 56 57 61 64 65 63 60 57 91 88 93 89 // 29 24 27 22 22 22 23 25 24 25 24 13 17 17 19 21 18 18 10 13 10 14
21 |56 55 55 54 51 49 50 48 48 51 54 51 53 53 50 55 55 56 // 22 24 21 22 21 19 21 22 18 17 17 16 23 21  8  7 10 16
22 |58 59 51 59 55 53 60 55 57 54 59 63 59 54 57 49 51 51 52 49 53 // 60 22 24 25 23 25 25 20 27 17 17 17 19 20 22 19  7 12 14  8
23 |57 54 52 53 53 51 61 53 55 53 60 60 55 51 55 47 50 50 49 51 54 79 // 25 25 25 28 25 27 23 31 15 25 22 22 23 21 21  9 15 16  8
24 |48 50 49 50 49 47 51 50 43 49 51 48 51 52 51 49 46 49 49 49 47 50 48 // 65 67 27 23 24 29 21 15 25 19 21 20 24 26 10  6  9 10
25 |53 51 51 49 49 45 49 52 45 46 53 49 48 49 51 44 46 45 47 45 43 51 51 79 // 67 27 27 26 27 21 15 21 21 21 23 23 24 11 12 13  7
26 |50 50 47 51 51 46 49 49 45 48 52 49 51 49 50 49 49 48 49 47 43 51 47 85 84 // 27 27 27 25 21 17 21 21 22 21 24 23  8  7 11  7
27 |57 52 52 51 56 55 53 49 45 56 57 49 57 58 55 51 52 54 51 54 47 51 52 58 57 56 // 39 45 29 26 19 24 21 21 22 19 23  9  9  9 11
28 |58 52 53 51 50 49 52 50 46 49 55 49 50 49 47 48 53 51 51 48 53 53 52 55 54 66 // 73 27 21 18 19 19 17 22 21 21 10  7  8 12
29 |55 51 51 51 50 51 51 49 48 51 56 47 50 49 49 47 51 51 50 51 50 53 51 53 53 55 69 91 // 25 20 18 24 19 20 20 21 25 13  9 11  9
30 |53 55 51 53 53 50 55 56 47 51 58 50 51 52 47 50 51 55 51 54 47 49 49 55 55 57 59 55 57 // 20 18 29 15 16 19 21 23 11 11 11 10
31 |50 53 47 50 53 44 52 46 45 48 52 47 56 53 52 49 50 53 48 53 50 53 52 48 46 49 52 49 49 51 // 15 21 21 18 22 21 24  8 12 13 13
32 |47 48 47 47 46 41 43 42 38 45 47 45 49 44 46 37 43 42 41 40 47 45 45 42 43 42 45 45 43 47 47 // 17 15 18 17 16 19 11  9  7 13
33 |48 49 46 47 48 47 49 50 44 46 52 47 45 46 41 45 44 48 45 45 43 45 53 49 50 45 51 45 46 58 45 39 // 20 25 19 18 25 11 13 10  9
34 |49 49 41 51 47 45 45 39 40 45 52 46 49 50 47 45 46 45 47 43 50 51 46 48 49 50 43 45 43 48 38 45 // 71 71 21 23  6  9  6  8
35 |49 49 44 51 47 45 45 40 43 45 52 48 51 52 49 47 48 47 48 49 44 49 47 50 49 49 43 44 44 46 40 47 89 // 67 23 22  7 10  9  7
36 |51 52 45 51 49 45 47 41 44 47 52 47 53 52 52 47 45 45 45 47 45 51 52 47 50 49 50 41 43 48 50 44 47 85 89 // 23 21  7  9  5  7
37 |53 49 46 47 45 49 49 48 49 49 49 49 46 48 45 43 45 45 41 41 45 52 48 48 53 52 45 42 45 43 47 36 40 44 47 47 // 39  9  7  9 10
38 |51 47 47 48 47 45 51 50 49 45 48 45 42 45 47 46 47 49 45 47 46 48 48 51 53 52 51 46 49 49 49 44 47 47 46 51 61 // 13 10  9 11
39 |35 40 31 43 35 33 41 35 39 35 36 39 35 35 33 33 37 35 37 35 32 35 39 35 37 32 31 33 32 35 35 31 30 27 27 28 32 38 // 38 35  7
40 |35 38 33 39 31 31 36 35 36 31 40 37 37 35 35 31 34 34 33 33 35 37 40 28 35 31 32 35 33 36 35 31 35 28 30 30 28 31 65 // 69 11
41 |37 39 37 39 35 37 41 37 39 38 41 42 39 37 37 34 37 37 36 36 38 40 39 31 39 35 36 37 37 38 37 34 31 27 29 28 32 34 59 86 // 11
42 |29 37 30 33 27 34 31 33 29 37 36 30 35 33 28 29 35 32 33 33 37 31 32 27 26 27 31 33 33 31 30 33 29 29 29 27 28 29 30 37 35 //
```

```
( 1) 5ht1a#Hum;( 2) 5ht1c#Rat;( 3) 5ht1d#Can;( 4) 5ht2##Rat;( 5) His2#Cani;( 6) D1dr#Huma;( 7) D2dr#Huma;( 8) D3dr#Ratt;
( 9) D4dr#Huma;(10) D5dr#Huma;(11) Alar##Huma;(12) A2ar#Huma;(13) B1ar#Huma;(14) B2ar#Huma;(15) B3ar#Huma;(16) Acm1#Huma;
(17) Acm2#Huma;(18) Acm3#Pig;(19) Acm4#Huma;(20) Acm5#Huma;(21) Cannabino;(22) Aden1#Can;(23) Aden2#Can;(24) Nk1r##Rat;
(25) Nk2r#bovi;(26) Nk3r##Rat;(27) Bombesinr;(28) Endo1rece;(29) Endo2rece;(30) Rdc1#Cani;(31) Trh#Mouse;(32) Txr#Human;
(33) Paf#Guine;(34) Lshr#Huma;(35) Tshr#Huma;(36) Fshr#Huma;(37) Opsr#Huma;(38) Opsd#Huma;(39) Bac#Halha;(40) Bachh#alo;
(41) Bachn#att;(42) Mglutarec;
```

Figure 3. Percentage identity and similarity between the transmembrane regions of different G-protein-coupled receptors. Identity values are given above the diagonal.

to a G-protein but their agonists do not possess a cationic group. Given these different binding modes, our basic assumption was that amino acids invariant within all GPCRs play an important role in the overall folding and function of the receptors whereas amino acids conserved only in specific classes of receptors should be responsible for the binding of the corresponding ligands and for some of their specific triggering mechanisms.

Conserved Pro and Gly Residues. Among the amino acids conserved throughout the GPCRs, it is striking to find highly invariant Pro and Gly residues in transmembrane regions (TM) 1, 4, 5, 6, and 7. As already postulated, such amino acids could be involved in conformational changes during ligand binding. Furthermore, TM1 and TM7 possess homologous Gly_{116}-Asn_{117} (or Thr_{116}-Asn_{117} in mammalian opsins) and Asn_{721}-Pro_{722} sequences. (Figure 2: the numbering of the sequences is taken from our alignment and does not correspond to the numbering of the individual receptor sequences. The first digit corresponds to the helical transmembrane domain, the next two digits indicate the position of the residue in the helix). Gly and Pro residues occur rarely in α-helices since they

induce bends and kinks in the helical backbone, although some are found.[60] Interestingly, the presence of a Pro residue in a transmembrane α-helix has been studied experimentally and has been shown to have a hinge function, inducing oscillations of the two helical arms.[61] Asn residues could stabilize the α-helical conformation around Pro residues by involving its side chain in a hydrogen bond to the peptide backbone. Therefore an Asn next to a Pro or a Gly might play a role in the proper folding of the protein as well as in stabilizing different conformations of the receptor.

Conserved Aspartic Acid Residues. All cationic neurotransmitter receptors possess three invariant acidic

(60) Finkelstein, A. V.; Ptitsyn, O. B. Statistical analyses of the correlation among amino acid residues in helical, β-structured and non-regular regions of globular proteins. *J. Mol. Biol.* 1971, 62, 613–624.

(61) Riegler, R. "Molecular recognition, allosteric receptors and drug design". Molecular structure, dynamics and interactions in bioactive membrane peptides. *Round Table Roussel-Uclaf* 1990, 67, 19–20.

residues. Two of them, Asp_{207} and Asp_{328}, are highly conserved across the whole GPCR family, indicating that these two residues probably play an essential role in the folding and/or in the function of the receptor.

Asp_{207} is located in TM2 in a conserved sequence incorporating the C-terminal amino acids of the first cytoplasmic loop. Mutation of this residue to Asn in the muscarinic m_1 receptor[62] produced mutant receptors that had normal antagonist but lower agonist affinity. These data are similar to those obtained from site-directed mutagenesis of the analogous Asp_{207} in the β_2-adrenergic receptor.[63] From these results, the authors suggested that Asp_{207} is involved in the binding of the natural agonists. As stated above and in agreement with Hulme and coworkers, the fact that this amino acid is conserved in the cationic neurotransmitter receptors as well as in the other GPCRs makes it more plausible that Asp_{207} rather plays a specific role in the function of the receptor. Our hypothesis received strong support from recently published mutagenesis experiments of the corresponding aspartates on the α_{2a} and D_2 receptors. In the α_{2a}-adrenergic receptor, Asp_{207} is required for allosteric regulation of the receptors by Na^+.[64] In the D_2 receptor, similar results are obtained, abolishing or decreasing the regulation of the affinity of D_2 receptors for agonists and substituted benzamine antagonists by Na^+ and pH.[65] These results support and extend our hypothesis by indicating that interaction of cations or protons with the aspartate residue modulates receptor conformation. Thus, it seems likely that Asp_{207} insures a crucial allosteric role for the G-protein-coupled function of the receptor.

The second fully conserved acidic residue, Asp_{328} (or Glu), is located at the end of TM3 near the intracellular domain in a conserved Asp-Arg-Tyr (DRY) sequence. Mutation of this Asp residue to Asn in the muscarinic m_1 receptor produced a mutant receptor that had normal affinity for antagonists, but a 3.2-fold higher affinity for carbachol. However, the EC_{50} for agonist stimulation of the PI turnover was decreased approximately by a factor of 10. The discrepancy between the increase in receptor affinity and the decreased efficacy of full and partial agonists to ellicit maximal responses suggest a change in the efficiency of the receptor coupling to the G-protein.[62]

In contrast, Asp_{311}, which is located at the middle of TM3 near the extracellular domain, is only present in the cationic neurotransmitter receptors and absent in all other GPCRs. Two experimental findings indicate that Asp_{311} participates in the binding of the ligand's ammonium headgroup. Firstly, peptide mapping and sequencing studies pinpoint Asp_{311} as the major site at alkylation of purified forebrain muscarinic acetylcholine receptors by [^{3}H]propylbenzylcholine mustard ([^{3}H]PrBCM), a benzylcholine analogue in which the quarternary ammonium

headgroup is replaced by a chemically reactive aziridinium moiety.[66] Secondly, mutation of Asp_{311} to Asn produced a muscarinic receptor with drastically decreased affinity for agonists and antagonists.[62] Similar results were obtained in the hamster β-adrenergic receptor.[63]

On the basis of the results of site-directed mutagenesis of the conserved aspartates, it has been suggested that there are two different ligand binding sites involving Asp_{311} and Asp_{207} as counterions for the positively charged nitrogen of the ligands.[4] Our primary structure comparison together with recently published mutagenesis experiments suggests rather that there is essentially one binding site in which Asp_{311} on TM3 binds the cationic headgroup of the neurotransmitter, while Asp_{207} on TM2 and Asp_{328} on TM3 have a structural and allosteric functional role. This important problem will be addressed in the 3D models.

Conserved Aromatic Residues. Most GPCRs contain invariant aromatic residues on transmembrane regions 4, 5, 6, and 7. It is generally known that the aromatic residues Trp or Tyr can be involved in important internal cross-linking hydrogen bonds and conformational changes.[67] Thus we would propose that such residues can play a similar role in GPCR function and, in particular, mediate a transition between different conformations.

The cationic neurotransmitter GPCRs additionally possess a number of conserved aromatic residues which are absent in the mammalian opsins, the peptide and cannabinol receptors. In particular, a conserved Trp, Phe, or Tyr_{307} residue is located at the beginning of TM3, one helix turn before Asp_{311}. TM5 shows a conserved Phe_{509}-Tyr_{510} sequence. TM6 has a conserved Phe_{615} or Tyr_{615} in the direct neighborhood of the two aromatic residues Phe_{609} and Trp_{613} which are conserved in most GPCRs. TM7 possesses a Tyr_{715} residue which is particularly interesting since its position corresponds to that of the retinal attachment lysine residue in the opsin family. In the mammalian opsins, a Glu_{307} residue occupies a position very similar to that of the TM3 Asp_{311} in the cationic neurotransmitter receptors and is postulated to act as a counterion to the protonated retinal Schiff base, suggesting a close interaction/association between TM3 and TM7. In agreement with Hulme et al., we would argue that Tyr_{715} could play a homologous role in the cationic amine receptors in bridging TM3 and TM7. This residue may be crucial for ligand binding and receptor activation. This will be studied in more detail in the 3D models.

Conserved Residues in Specific Cationic Neurotransmitter Receptor Types. In addition to fully conserved amino acids and those conserved only within the cationic neurotransmitter receptors, amino acids are found which are conserved only within the catecholamine, indolamine, or muscarine receptor subtypes (Figure 2).

For example, TM3 contains a conserved Val_{312}-Leu_{313} sequence in the aromatic neurotransmitter receptors which is exchanged for a Tyr-Val sequence in the muscarinic receptors. Additionally, Thr_{316} is exchanged to Asn in the muscarinic receptors. Interestingly, the 5-HT and α_2 receptors contain a conserved His_{321} on TM3. Instead of this, the muscarinic receptor subtypes show a conserved Asn_{321}. In TM4, Ser_{410} is conserved in the catechol and indolamine receptors but exchanged for a Trp in the muscarinic receptors. Ser residues on helix 5 (Ser_{505} and Ser_{508}) are

(62) Fraser, C. M.; Wang, Cheng-Dian; Robinson, D. A.; Gocayne, J. D.; Venter, J. C. Site directed mutagenesis of m_1 muscarinic acetylcholine receptors: conserved aspartic acids play important roles in receptor function. *Mol. Pharmacol.* **1990**, *36*, 840–847.

(63) Dixon, R. A. F.; Strader, C. D.; Sigal, I. S. Structure and function of G-protein coupled receptors. *Annu. Rep. Med. Chem.* **1988**, *23*, 221–233.

(64) Horstman, D. A.; Brandon, S.; Wilson, A. L.; Guyer, C. A.; Cragoe, E. J.; Limbird, L. E. An aspartate conserved among G-protein receptor confers allosteric regulation of α_2-adrenergic receptors by sodium. *J. Biol. Chem.* **1990**, *265* (35), 21590–21595.

(65) Neve, K. A.; Cox, B. A.; Henningsen, A. R.; Spanoyannis, A.; Neve, R. L. Pivotal role for aspartate-80 in the regulation of dopamine D_2 receptor affinity for drugs and inhibition of adenylyl cyclase. *Mol. Pharmacol.* **1991**, *39*, 733–739.

(66) Curtis, C. A.; Wheathy, N.; Bansal, S.; Birdsall, N. J. M.; Eveleigh, P. Transmembrane helix 3 of the muscarinic receptor. *J. Biol. Chem.* **1989**, *264*, 489–495.

(67) Burley, S. K.; Petsko, G. A. Aromatic-aromatic interactions. A mechanism of structure stabilization. *Science* **1985**, *229*, 23–28.

conserved as a pair among catecholamine receptors but not in the neurotransmitter receptors with endogenous ligands which do not have the catechol moiety. Experimental evidence suggesting that Ser_{505} and Ser_{508} are involved in hydrogen bonds to the catechol moiety of the ligand came again from site-directed mutagenesis. Mutants of the β_2 receptor in which Ser_{505} and Ser_{508} are substituted by Ala were found to bind isoproterenol with a 10–100-fold decreased affinity compared to the wild type receptor.[68] It is interesting to note that the serotonin receptor subtypes, which bind a substrate with only one hydroxyl group on the aromatic ring, possess only Ser_{505} or Thr_{505} while Ser_{508} is exchanged for an Ala residue. The dopaminergic and the α- and β-adrenergic receptors, which bind natural ligands with two hydroxyl groups, possess both Ser_{505} and Ser_{508}. Instead of these Ser residues, two hydrophobic Ala residues are present in the muscarinic receptors, for which the ligand, acetylcholine, contains no phenol to interact with.

In the catecholamine and indolamine receptor subtypes, TM6 possesses a conserved Pro_{615}-Phe_{616}-Phe_{617} sequence. Site-directed mutagenesis showed that substitution of Phe_{616} and Phe_{617} by an Ala residue in the β-adrenergic receptor leads to a 10-fold decrease in the affinity of the receptor for isoproterenol. This led to the assumption that the two Phe residues are involved in hydrophobic interactions with the catecholamine ligands.[63] In contrast, the muscarinic receptor subtypes show a conserved Pro_{615}-Tyr_{616}-Asn_{617} sequence. The high conservation of Asn within the muscarinic subtypes suggests that this residue also plays a crucial role in the binding of acetylcholine, but this role remains to be understood.

While numerous labeling or mutagenesis studies have allowed speculation on the functional role of the Asp and of some Ser residues in the transmembrane domains, it remains more difficult to propose a precise role for most of the other conserved residues listed above. In particular, the function of the aromatic residues could not be evaluated or predicted, with the exception of Tyr_{715} on helix 7 and Phe_{616} and Phe_{617} on helix 6.[63,58]

Clearly, very interesting information has been obtained from mutagenesis, labeling data, and extensive comparison of primary sequences. However, 3D models of GPCRs are necessary to provide new insights into the processes involved in receptor function and activation.

3D Models of the Cationic Neurotransmitter–Receptor Complexes

i. 3D Models of the Cationic Neurotransmitter GPCRs. 3D models were constructed for the transmembrane portion of the GPCRs as derived from our primary structure comparison and hydropathicity analyses. The high degree of similarity within these hydrophobic stretches leads to the assumption that the homologous transmembrane regions in all GPCR have the same secondary structure and fold in the same way.

The seven helices were constructed with φ/ψ angles of $-59°$ and $-44°$, respectively, as suggested for an α-helix in a hydrophobic environment[69] and geometry optimized. The distribution of the conserved and charged amino acids on the same face of the α-helices implies that the as-
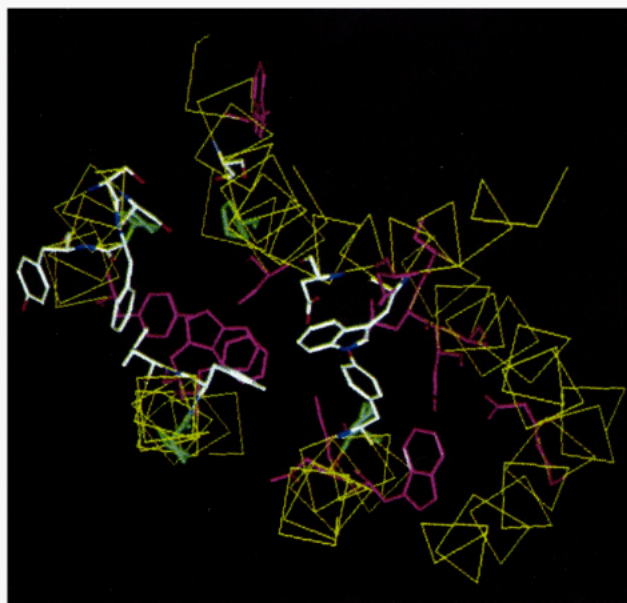
Figure 4. View of the seven transmembrane regions of the β_2-adrenergic receptor from the exoplasmic site of the cell. The seven helices are indicated by a c_α-carbon chain trace. Only the conserved residues are displayed.

sumptions made above are correct. As mentioned above, it is not possible to localize precisely the starting and ending of the transmembrane helices from the hydropathicity analysis. For the positioning of TM3, TM6, and TM7, the low homology found between the mammalian opsins and bacteriorhodopsin on TM3, TM6, and TM7 (see alignment Figure 2) was used. TM1, TM2, TM4, and TM5 were positioned so that their apolar portions begin and end near the membrane surfaces according to the results from primary structure comparison. Helices were oriented in such a way that all charged amino acids as well as the conserved residues point towards the inside of the receptor following the general rule for membrane protein folding.

The model derived from the cryoelectron microscopy study of bacteriorhodopsin was used as a template for the relative positioning of the α-helical main axes. In the bacteriorhodopsin structure, the 7 transmembrane helices are very tightly packed with the amino acid side chains of adjacent helices closely stacked. Helices 2, 3, and 4 are slightly tilted relative to the others. The overall dimensions of the model are $25 \times 35 \times 40$ Å with the longest dimension perpendicular to the plane of the membrane. Retinal is bound in its trans conformation and forms a Schiff base with Lys on TM7. The ionone ring of retinal is located in the vicinity of the extracellular surface in a large hydrophobic pocket defined by TM4, TM5, and TM6.[11]

For the modeling of the GPCRs, the relative position of helices 1, 2, 3, 5, 6, and 7 was directly taken from the 3D model of bacteriorhodopsin. As mentioned by Henderson, the interpretation of helix D (TM4) is less certain, and we found the interations with helix 5 and 3 not really favorable. Thus the position of helix 4 was changed slightly to improve the helix–helix interactions.

In this report we will focus on the 3D models of the cationic neurotransmitter receptors. Models for the other classes of GPCRs have been constructed in a similar way and will be discussed elsewhere.

Using the arrangement of the α-helices found in bacteriorhodopsin, it was possible to construct models possessing a number of features which we believe to be es-

(68) Strader, C. D.; Candelore, M. R.; Hill, W. S.; Sigal, I. S.; Dixon, R. A. F. Identification of two serine residues involved in agonist activation of the β-adrenergic receptor. *J. Biol. Chem.* **1989**, *264*, 13572–13578.

(69) Blundell, T.; Barlow, D.; Borlakoti, N.; Thornton, J. Solvent induced distortions and the curvature of α-helices. *Nature* **1983**, *306*, 281–283.

sential for this class of membrane-embedded receptors: the seven α-helices are tightly packed and define a central narrow, dihedral cleft. Charged amino acids and those conserved in all GPCRs are located on the inside of the receptors (Figure 4). Additionally, most of the aromatic residues are also located on the inside. Interestingly, they are not distributed randomly, but are rather concentrated on helices 4, 5, 6, and 7, indicating that this region plays an important role in the formation of hydrophobic aromatic interactions. In contrast, TM1 and TM2 possess no conserved aromatic residue, and in TM3, only Trp_{307} (or Phe or Tyr) is present. Thus, the core of cationic neurotransmitter receptors contains negatively charged Asp residues on helices 2 and 3 and a concentration of aromatic residues on helices 4, 5, 6, and 7. Furthermore, amino acids conserved only within the cationic neurotransmitter receptors are essentially located in the central cleft near the extracellular surface and in the immediate vicinity of Asp_{311}. As mentioned above, primary sequence analysis and recently published site-directed mutagenesis led us to propose that Asp_{209} on TM2 is not directly involved in ligand binding. This received strong support from the model since Asp_{207} is deeply buried in the structure at about 25 Å from the entrance of the binding cleft. In addition, a very limited volume is left accessible in the neighborhood of this acidic residue. In contrast Asp_{311} is located in a very favorable position on TM3 near the extracellular surface of the receptor and is surrounded by two dissymetrical hydrophobic pockets, one of them containing the residues invariant within the cationic neurotransmitter receptors. From this arrangement, we propose that the binding site consists of Asp_{311}, representing the negative counterion for the positively charged nitrogen of the ligands, and a large hydrophobic pocket containing the conserved residues Trp_{307} (or Phe, Tyr) on TM3, Trp_{403} and Ser_{410} or Trp on TM4, Ser_{508} (or Ala), Phe_{509}, and Tyr_{510} (or Phe) on TM5, Phe_{609}, Trp_{613}, Phe_{616} (or Tyr), and Phe_{617} (or Asn_{616}) on TM6, and Tyr_{715} on TM7. Interestingly, three conserved proline residues are also found in the proposed binding pocket in the direct neighborhood of the conserved aromatic residues. Thus it seems possible that the proline residues together with the aromatic residues are directly involved in the conformational change of the receptor upon ligand binding (Figure 5).

The overall conserved functional amino acids (also found in the other GPCRs) are distributed throughout the 7 helices and are located in the binding region described above as well as near the cytoplasmic side of the receptor.

ii. 3D Models of Receptor–Neurotransmitter Complexes. Using the conclusions of the primary structure comparison, published site-directed mutagenesis data, and the 3D models described above, we have attempted to find an optimal fit for the natural agonists into their binding site in order to identify the exact binding mode responsible for their affinity, selectivity, stereospecificity, and efficacy.

The positively charged nitrogen of the cationic neurotransmitter ligands was positioned in the active site close to the negatively charged Asp_{311}. The rest of the ligand was docked into the adjacent hydrophobic pocket defined by TM4, TM5, TM6, and TM7. The interactions of the aromatic side chains with the natural ligands were modeled interactively taking general rules of aromatic–aromatic interactions into account[67] since molecular mechanics calculations do not account very well for the spatial charge distribution of π-electron systems.

The complexes thus obtained were geometry optimized and analyzed in terms of interaction energies and conformational properties. The energy-minimized complexes
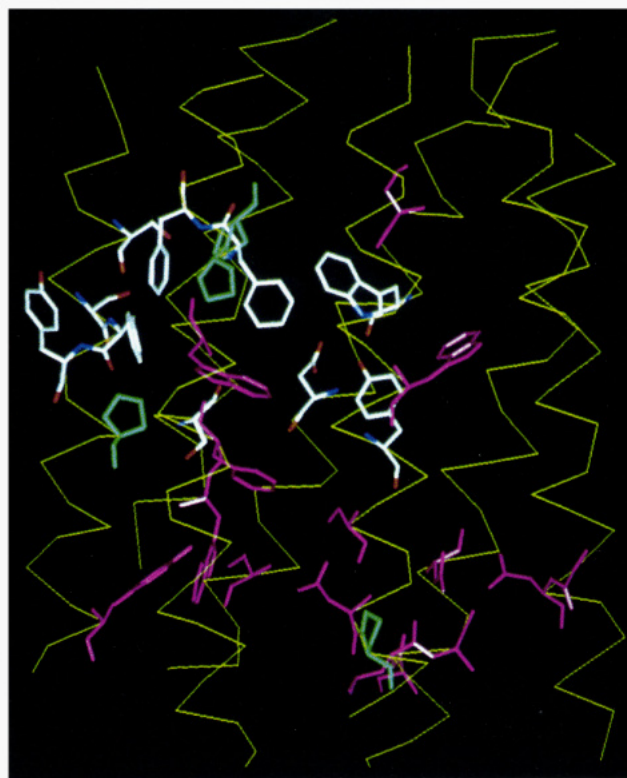

Figure 5. View of the seven membrane regions of the β_2-adrenergic receptor from the side. The seven helices are indicated by a c_α-carbon chain trace in yellow. Only the conserved residues are displayed. Residues conserved in the cationic neurotransmitter receptors only are colored by atom type. The residues which are conserved in all GPCR are colored in red-orange. Proline residues are shown in green. The residues which are conserved only in the cationic neurotransmitter receptors are concentrated around the putative agonist binding site.

with the most favorable interactions are shown in Figures 6–10. The results appeared to be much better than originally expected since the steric and electrostatic complementarity of the receptors with their ligands is striking and fits both with general rules of molecular recognition and with all published experimental data. Residues likely to be responsible for the binding potency, selectivity, stereospecificity, and efficacy could be identified. They were essentially of three types: acidic Asp residues, Ser and Asn residues, and aromatic residues (Figure 6–10).

The Asp Acidic Residues. The docking attempts for the five neurotransmitters reinforced that Asp_{311} on helix 3 belongs to the agonist binding site. No satisfactory receptor–ligand complex could be generated with Asp_{207}, essentially for reasons of steric hindrance. In contrast, ion pairs between Asp_{311} and the neurotransmitter cationic head could be made easily, including the hindered quarternary ammonium group of acetylcholine.

The Ser Connections. Examination of the receptor–ligand complexes indicated that several Ser residues on TM4 and TM5 might be important for the ligand affinity, selectivity, and stereospecificity. We observed in the model that Ser_{505} (Cys in the α_2 receptor) and Ser_{508} are ideally located to make hydrogen bonds with the catechol moiety of dopamine, noradrenaline, and adrenaline in the D_2-, α_2-, and β_2-receptor agonist binding sites, respectively. The m-hydroxyl group of these ligands interacts with Ser_{505} (or Cys) while the p-hydroxyl group can interact with Ser_{508} located on the same helix. The hydrogen bond with the m-hydroxyl group is very strong and possesses the ideal geometry. However, the hydrogen bond to the p-hydroxyl
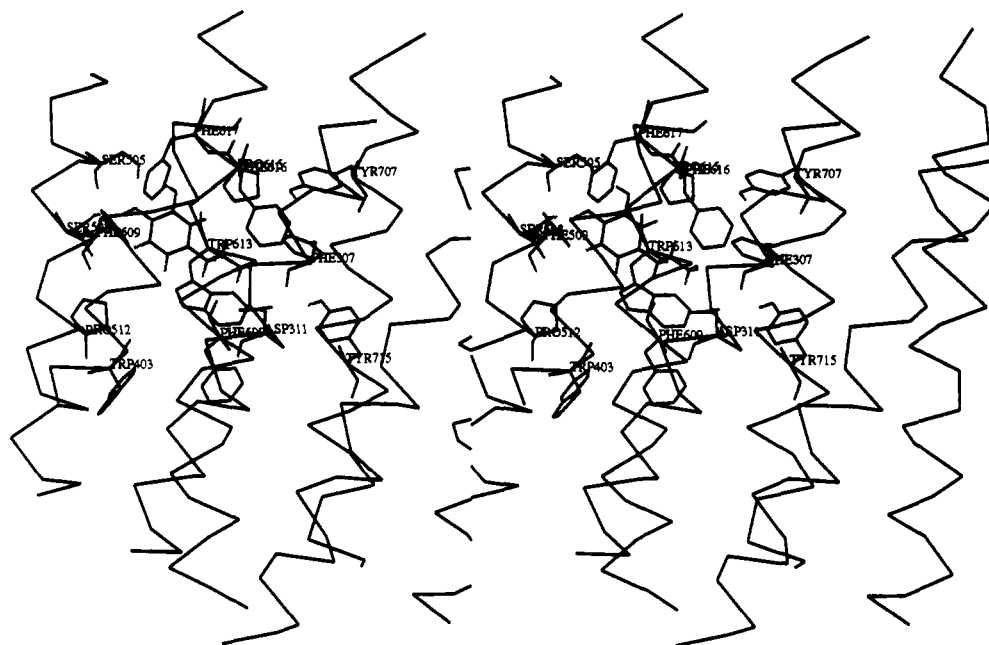
Figure 6. Stereoplots of the five neurotransmitters in their respective binding sites. Only residues of the binding site and the trace of the main chain are displayed. 6/ dopamine D_2; 7/ serotonin 5-HT_2; 8/ noradrenaline α_2; 9/ adrenaline β_2; 10/ muscarinic m_2. The cationic neurotransmitters form ionic interactions with the conserved Asp on helix 3.[6-10] This ion pair is surrounded by a cluster of three conserved aromatic residues (613, 616, and 307) belonging to helices 6 and 3. Catecholamine's hydroxy substituents can interact with the pair of conserved Ser_{505}, Cys_{505}, or Ser_{508} residues on helix 5.[6,8,9] The single hydroxy substituent of serotonin interacts with Ser_{505} (b). The β-hydroxy group of noradrenaline or adrenaline can stereoselectively form a hydrogen bond with Ser_{410}.[8,9] The indole nitrogen atom of serotonin can interact with Ser_{406}.[7] A conserved aromatic residue, Phe_{617}, can interact with the aromatic nucleus of the aromatic neurotransmitter.[6,7,9,10] In the case of acetylcholine receptors Phe_{617} is replaced by Asn_{617} which can make hydrogen bonds with the neurotransmitter ester group.
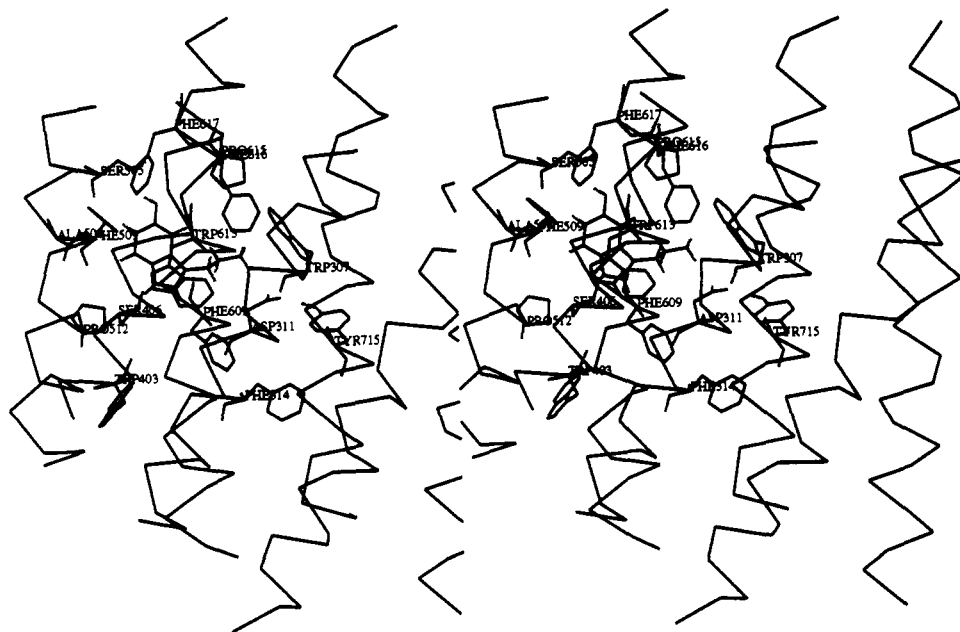
Figure 7.

group is much weaker with a longer bond length and a nonideal angle. This arrangement is in full agreement with published site-directed mutagenesis data discussed above.[68] Strikingly, only one of these two Ser residues, Ser_{505}, is present in the 5-HT receptors, Ser_{508} being replaced by Ala. Accordingly, we observed that the single phenol hydroxyl group of 5-HT in its receptor-bound conformation could make a hydrogen bond with Ser_{505}. Furthermore, both Ser residues are replaced by Ala in the muscarinic receptor for which the natural ligand, acetylcholine, is lacking phenolic hydroxyl groups. This shows the perfect complementary

between the ligands and their respective binding sites.

Two other Ser residues are essential to make selective or stereospecific interactions: Ser_{413} on TM4 in the α_2 and β_2 adrenoreceptors occupies the ideal position in the recognition site to form a stereospecific hydrogen bond with the β-hydroxyl group of the adrenaline and noradrenaline side chains; Ser_{406}, also on TM4, is perfectly located to accept a hydrogen bond from the indole NH of 5-HT.

It thus seems possible with the 3D models to explain the involvement of Ser_{505} and Ser_{508} in the catecholamine and 5-HT binding and to suggest that Ser_{406} and Ser_{410} can
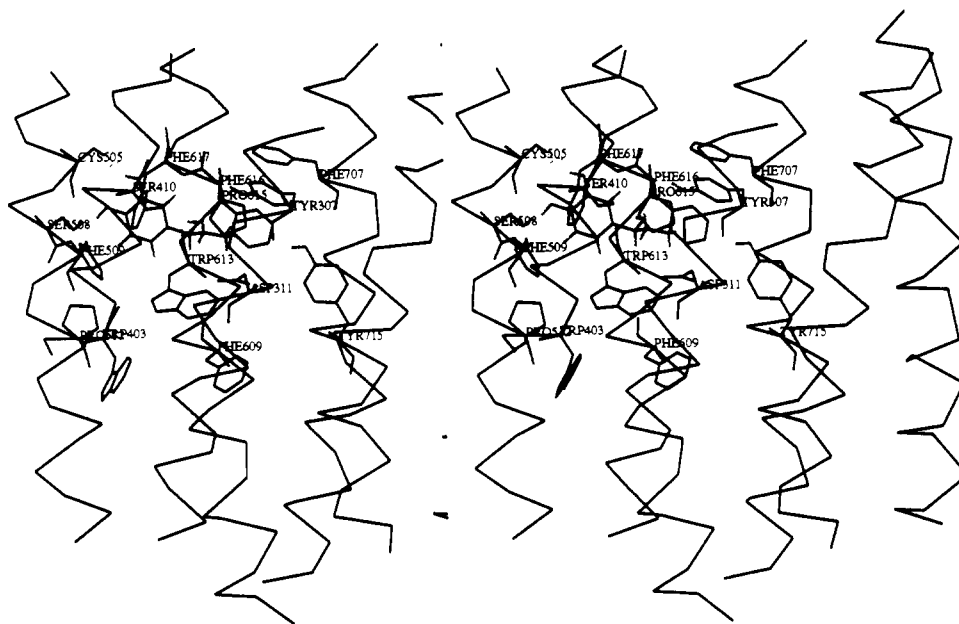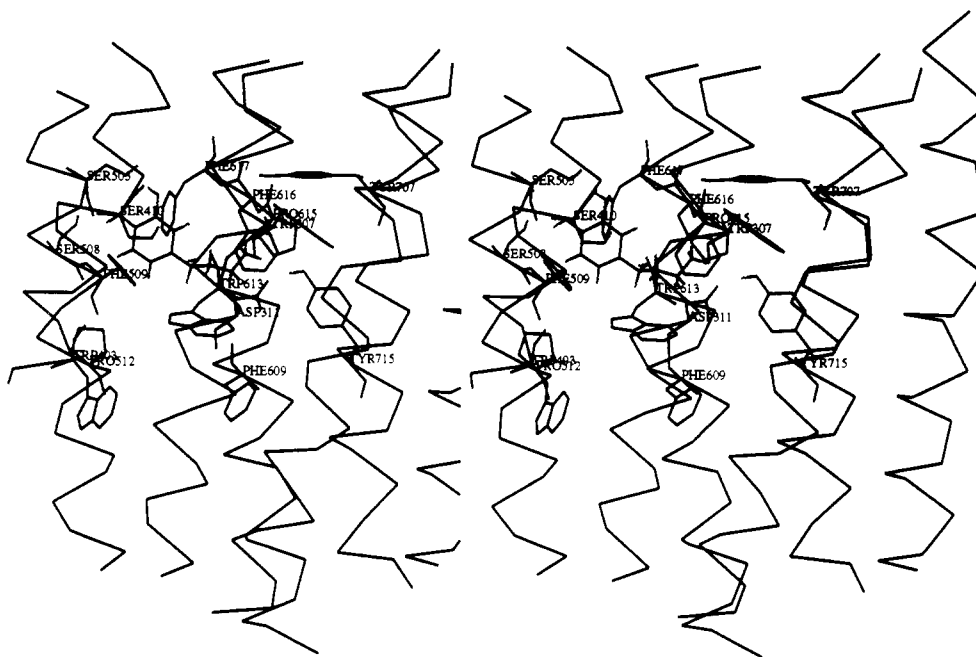
Figure 8.

Figure 9.

interact selectively or stereospecifically with 5-HT and adrenergic neurotransmitters, respectively. It is interesting to note that the essential anchoring points described above were similar in the catechol and indolamine receptors, with the exception of Ser_{505}. This is in agreement with the cross-activity of the corresponding ligands and the lack of selectivity of numerous synthetic analogues.

The Aromatic Connections. As mentioned above, despite the fact that many aromatic residues are highly conserved within all GPCRs or within GPCR subclasses, hypotheses concerning their functional role were formulated only for a very limited number of them. One of the most interesting observations derived from the GPCR 3D models concerns the interactions of the natural ligands with the aromatic residues of their receptors.

The most striking feature which is common to all cationic neurotransmitter GPCR is the presence of a cluster of conserved aromatic residues which encages the ammonium–aspartate ion pair: the positively charged ammo-

nium group of the ligand and its receptor counteranion (Asp_{311}) are surrounded by the conserved aromatic residues Trp_{307} (or Phe or Tyr) on TM3 and Trp_{613} and Phe_{616} (or Tyr in muscarinic receptor) on TM6. All of them can adopt a favorable geometry to form charge-transfer interactions with the positively charged ammonium group. This aromatic hydrophobic cluster probably strongly reinforces the stability of the receptor–ligand complex by strengthening the ionic interaction and exerting a shielding effect. Tyr_{715} on TM7 additionally can stabilize the complex by interacting with its phenol group with the positive headgroup of the ligand. Tyr_{707} (or Phe or Trp) on TM7 seems also able to interact with the charged headgroup of the ligand. This aromatic residue is located above Trp_{307} (or Tyr or Phe) in the 3D models and can form cation–π interactions with the ligand in a similar way. This interaction is only possible when residue 307 rotates to make space. However, the aromatic residue 707 on TM7 is conserved only on the catecholaminergic and cholinergic
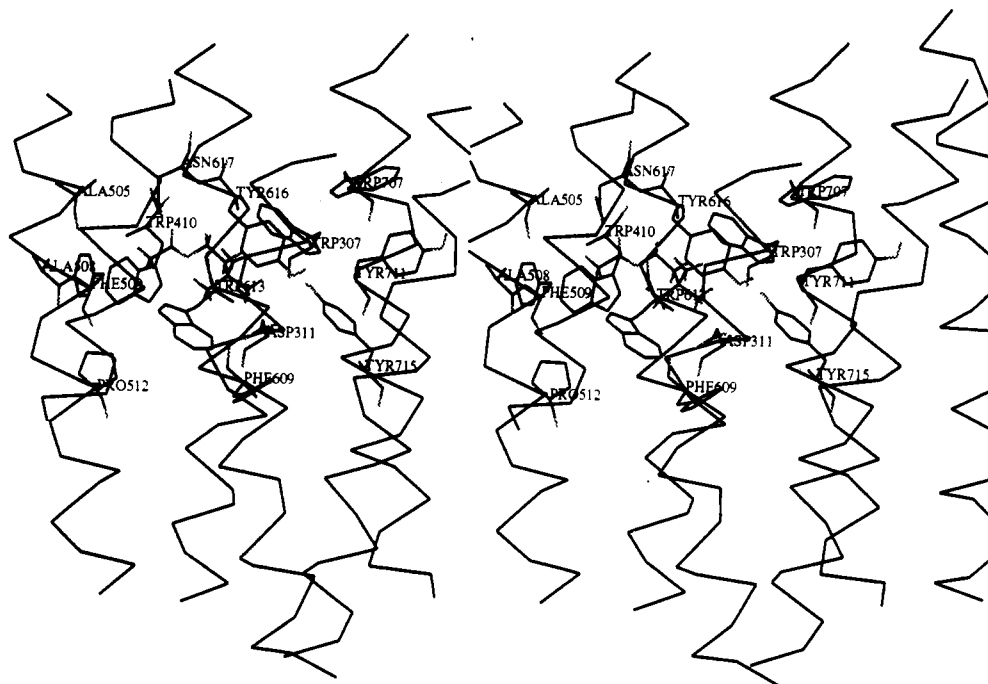
Figure 10.

receptors and is absent in the serotonergic receptors (with the exception of the 5-HT$_{1D}$ receptor). It seems very likely, that residue 707 together with neighbor residues are to a certain extent responsible for the selectivity of the receptor types and subtypes. Another important aromatic residue is Phe$_{617}$ on TM6 of the catecholamine and indolamine receptors. This Phe can form an additional stabilizing orthogonal interaction with the aromatic nucleus of the ligands and simultaneously interact with Trp$_{613}$ and Phe$_{508}$ in a similar manner. Together with Phe$_{509}$ on TM5 it defines a narrow cleft for the flat aromatic part of the ligands in the catecholamine and indolamine receptors.

Of prime interest is the substitution of Phe$_{617}$ by Asn in the muscarinic receptors whose natural ligand, acetylcholine, lacks an aromatic ring. Asn$_{617}$ is in fact ideal to interact with the ester group of acetylcholine via hydrogen bonding. It is also worth mentioning that five conserved tyrosine residues can be found in the muscarinic receptors in the direct neighborhood of the quaternary ammonium headgroup binding site (Tyr$_{221}$, Tyr$_{312}$, Tyr$_{616}$, Tyr$_{711}$, and Tyr$_{715}$).

Additionally, Ser$_{410}$ on TM4, which makes the stereoselective hydrogen bond in the catechol and indolamine receptors, is exchanged to Trp$_{410}$ which can form π–σ and π–π interactions with the ester group of acetylcholine.

The Signal Transduction. One of the major problems to be addressed is the understanding of the transduction from the neurotransmitter to the corresponding G-protein mediated by a conformational change in the receptor. The importance of some residues has been highlighted by mutagenesis studies but the mechanisms involved remain unclear. A careful analysis of the 3D models described above prompted us to suggest possible mechanisms, involving a network of the aromatic residues Trp$_{403}$, Phe$_{509}$, and Phe$_{617}$, the cluster of aromatic residues around the ammonium aspartate ion pair (Trp$_{307}$ or Phe or Tyr; Trp$_{613}$; and Phe$_{616}$ or Tyr) and Tyr$_{715}$ connecting the agonist binding domain to the helices 1, 2, and 7. The conformation of the side chain of these residues is necessarily modified upon ligand binding and could thus directly affect the receptor backbone conformation at the level of the neighboring Pro or Gly hinge residues. This would result

in a change in the accessibility of charged residues on the inner loops and thus in an altered coupling to G-proteins.

For example, in one of the most stable conformations of the receptor without ligand, Asp$_{311}$ can form a hydrogen bond with the conserved Trp$_{613}$ on TM6 and Tyr$_{715}$ on TM7. These interactions must be broken during the ligand complexation process, since the side chains must be reoriented in order to build the cage of aromatic residues around the ion pair as described above. Tyr$_{715}$ could then interact with the cationic headgroup of the ligand. It is also important to note, that two of the residues forming the aromatic cage (Trp$_{613}$ and Phe$_{616}$ or Tyr) are adjacent to Pro$_{615}$ on TM6. It thus appears very likely, that the building of this aromatic cluster together with Pro$_{615}$ represent a molecular switch able to trigger the receptor conformational change upon agonist binding. More particularly, Trp$_{613}$ seems to be a key residue in triggering the conformational change, since it is conserved in most of the GPCR. Mutation of the corresponding Trp$_{613}$ (Trp$_{265}$ numbering in bovine rhodopsin) to Phe or Ala in bovine rhodopsin produced mutant receptors with abnormal bleaching behavior and only a very low activation of transducin. From these results, the authors suggest that Trp is located close to retinal and may be involved in the crucial transition step that leads to the active form of rhodopsin.[70] Involvement of aromatic residues in local conformational changes of photoactivated rhodopsin has also been suggested by spectroscopic studies. In particular, from circular dichroism studies it was proposed that one tryptophan residue rotates during the metarhodopsin I to II transition.[71,72] This is in full agreement with the

(70) Tomoko, A.; Khorana, H. G. Mapping of the amino acids in membrane embedded helices that interact with the retinal chromophore in bovine rhodopsin. *J. Biol. Chem.* **1991**, *266*, 4269–4275.

(71) Rafferty, C. N. Light-induced perturbation of aromatic residues in bovine rhodopsin and bacteriorhodopsin. *Photochem. Photobiol.* **1991**, *29*, 109–120.

(72) Chabre, M.; Breton, J. The orientation of the chromophore of vertebrate rhodopsin in the "Meta" intermediate states and the reversibility of the II- Meta III transition. *Vision Res.* **1978**, *9*, 1005–1018.
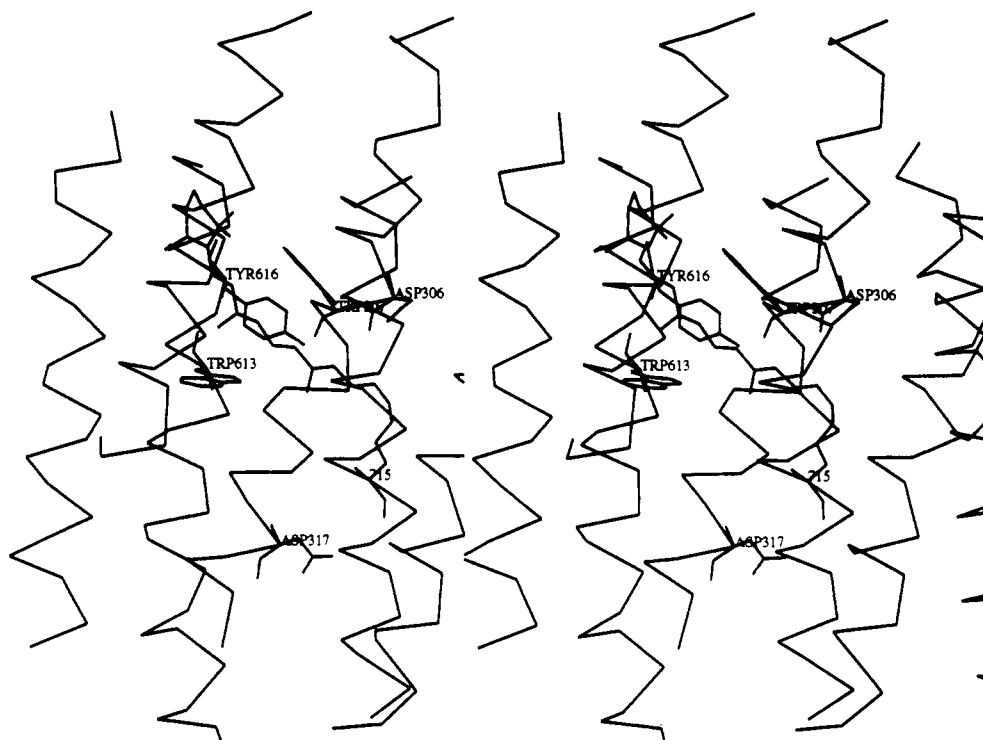
Figure 11. Stereoplot of the Henderson model of bacteriorhodopsin. Residues of the binding site and the trace of the main chain are displayed. Retinal is bound in its trans conformation and makes a Schiff base with Lys_{715} on TM7. The ionone ring of retinal is located in a hydrophobic pocket defined by TM4, TM5, and TM6. The chromophore is fixed by the three conserved aromatic residues Trp_{307} on TM3 and Trp_{613} and Tyr_{616} on TM6. Lys_{715} on TM7 and Trp_{613} and Tyr_{616} on TM6 are also found in the G-protein-coupled mammalian opsins. The aromatic residues are also conserved within the cationic neurotransmitter receptors and are able to form similar interactions with the cationic part of the agonist, as does the chromophore of retinal.

findings from the 3D models.

Similarly, Phe_{509} on TM5, which is also conserved in the cationic neurotransmitter GPCR, was found in the model to interact directly both with the ligand and with Trp_{613} on TM6. In addition, Phe_{509} is located in the neighborhood of Pro_{512} on TM5 which is also adjacent to Trp_{403} on TM4. Modifications in the side-chain conformation of these aromatic residues might be responsible for a change in geometry in TM5 and represent another possible component of the activation mechanism. Finally, Tyr_{715} on TM7 seems also to be involved in the conformational change. During the ligand binding, the Tyr residue side chain has to be reoriented. This rearrangement can induce conformational changes at the level of the conserved Asn_{721}-Pro_{722} located on the same helix. It thus appears from the 3D models that the binding can induce a cascade of events involving a number of conserved aromatic and proline residues.

It seems very likely that the mechanisms described above, in conjunction with more subtle dynamic changes, in fact occur in a concerted manner. Preliminary molecular dynamics simulations seem to provide some support for these hypotheses, but extensive studies will be necessary to further explore the relevance of these mechanisms.

iii. Relevance of Bacteriorhodopsin As a Template. The weakest starting hypothesis in this study is the choice of bacteriorhodopsin as a template for the packing of the seven α-helices, since there is only low primary structure homology between this protein and the GPCRs. However we could observe from the 3D models that the interactions between the cationic neurotransmitter ligands and their receptors are similar to those found between retinal and bacteriorhodopsin. The receptor-bound neurotransmitters and retinal in its trans conformation occupy identical positions within the transmembrane core of their respective receptor (Figure 11). In analogy with the location of the

ionone ring of retinal, the aromatic ring of the catechol and indolamines is located near the extracellular surface in a hydrophobic binding pocket, consisting of TM4, TM5, and TM6. Furthermore, the side chain with the quaternary nitrogen of the neurotransmitter ligands is located at the same relative position as the chromophore of retinal in bacteriorhodopsin. Three of the conserved aromatic residues found in cationic neurotransmitter receptors around the ion pair are found in bacteriorhodopsin in the same geometrical arrangement of TM3 and TM6 where they interact with the chromophore of retinal (see alignment Figure 2, TM6 and TM3). In analogy with Trp_{613} in the neurotransmitter receptors, Trp_{613} in bacteriorhodopsin is located directly below the chromophore. Tyr_{616} in bacteriorhodopsin is positioned at the side of the chromophore, similarly to Phe or Tyr_{616} which is positioned at the side of the positive ammonium headgroup in the cationic neurotransmitter receptor complexes. Trp_{307} in bacteriorhodopsin is positioned at the same place as is Trp_{307} in the cationic neurotransmitter receptors. The aromatic residues are pointing their π-electrons in the direction of the chromophore and thus perform favorable π-σ and π-π interactions.

To further validate the use of bacteriorhodopsin as a template the G-protein-coupled mammalian opsins were modeled. They represent a crucial link between bacteriorhodopsin and the G-protein-coupled receptors. Both receptors share a common ligand and have the same activation mechanism. In addition, the mammalian opsins transfer protons like bacteriorhodopsin.[72]

Construction of the 3D models of the G-protein-coupled mammalian opsin receptors indeed shows that there is an excellent structural homology to bacteriorhodopsin (Figure 12). The interactions of retinal with the protein are identical in both bacteriorhodopsin and the mammalian opsins. In both receptors Lys is located at the same
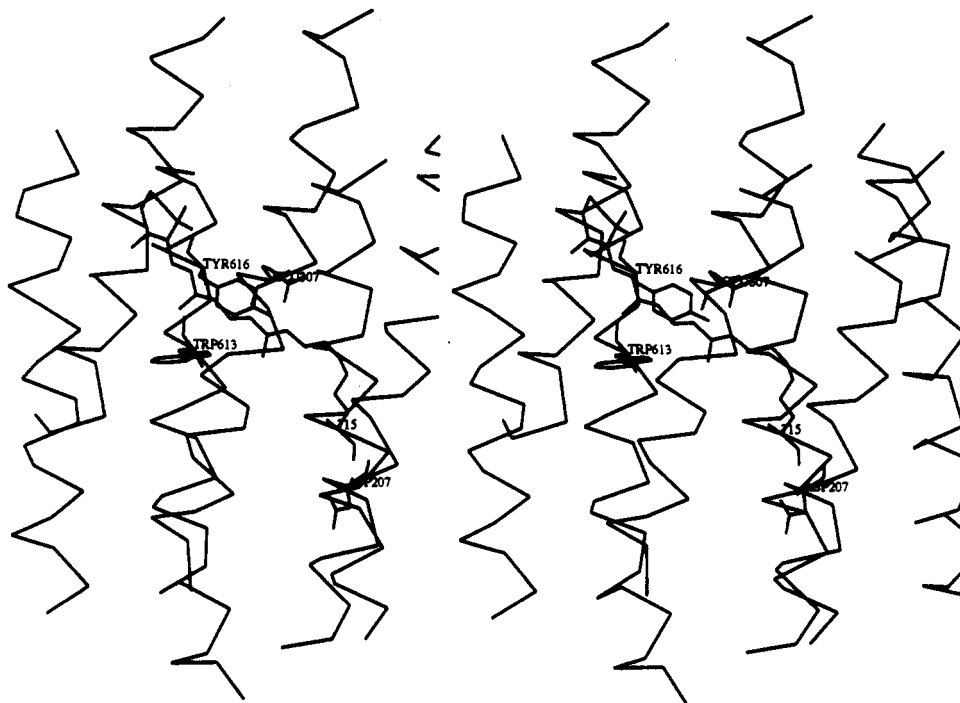
Figure 12. Stereoplot of the mammalian red-sensitive opsin model. Residues of the binding site and the trace of the main chain are displayed. The interactions of *trans*-retinal with the protein are identical to bacteriorhodopsin. Only Trp_{307} on TM3 is missing. Asp_{207} on TM2 and Glu_{307} on TM3, which could be involved in the proton transfer, are also displayed.

position on TM7, forming a Schiff base with retinal. Trp_{613} and Tyr_{616} on TM6 are also present in both receptors and form π–σ interactions with the chromophore of retinal. An acidic Asp or Glu is found in both receptors at geometrically the same place on TM3, acting as a counterion for the protonated Schiff base. The fact that the mammalian opsins are also able to transport protons during ligand binding[72] can be explained with the 3D models. Asp_{207} in the mammalian opsins could have the same proton transfer function as Asp_{317} (original numbering Asp_{85}) in bacteriorhodopsin, because it is located in geometrically the same place at the bottom of the receptor and at the same distance from the Schiff base. According to the model of bacteriorhodopsin, Glu_{307} on TM3 in the mammalian opsins could be protonated by the chromophore Schiff base, resulting in the release of a proton in the extracellular medium, whereas Asp_{207} on TM2 is involved in the reprotonation of the Schiff base and uptake of a proton from the cytoplasmatic medium. The allosteric function of Asp_{207} mediated by Na^+ or H^+ in α_2-adrenergic and D_2-dopaminergic receptors presents a striking homology with the proton transfer function insured by the homologous Asp_{207} in the mammalian opsins.

This structural similarity begs the question of an evolutionary relationship between the bacterial and mammalian opsins and all other GPCR. Since there is a very low sequence homology between bacteriorhodopsin and the GPCR, evolutionary relationships between them could only emerge from a conservation of their three-dimensional structure.

Examples are known for a number of protein families where the three-dimensional skeleton and the biochemical functions remain the same although the amino acid sequences diverge widely. In general, only few amino acids, which are directly involved in ligand binding, are conserved.[73]

Considering that bacteriorhodopsin and the mammalian opsins have identical ligands and activation mechanisms, it does not therefore seem surprising that the receptors possess the same three-dimensional arrangement and similar or identical amino acids responsible for interactions with retinal. In this case, the conservation of three-dimensional structure could reflect the importance of the environment of the ligand for the biological activity of the molecule. The conservation of the three-dimensional structure in evolution has been controlled by the ligand. In contrast, the striking sequence similarities and common three-dimensional arrangements within all GPCR reflects the importance of the three-dimensional structure for the biochemical function of these proteins, i.e. signal transduction upon extracellular stimulation.

Clearly the experimental data as well as the modeling techniques, which were used at this stage of the study, do not allow the definition of models at the atomic resolution. However, the most reasonable representation has to be generated in order to study the potential interactions between the ligand and the receptor chemical functions and to propose precise validating experiments to be performed.

Interaction Homologies with Other Cationic Neurotransmitter Binding Proteins. When these models were refined, one of their most interesting and original characteristics was the stabilizing interaction between the ligand ammonium group and the cluster of aromatic residues. We were not aware at that time of experimental data showing the feasibility and the magnitude of such an interaction, with the exception of photoaffinity labeling results.[74,75]

(73) Rao, S. T.; Rossmann, M. G. Comparison of super-secondary structures in proteins. *J. Mol. Biol.* **1973**, *76*, 241.

(74) Kieffer, B.; Goeldner, M.; Hirth, C.; Aerbersold, R.; Chang, J. Y. Sequence determination of a peptide fragment from electric eel acetylcholinesterase involved in the binding of quaternary ammonium. *FEBS Lett.* **1986**, *202*, 91–96.

(75) Galzi, J. L.; Revah, F.; Black, D.; Goeldner, M.; Hirth, C.; Changeux, J. L. Identification of a novel amino acid α-tyrosine 93 within the cholinergic ligand sites of the acetylcholine receptor by photoaffinity labelling. *J. Biol. Chem.* **1990**, *265*, 10430–10437.

Several publications have appeared since, highlighting examples of such interactions[76] or providing novel structural biological[77] or synthetic[78] examples of ⟨ammonium–aromatic cage⟩ interactions. Thus two X-ray structures of proteins binding a cationic neurotransmitter are now available: the three-dimensional structure of acetylcholinesterase and the phosphocholine binding FAB McPC603. These structures support the models.

In the three-dimensional structure of acetylcholinesterase, the binding site of acetylcholine is deeply buried in the protein structure. This site is of hydrophobic nature and consists of a number of electron-rich tyrosine and phenylalanine residues which can interact in a similar manner with acetylcholine[77] as in our muscarinic neurotransmitter receptor model. This is also in agreement with earlier photoaffinity labeling studies.[75]

Similar interactions are proposed from labeling experiments for the nicotinic acetylcholine receptor where a region in the α-subunit has been determined which forms part of the ACh binding site. Along with two cysteine residues (192, 193), a number of aromatic residues (Tyr$_{93}$, Tyr$_{149}$, Tyr$_{190}$, Tyr$_{151}$, Tyr$_{198}$) were labeled.[79,80,74] All these residues are conserved in the α-subunits of muscle acetylcholinesterases from all species examined to date. From these results the authors propose that the lone pairs of the oxygen of Tyr, the nitrogen of Trp, and the sulfur of Cys stabilize the charge of acetylcholine.

The second choline binding site for which detailed structural information is available is the phosphocholine (PCh)-binding immunoglobin Fab McPC603.[78,76] Two anionic residues (Asp^{971} and Glu^{61H}) are found near the ammonium headgroup of phosphocholine. However much closer contacts are made with Trp^{107H}, Tyr^{33H}, and Tyr^{100I}.

Thus, the trimethylammonium headgroup is surrounded by these three aromatic "walls". The interaction of Trp^{107H} with the ammonium is especially striking. In particular one of the ammonium N-methyl groups is located directly above the center of the benzene ring of Trp^{107H} with distances to the six aromatic carbon atoms ranging from 3.2 to 3.4 Å.

Although these are only a few examples, they reflect an emerging trend that a cluster of aromatic residues is found frequently at acetylcholine ammonium binding sites. It is clear from the GPCR models, that this seems to be not only an essential feature for the acetylcholine binding site but also for the catechol and indolamine binding sites.

The high homology in the recognition sites of different neurotransmitter receptors leads to the question of the origin of their selectivity for certain ligands. Together with other authors[78] we would propose that a cluster of aromatic residues produces a hydrophobic binding site which is able to strongly bind primary, secondary, tertiary, and quaternary ammonium ligands through a stabilizing cation–π interaction and an energetically favorable desolvation process. The selectivity of the recognition for a certain ligand results probably from interactions with other residues in the direct neighborhood of the primary binding site. They alter the shape and polarity of the receptor in order to position the ligand precisely and to take full advantage of electrostatic, hydrogen bonding, and Van der Waals interactions.

Conclusion

Starting from the refined atomic model of bacteriorhodopsin, it was possible to construct 3D models for GPCRs. Five related cationic neurotransmitter receptors and one mammalian opsin receptor were studied in detail. The models account for labeling experiments, mutagenesis experimental data, and traditional structure–activity relationship studies and permit a more detailed understanding of the structure of this class of transmembrane receptors. Receptor–ligand interactions could be proposed, leading to a hypothesis for the putative function of most of the conserved amino acids. Additionally, amino acids which might be involved in the conformational change upon ligand binding are proposed.

These models thus represent a novel important working hypothesis which will have to be validated by site-directed mutagenesis and labeling experiments. They might also represent a major breakthrough for drug design.

(76) Satow, Y.; Cohen, G.; Padlan, E.; Davies, D. Phosphocoline binding immunoglobin. Fab Mc PC603. An X-ray diffraction study at 2.7Å. *J. Mol. Biol.* **1986**, *190*, 593–604.

(77) Sussman, J. L.; Harl, M.; Frolau, F.; Oefner, O.; Goldman, A.; Toker, L.; Silman, I. Atomic structure of acetylcholine esterase from *Torpedo californica*: A prototypic acetylcholine binding protein. *Science* **1991**, *253*, 872–878.

(78) Dougherty, D. A.; Stauffer, D. A. Acetylcholine binding by a synthetic receptor: Implications for biological recognition. *Science* **1990**, *250*, 1558–1560.

(79) Kao, P.; Dwork, A.; Kaldany, R.; Silver, M.; Wideman, J.; Stein, S.; Karlin, A. Identification of the α-subunit half-cystine specifically labelled by an affinity reagent for the acetylcholine receptor binding site. *J. Biol. Chem.* **1988**, *259*, 11662–11665.

(80) Dennis, M.; Giraudat, J.; Kotziba-Hibert, F.; Goeldner, M.; Hirth, C.; Chang, J. Y.; Lazure, C.; Chrétien, M.; Changeux, J. P. Amino acids of the *Torpedo marmorata* acetylcholine receptor and subunit labelled by a photoaffinity ligand for the acetylcholine site. *Biochemistry* **1988**, *27*, 2346–2357.