Structural characterization and binding sites of G-protein-coupled receptors

Annette G. Beck-Sickinger

G-protein-coupled proteins constitute a superfamily of integral membrane proteins encompassing hundreds of receptors for all types of chemical messengers, as well as, for example, the key molecules of our light and smell sensory systems, bioactive amines, peptide hormones, neurotransmitters and even proteins. Because of their complicated organization with the characteristic seven transmembrane segments (7TM) it has as yet been impossible to structurally characterize these proteins by crystallography or magnetic resonance. However, a number of indirect methods to study the structure, ligand binding and signal transduction capacity of these proteins are known and summarized in this review.

The superfamily of G-protein-coupled receptors are integral membrane proteins characterized by amino acid sequences that contain seven hydrophobic domains (Figure 1). These are predicted to represent the transmembrane-spanning regions of the proteins and are the reason for the second name of this superfamily: 7TM- or heptahelix receptors. G-protein-coupled receptors are found in a wide range of organisms and many kinds of chemical messengers act through them, from classical monoamine messengers (adrenaline, acetylcholine, serotonin, histamine, dopamine, etc.) via lipids (prostaglandins, endogenous cannabinoids, etc.), neuropeptides [neuropeptide Y (NPY), substance P (SP), cholecystokinin (CCK), opioids, etc.] and peptide hormones (glucagon, angiotensin, bradykinin) to small proteins, such as chemokines, and large proteins (glycoprotein hormones, thrombin, etc.). All these messengers are involved in the transmission of signals to the interior of the cell through interaction with heterotrimeric G proteins. In addition, important sensory proteins, such as rhodopsin and the olfactory receptors (alone several hundred different molecules), belong to this superfamily.

Structural characterization of G-protein-coupled receptors

The overall size of G-protein-coupled receptors can vary significantly: from less than 300 amino acids (adrenocorticotrophin hormone receptor) to between 300–370 residues (odorant receptors and mammalian rhodopsins), up to 350–600 residues for peptide hormone and monoamine messenger receptors), 650–750 amino acids for protein receptors and even more than 1,100 amino acids for the metabotropic glutamate receptors. The hydropathy plots suggest that all of the members of this family contain seven hydrophobic domains of 20–25 amino acids in length, which are believed to represent the transmembrane regions. Based on structural similarities with the extensively characterized protein bacteriorhodopsin, for which electron diffraction data are available[1], these regions are predicted to be α-helices and to be orientated to form a ligand binding pocket (Figure 1). Of the 7TM receptors themselves only rhodopsin has been structurally characterized by cryoelectron microscopy analysis, which has confirmed the seven-helix transmembrane

Annette G. Beck-Sickinger, Department of Pharmacy, Federal Institute of Technology (ETH) Zürich, Winterthurerstr. 190, CH 8057 Zürich, Switzerland. tel: + 41 1 257 6063, fax: +41 1 262 1580, e-mail: beck-sickinger@pharma.ethz.ch

  1359-6446/96/$15.00. PII:S1359-6446(96)10042-8

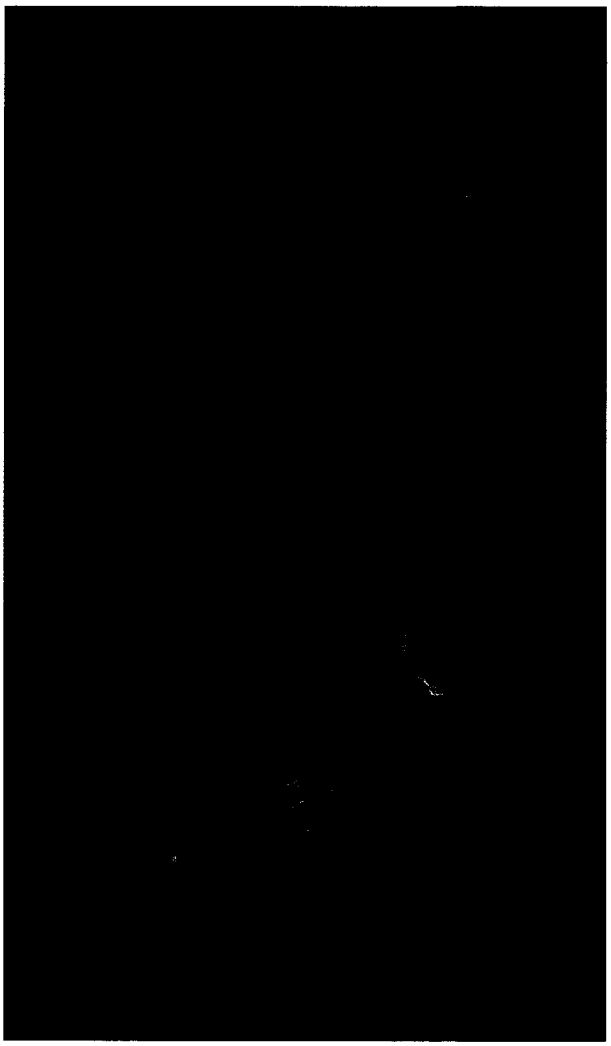

Figure 1. *3D Model of a 7TM, G-protein-linked receptor. Above: side view (extracellular part up, cytoplasm down); N- and C-terminal segments are cut. Below: view from top, the putative binding pocket of nonpeptide antagonists is indicated by an arrow.*

bundle[2]. However, the resolution is too low to give detailed information. Recently, site-directed electron paramagnetic resonance, which has given much information on bacteriorhodopsin, has been applied to rhodopsin, a true G-protein-coupled protein. One great advantage of site-directed electron paramagnetic resonance over other structure determination techniques is that it is suitable for the investigation of the dynamics in protein structures[3].

Protease digestion studies and detailed immunological mapping have provided evidence for the 7TM-model, including direct evidence that the N-terminal sequence is extracellular and the C-terminal sequence intracellular. Thus, each receptor is believed to have an extracellular N-terminal sequence that can vary in length from less than 10 amino acids (adenosine receptors) to several hundred (glycoprotein hormone receptors, metabotropic glutamate receptors), which is followed by three sets of alternate intracellular and extracellular loops and a final intracellular C-terminal sequence. The majority of intracellular and extracellular loops are predicted to be between 10 and 40 amino acids in length, although the third intracellular loop and the C-terminal sequence can have more than 150 residues, a common occurrence in the receptors for bioactive amines.

Whereas the seven helices of rhodopsin were characterized from electron density maps, the sense of the orientation (clockwise or anti-clockwise) remained unclear. However, the anti-clockwise orientation (seen from outside) has recently become much more probable, following evidence from a set of mutation experiments: by the incorporation of two His residues on different transmembrane helices in the tachykinin receptor, artificial Zn^{2+}-binding sites were created, if the distance between the His residues was close enough. Only the anti-clockwise orientation was in agreement with the identified positions on TM 3 and 5 or 2 and 6 (Ref. 4).

Common features of G-protein-coupled receptors

G-protein-coupled receptor subfamilies are characterized according to the lengths of the extra- and intracellular segments and the identity of specific residues. However, some structural features are common for all receptors. A higher molecular weight than that predicted from the amino acid sequence of a receptor is frequently found, and suggests that the receptor is glycosylated. This is supported by the observation that almost all 7TM receptors have one or more asparagine residues, which are present in a glycosylation consensus sequence N-x-T/S. Most of them are found in the N-terminus, though some receptors could be glycosylated in other extracellular loops. Glycosylation can increase the molecular weight by up to 20 kD and the added glycoside can constitute up to 30% of a receptor's total molecular weight. However, it is thought that glycosylation plays only a minor role in the agonist binding of most receptors, though it may be important in determining the correct distribution of the receptor in the cell and in controlling receptor expression. Expression of G-protein-coupled receptors in *Escherichia coli* or the treatment of receptors with endoglycosidases leads to receptors with similar

binding properties. Glycosylation-deficient mutant β_2-adrenoceptors exhibit a significant reduction in the part of the receptor appearing on the cell surface. In addition, mutation of rhodopsin Thr17, which is part of a glycosylation site, to Met has been identified in patients suffering from autosomal dominant retinitis pigmentosa, and it is thought that these patients express much less protein[5].

A cysteine residue at the carboxyl end of the first extracellular loop and a second one in the middle of the second extracellular loop are present in most G-protein-coupled receptors, and it is suggested that an important disulphide bridge is formed. In rhodopsin, it has been shown that this bridge is essential for obtaining the correct folding. Disruption of the bridge either by chemical reduction of the linkage or by replacing cysteine using site-directed mutagenesis has been shown to affect ligand binding. Additional disulphide bridges, involving different extracellular loops or the N-terminal segment, are present in some receptors; their position varies from subtype to subtype[6].

Subfamilies of G-protein-coupled receptors

The over 300 G-protein-linked receptors of which the primary structure is known can be divided into several subfamilies by various criteria. The type of the ligand, and thus the hypothetical binding mode, or the type of G protein a specific receptor can interact with are possibilities for further characterization of this huge protein family. Most frequently, however, the receptors are classified by primary-sequence homology and subfamilies are named after well-characterized members. While only low homology is found in the loop segments, the transmembrane helices contain a number of residues that are conserved for all or several receptors. An evolutionary tree can be drafted based on this homology.

Three main types of receptors can be distinguished: the rhodopsin subfamily (Figure 2a), the calcitonin subfamily (Figure 2b) and the glutamate-metabotropic-receptor subfamily. The only feature that all three subtypes have in common is the disulfide bridge between a cysteine residue at the top of TM3 and another in the second extracellular loop, which seems to be important for the overall structure. Bacteriorhodopsins, which are not G-protein-coupled proteins but proton pumps, are totally different with respect to amino acid sequence but have a seven-helix bundle similar to the G-protein-coupled receptors' helical portion[7].

Most of the receptors cloned so far belong to the *rhodopsin-like* subfamily, which includes receptors that bind bioactive amines, lipids, purines, kinins, endothelin, NPY, opsins and rhodopsins (Table 1). The N-terminal segment is highly glycosylated, which secures folding and surface expression. A number of polar residues within the transmembrane segments are conserved: Asn18 on TM1, Asp10 on TM2, Arg26 on TM3 and Asn16 on TM7. These residues probably create a network of hydrogen bonds that is important for the structure of the receptor. The charged residues Asp10 on TM2 and Arg26 on TM3 may also play an important role in signal transduction. Conserved proline residues in TMs 4–7 are possibly involved in the dynamics of the receptors, for example in the change between active and inactive receptor. Conserved hydrophobic residues are also found. A fourth intracellular loop, created by a palmitoylated Cys-residue in the C-terminal segment, could be involved in receptor desensitization. The second intracellular loop and parts of the third intracellular loop and of the C-terminal segment are involved in G-protein interaction, and therefore in signal transduction. In addition to the highly conserved residues mentioned above, closely related receptors have further identical residues. For example an Asp residue in TM3 is found in monoamine receptors, but not in peptide, purine or protein receptors.

Mainly peptide hormones and neuropeptides bind to receptors of the *calcitonin-like* subfamily. Glucagon-like peptide I, glucagon, gastrin inhibitory peptide, secretin, vasoactive intestinal peptide (VIP), calcitonin and calcitonin-gene-related peptide are some of the ligands characterized so far. The large N-terminal segment (100–150 residues) of this subfamily contains six cysteine residues, which could be interconnected by a number of disulphide bridges hypothetically forming a globular ligand-binding domain. There are also several conserved proline residues, located at different positions from those in the rhodopsin-like receptors.

The third and so far smallest group of G-protein-coupled receptors are characterized by a huge N-terminal segment (500–600 amino acids) and a similarly large C-terminal domain. The G-protein-coupled *glutamate receptors* (so called metabotropic receptors) and the calcium sensor of the parathyroid receptor belong to this subfamily. The transmembrane segments are connected by short loops and share no sequence homology. Interestingly, the large extracellular domain resembles that of a family of bacterial transporting proteins with known X-ray structure. Mutation analysis of metabotropic receptors suggest that glutamate binds in between two large subdomains in the N-terminal

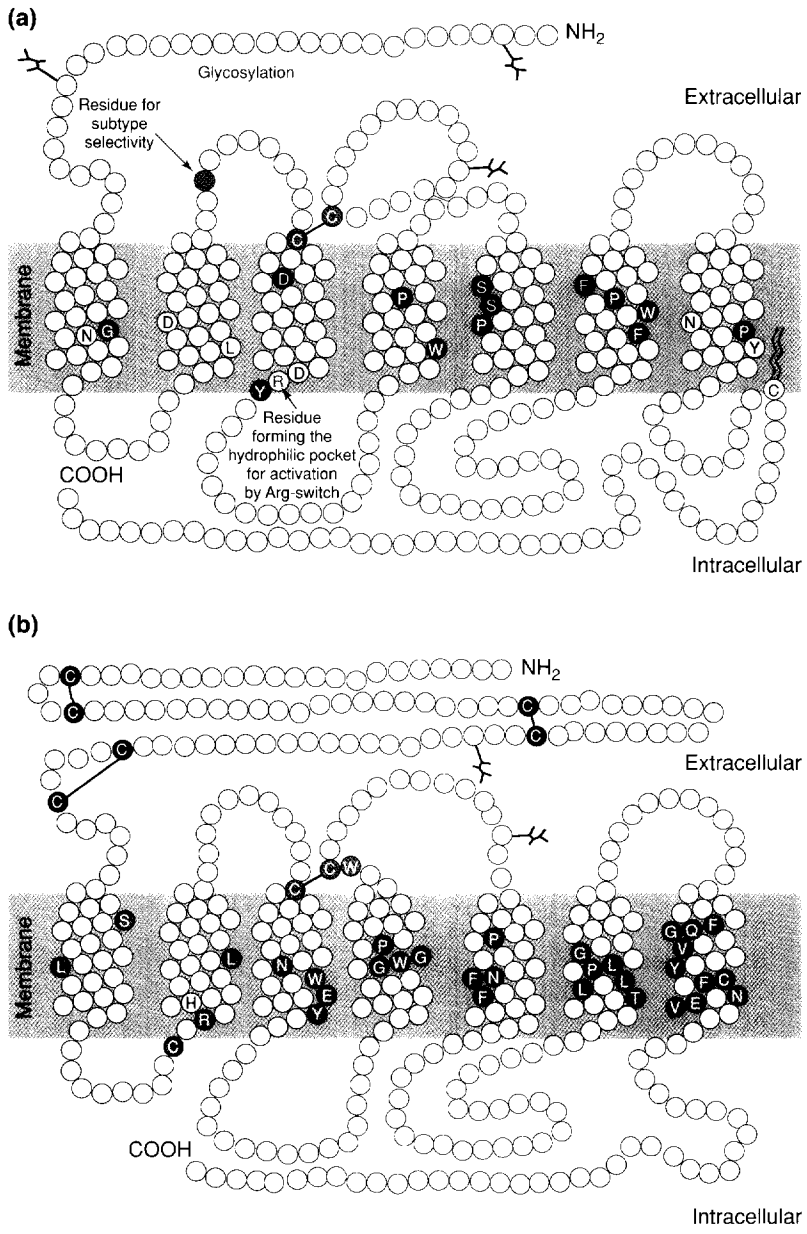

Figure 2. *Characteristics of two main G-protein-coupled receptor subfamilies: (a) Rhodopsin-like receptors frequently have a small N-terminal segment that is glycosylated and highly conserved residues (red) in the transmembrane segments. A residue that was identified for subtype selectivity of peptide receptors is shown in green, residues important for binding of adrenaline in the β_2-receptor are shown in blue. Residues that play a role in signal transduction are marked in yellow, and the very conserved Arg in transmembrane helix 3, which has been found to act as a switch, is labelled in yellow/green. (b) Calcitonin-like receptors have a large N-terminal segment with glycosylation sites, six cysteine residues (green) and a number of conserved residues (red) in the transmembrane segments. A conserved His residue (yellow) in transmembrane helix 2 has been identified recently to be involved in signal transduction[13].*

segment in the same way that amino acids bind to their transporting proteins in bacteria[18].

Ligand binding: agonists, antagonists and inverse agonists

So far, two main hypotheses that describe the interaction of a ligand and its receptor have been proposed. For a long time, it was believed that agonists and antagonists bind to the receptor in a similar way. An agonist induces a conformational change that leads to signal transduction, whereas for an antagonist, this is not possible (or favourable). By contrast, a number of recent experiments support the hypothesis of conformational selection[9]. According to this hypothesis, 7TM receptors exist in at least two conformations, of which at least one binds G proteins and is, therefore, active (R*). In other conformations the receptor cannot interact with G proteins, and is referred to as an inactive, resting or uncoupled receptor (R) (Figure 3). Usually, uncoupled receptors are dominant, and the equilibrium constant L, which is defined by the equation L = [R]/[R*], is large[10]. However, small numbers of active receptors are responsible for a low, but consistent signal, even when no ligand is available. Ligands have different affinities for the two states of a receptor; these affinities are characterized by K_A and K_A^*. If ligands bind to the active conformation with high affinity, $K_A/K_A^* >> L$ (K_A^* = [R*] [Lig]/[R*Lig]; K_A = [R] [Lig]/[RLig]), active receptor molecules are then dominant and the compound is called an agonist. Compounds that bind to the inactive conformation can be found, too. A K_A/K_A^* ratio that is much smaller than the equilibrium constant L results in a reduction of the number of constitutively active receptors; in this case the ligand functions as an inverse agonist[11,12]. The latter are characterized by the fact that they lower the level of second messenger compared with the basal level. Ligands that bind to both receptor

Table 1. Some of the most important ligands that bind to G-protein-coupled receptors and their predominant signal transduction pathways[46,47]

Receptor subfamily	Ligand or ligand family	Receptor	G Proteins (α-subunit)	Effect
Rhodopsin-like (monoamine)	Acetylcholine	M_1, M_3, M_5	G_q/G_{11}	PLC↗
		M_2, M_4	G_i/G_o	cAMP↘, Ca^{2+}ch↘
	Adrenaline, noradrenaline	α_{1A}, α_{1B}, α_{1D}	G_q/G_{11}	PLC↗
		α_{2A}, α_{2B}, α_{2C}	G_i/G_o	cAMP↘, Ca^{2+}ch↘
		β_1, β_2, β_3	G_s	cAMP↗
	Dopamine	D_1, D_5	G_s	cAMP↗
		D_2, D_3, D_4	G_i/G_o	cAMP↘, Ca^{2+}, K^+↗
	Histamine	H_1	G_q/G_{11}	PLC↗
		H_2	G_s	cAMP↗
	Melatonin	ML_{1A}, ML_{1B}	G_i/G_o	cAMP↘
		ML_2	G_q/G_{11}	PLC↗
	Serotonin	5-HT_{1A-F}	G_i/G_o	cAMP↘, other
		5-HT_{2A-C}	G_q/G_{11}	PLC↗
		5-HT_3	?	Cation channel
		5-$HT_{4,6,7}$	G_s	cAMP↗
		5-HT_5	?	
Rhodopsin-like (peptide)	Angiotensin	AT_{1A}, AT_{1B}	G_q/G_{11}	PLC↗
		AT_2	?	cGMP↘
	Bombesin family			
	Neuromedin B	BB_1	G_q/G_{11}	PLC↗
	Gastrin-releasing peptide	BB_2	G_q/G_{11}	PLC↗
	Ligand?	BB_3	?	
	Bradykinin	B_1, B_2	G_q/G_{11}	PLC↗
	CCK/gastrin family			
	CCK	CCK_A	G_q/G_{11}	PLC↗
	Gastrin	CCK_B	G_q/G_{11}	PLC↗
	Endothelin ET-1, ET-2, ET-3	ET_A, ET_B	G_q/G_{11}	PLC↗
	GnRH	GnRH receptor	G_q/G_{11}	PLC↗
	Melanocortin family			
	ACTH	ACTH receptor	G_s	cAMP↗
	α-, β-, γ-MSH	MSH receptor	G_s	cAMP↗
	Neurotensin	Neurotensin receptor	G_q/G_{11}	PLC↗
	Neuromedin N			
	Opioid family			
	Endorphin	μ-receptor	G_i/G_o	cAMP↘, other
	Dynorphin	κ-receptor	G_i/G_o	Ca^{2+}ch↘
	Enkephalins	δ-receptor	G_i/G_o	cAMP↘
	PP family			
	PP	PP_1/Y_4	G_i/G_o	cAMP↘
	Neuropeptide Y, peptide YY	Y_1	G_i/G_o	cAMP↘
		Y_2	G_i/G_o	cAMP↘, Ca^{2+}ch↘
		Y_5, Y_6	G_i/G_o	cAMP↘
	Somatostatin	SST_1, SST_2, SST_3,	G_i/G_o	cAMP↘
		SST_4, SST_5	G_i/G_o	cAMP↘
	Tachykinin family			
	Substance P	NK_1	G_q/G_{11}	PLC↗
	Neurokinin A	NK_2	G_q/G_{11}	PLC↗
	Neurokinin B	NK_3	G_q/G_{11}	PLC↗
	TRH	TRH receptor	G_q/G_{11}	PLC↗
	Vasopressin	V_{1A}, V_{1B}	G_q/G_{11}	PLC↗
		V_2	G_s	cAMP↗
	Oxytocin	OT	G_q/G_{11}	PLC↗

Rhodopsin-like (protein)	FSH	FSH receptor	G_s	cAMP↗
	LH/hCG	LH receptor	G_s	cAMP↗
	Thyrotrophin (TSH)	TSH receptor	G_s	cAMP↗
Rhodopsin-like (protease-activated)	Thrombin cleaves N-terminus of receptor = ligand: TFRIED	Thrombin receptor (PAR_1)	G_q/G_{11}, G_i/G_o	cAMP↘ and PLC↗
	N-Terminal segment: SLIGRL	PAR_2	?	
Rhodopsin-like (chemokine)	C-C Chemokines	CC CK_1,	G_i/o	PLC↗, Ca^{2+} mobilization
	(MIPs, MCPs, eotaxin, and others)	CC CK_{2B}, CC CK_3, CC CK_4, CC CK_5	?	
	C-X-C Chemokines	$IL\text{-}8_A$	G_i/G_o	PLC↗
	(IL-8, NAP-2, GROα)	$IL\text{-}8_B$	G_i/G_o	PLC↗
Rhodopsin-like (miscellaneous)	Leukotrienes	BLT,		
	LTB, LTC, LTD,	$CysLT_1$,		
	LTE, LTF	$CysLT_2$		
	Chemicals (odorous signals)	Odorant receptors		
	11-*cis*-Retinal (light)	Red, green, blue opsins, rhodopsin	Transducin	cGMP-phosphodiesterase↗
	Platelet activating factor (PAF)	PAF receptor	?	PLC↗
	Prostanoids			
	D	DP	G_s	cAMP↗
	F	FP	G_q/G_{11}	PLC↗
	I	IP	G_s	cAMP↗
	T	TP	G_q/G_{11}	PLC↗
	E	EP_1, EP_3	G_q/G_{11}	PLC↗
		EP_2, EP_4	G_s	cAMP↗
	Adenosine	A_1, A_3	G_i/G_o	cAMP↘
		A_{2A}, A_{2B}	G_s	cAMP↗
	Purine			
	ADP	P_{2T}	G_i/G_o	cAMP↘
	ATP	P_{2Y}	G_q/G_{11}	PLC↗
	UTP	P_{2U}, P_2	G_q/G_{11}	PLC↗
Calcitonin-like	Calcitonin family			
	Calcitonin	Calcitonin receptor	G_s	cAMP↗
	CGRP	$CGRP_1$	G_s	cAMP↗
	Amylin	?		
	Adrenomedullin	Adrenomedullin receptor	G_s	cAMP↗
	Glucagon	Glucagon receptor	G_s	cAMP↗
	GLP-1	GLP-1 receptor	G_s	cAMP↗
	Parathyroid hormone (PTH)	PTH receptor	G_s	cAMP↗
	VIP family			
	VIP	VIP_1, VIP_2	G_s	cAMP↗
	PACAP	$PACAP_1$	G_s	cAMP↗
	GRF			
	PHM, PHI			
	Secretin			
Metabotropic glutamate receptors	Glutamate	$mGluR_1$, $mGluR_5$	G_q/G_{11}	PLC↗
		$mGluR_{2,3,4}$, $mGluR_{6,7,8}$	G_i/G_o	cAMP↘

Abbreviations: PLC, phospholipase C; Ca^{2+}ch, Ca^{2+}-channels; ↘, inhibition; ↗, activation; ACTH, adrenocorticotrophin; CGRP, calcitonin gene related peptide; CCK, cholecystokinin; FSH, follicle stimulating hormone; GLP, glucagon-like peptide; GIP, gastrin inhibitory peptide; GnRH, gonadotrophin releasing hormone; GRF, growth hormone releasing factor; LH, luteinizing hormone; hCG, human choriogonadotrophin; MCP, monocyte chemotactic protein; MIP, macrophage inflammatory peptide; MSH, melanocyte stimulating hormone; NAP, neutrophil-activating peptide; PACAP, pituitary adenylyl cyclase activating peptide; PAR, protease-activated receptor; PHI, peptide histidine isoleucine; PHM, peptide histidine methionineamide; PP, pancreatic polypeptide; TRH, thyrotrophin releasing hormone; VIP, vasoactive intestinal peptide.

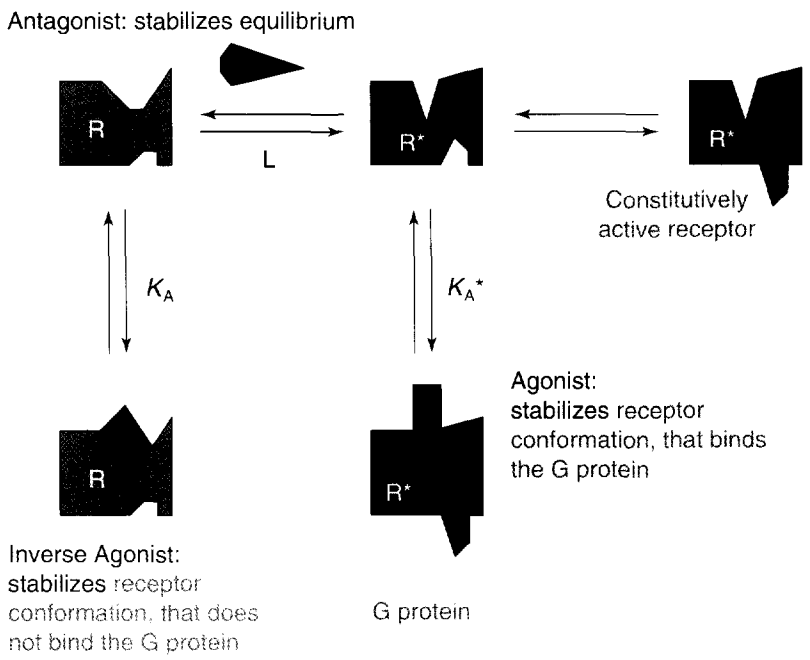

Figure 3. *Model of receptor–ligand interaction according to Ref. 10. The equilibrium between inactive (R) and active receptor conformations (R*) is characterized by the constant L. Agonists and inverse agonists stabilize one conformation and shift the equilibrium, whereas competitive antagonists stabilize both conformations equally.*

conformations and do not change the number of activated receptors ($K_A/K_A^* = L$) are called silent and function as competitive antagonists. By contrast, if a compound binds to the active conformation with K_A/K_A^* only slightly higher than L, it would be able to change the concentration of activated receptor only slightly and display low, but detectable efficacy; compounds of this type are called partial agonists. Similar effects can be found if a ligand binds slightly better to the active than the inactive conformation.

Energy differences between R* and R are generally small. This is shown by single amino acid mutations that lead to constitutively active receptors, thus decreasing the value of L significantly[13–15]. A further indication of this labile equilibrium is that antibodies can act as agonists[16] and very slight chemical differences can turn an agonist into an antagonist.

A recent debate suggests that there is no dichotomy between the conformational induction and conformational selection mechanisms. Both can be seen as two extremes of the same mechanism, assuming that there are more than two receptor conformations available and an agonist can stabilize various conformations by forming thermodynamically favourable complexes. Conformational induction thus can be seen as an extreme case of conformational selection, when an agonist stabilizes a rare active conformation[17].

Binding-site locations in G-protein-coupled receptors

The location of the binding site differs from subtype to subtype and also depends on the physical properties of the ligand (Figure 4). A unique property of the rhodopsin and the opsin receptors is that they bind ligands covalently and that receptor activation occurs in response to the absorption of photons. The former is illustrated by the reaction of the aldehyde function of 11-*cis*-retinal (a rhodopsin ligand) with Lys296 in TM7 to form a Schiff-base linkage. The photon-induced receptor activation is thought to proceed via a conformational change in the photo-excited rhodopsin involving TM3 and TM5 (Ref. 18).

Monoamines (histamine, adrenaline, dopamine, serotonin, acetylcholine), lipids and purines bind to sites located in the upper part of the TM helices. For example, the direct binding of the adrenaline to Asp8 in TM3 was identified by site-directed mutagenesis. The Asp8→Ser mutation changed the functional-group selectivity from amine to carbonyl: the ionic interaction was replaced by a polar one. The catechol ring of adrenaline interacts with two serine residues at the top of TM5 of the β_2-receptor and also performs a π–π-interaction with Phe in TM7 (Ref. 19) (Figure 2a, blue). Other amines are believed to bind in a similar way[20] (Figure 5). Antagonists that are chemically very similar to the agonist have been shown to occupy much of the same space. However, additional sites of interaction, for example at the top of TM7 have been identified.

The binding of peptide ligands to the rhodopsin-like receptor has been studied extensively in the past few years. Important findings have been obtained for tachykinins, endothelin, opioid peptides, angiotensin, vasopressin and NPY (for reviews see Refs 6,21). Several binding modes with major interaction sites in the exterior segments were found. The top of TM2, the first and the second extracellular loop and part of the N-terminal segment were identified for binding to the tachykinins. In addition, the top parts of TMs 4–6 are

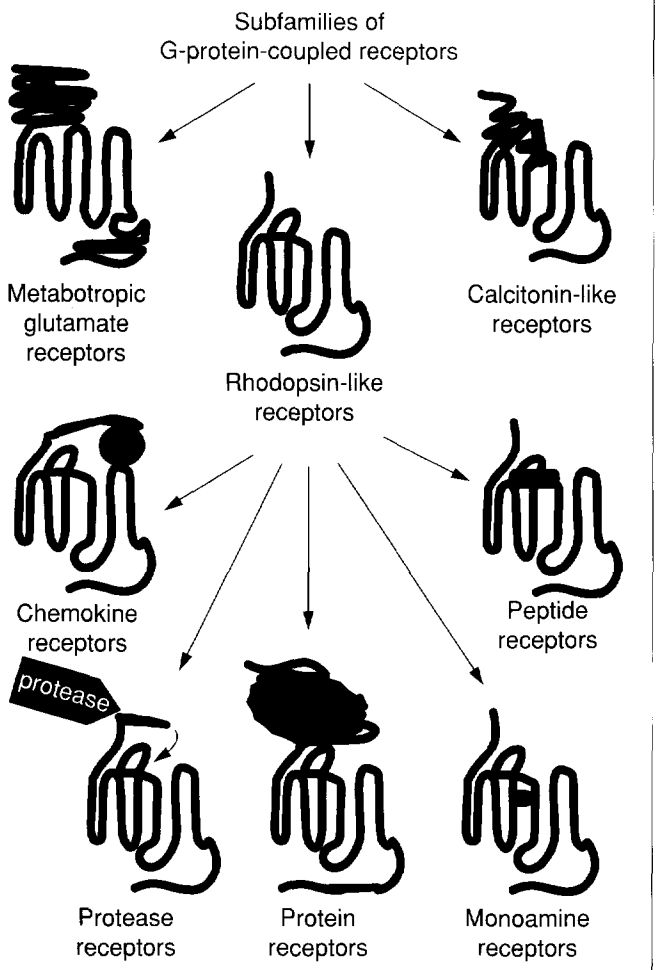

Figure 4. *Subfamilies of G-protein-coupled receptors (according to homology) and their putative ligand binding sites: ligands are shown in red, receptors in black.*

important for the binding of larger peptides, such as NPY. One residue of the first extracellular loop was found to be important for ligand selectivity, because subtype selectivity could be changed by mutation of this residue in the vasopressin/oxytocin system[22] (Figure 2a, green). Interestingly, the binding site of a number of nonpeptide antagonists was found to be located deeper into the membrane, interacting with residues from TMs 3, 4, 5, 6 and/or 7. Whereas agonist and antagonist binding sites are different for the small peptides (tachykinins, angiotensin)[23,24], larger and overlapping binding sites have been found for larger peptide hormones, such as NPY (Ref. 25) and endothelin, and their antagonists. The deep crevice between TMs 3, 5 and 6 is not accessible for peptides, as was shown for SP (Ref. 26).

Although the glycoprotein receptors have a rhodopsin-like structure, these receptors have a number of specific characteristics and thus represent a discrete subfamily. The N-terminal segment of glycoprotein receptors contains 9–14 copies of an imperfectly repeated sequence of 25 residues, called the leucine-rich repeat. The functional significance of this structure is not known. However, using chimeric receptors it could be shown that this N-terminal segment is important for ligand binding.

The chemokine receptor family, which has been discovered recently[27,28], binds homologous small proteins (4–8 kDa), the so-called chemokines. Among them, interleukin-8 is the best characterized ligand. The N-terminal regions of the chemokine receptors and the third extracellular loop have been identified to be important for ligand binding.

Activation of the thrombin receptor involves a novel proteolytic mechanism. Thrombin, a protease, cleaves its receptor to create a new segment at the N-terminus that functions as a tethered ligand to activate the receptor[29]. The second extracellular loop has been shown to be involved in the binding of this segment[30]. A receptor with high

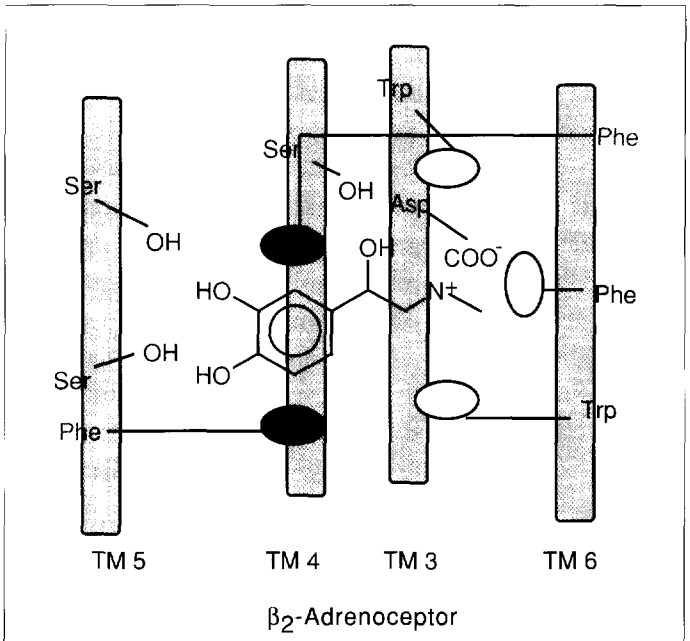

Figure 5. *Model of the interaction of adrenaline (red) with the β_2-adrenoceptor residues that were identified to be involved in binding: blue residues indicate polar/ionic interaction, green residues suggest hydrophobic interaction and yellow residues are involved in π–cation interactions.*

homology to the thrombin receptor, called proteinase-activated receptor-2 (PAR-2), was cloned recently[31]. Because this receptor is also activated by its N-terminal segment after protease cleavage, it seems that this particular activation pathway is rather more common than initially expected.

While the N-terminal segment plays only a minor role in the ligand binding of the rhodopsin-like receptors, it is responsible for ligand selectivity and affinity of calcitonin-like receptors. Using chimeric VIP/secretin[32] and glucagon/GLP-1 receptors[33], which both belong to the calcitonin-like subfamily, it was found that the receptors' large N-terminus with six cysteine residues (Figure 2b, green) and the first extracellular loop are involved in ligand binding and selection. Constitutively active mutants have been obtained for this receptor subfamily by mutation of a conserved His residue into Arg in TM 2 (Figure 2b, yellow)[13].

Activation of G-protein-coupled receptors

Binding of an agonist to the G-protein-coupled receptor stabilizes the latter's active conformation, which is characterized by the ability to bind the heterotrimeric G protein $G_{\alpha\beta\gamma}$ and cause its GTP-dependant dissociation into G_{α} and $G_{\beta\gamma}$. Thus, the conversion of the inactive heterotrimer $G_{\alpha\beta\gamma}$, which binds GDP in the G_{α}-subunit, to the active GTP-binding G_{α}-subunit is catalysed by the agonist–receptor complex. G_{α}-GTP is able to activate or inhibit further effector molecules, such as adenylate cyclase, phospholipase C and ion channels or phosphodiesterases, dependent on the subtype of G protein. Adenylate cyclase is inhibited by G_i- or G_o-proteins and activated by G_s-subtypes; G_q-proteins activate phospholipase C, which cleaves phosphaditylinositol 4,5-bisphosphate. The regulation of both enzymes leads to an increase or decrease in the release of a second messenger, such as cyclic AMP, inositol 1,4,5-trisphosphate or diacylglycerol. Na^+-, Ca^{2+}- and K^+-channels have been reported to act as effector systems and thus vary the intracellular concentration of certain cations. So all systems are pathways to transmit a signal from the extracellular side of the membrane to the cytoplasm by increasing or reducing the amount of a compound. As either enzymes are activated (or deactivated), or channels opened (or closed) the signal is additionally potentiated, because one agonist–receptor complex can activate several G proteins, and each of them will further activate an enzyme that produces a number of second messengers.

Responses to G proteins undergo a rapid desensitization during continuous stimulation and this is likely to be caused by receptor phosphorylation. In the intracellular segments of all G-protein-coupled receptors many serine and threonine residues are found, the majority in consensus sequences for the phosphorylation by G-protein-coupled receptor kinases. It has been shown that phosphorylation of the active conformation of the β_2-adrenoceptor does take place and allows a protein, called β-arrestin, to bind, which prohibits signal transduction.

Whereas many aspects of the signal transduction pathway are reported in textbooks, little is known about the molecular structure of active receptors. Site-directed mutagenesis, deletion and chimeric receptor studies have been used to identify the regions that couple to certain G proteins. Constitutively active mutants of the adrenoceptors have been identified in the third intracellular loop. The third intracellular loop was found to be important for the interaction with the G_s-protein of the β_2-receptor. However, residues in the second intracellular loop and the overall 3D structure on the cytoplasmic side of the membrane also seem to play a major role. Highly conserved residues, which frequently have a functional role in proteins, are found near the bottom of TM3: the motif Asp (or Glu)-Arg-Tyr is present in all rhodopsin-like receptors and is speculated to play a major role in signal transduction. Site-directed mutagenesis into Gly-Gly-Ala eliminates coupling to G-proteins[34]. The Arg is speculated to have a switch function, thus either binding to Asn in TM7 and Asp in TM2, the so-called hydrophilic pocket (inactive conformation), or shifting towards the cytosol (active conformation). The shift may be caused by a rearrangement of the pocket. Further residues involved in this hydrophilic pocket (Figure 2a, yellow) could be Asn (TM1) and Tyr (TM7), whereas Leu (TM2) could be necessary to orientate Arg by hydrophobic interaction[35,36].

Methods to study binding to G-protein-coupled receptors

Systematic variation and random screening of the ligand

To investigate the interaction of ligand and receptor, analogues of a high-affinity ligand are often synthesized. Through systematic variation of the ligand hundreds of slightly different compounds are prepared. Important functional groups can be distinguished from less important ones, and analogues with antagonistic properties can suggest segments that are relevant for signal transmission. Computer

models of 7TM receptors have been created based on the rhodopsin structural model[7,17,22,25,37]. These have been used to suggest new ligands, thus initiating a rational approach for the development of new drugs and reducing the number of organic syntheses required.

New strategies for the synthesis of peptide fragments have been developed in the past few years, which make it possible to synthesize a large number of peptides in parallel[38]. This has led to a new area in peptide hormone research. Systematic approaches to peptide synthesis have resulted in the identification of binding sites of peptidic ligands[39] (Figure 6). However, peptides can adopt several conformations in aqueous solutions, but adopt only one specific conformation at their receptors. Different receptor subtypes may recognize different conformations of the same peptide. In order to characterize receptor subtypes, and also to find smaller selective peptides or eventually nonpeptide drugs, the knowledge of the bio-active conformation of a neuropeptide agonist or antagonist is essential. Various strategies to constrain the conformation of small peptides have been tried, including cyclization and the incorporation of unnatural amino acids or spacer templates.

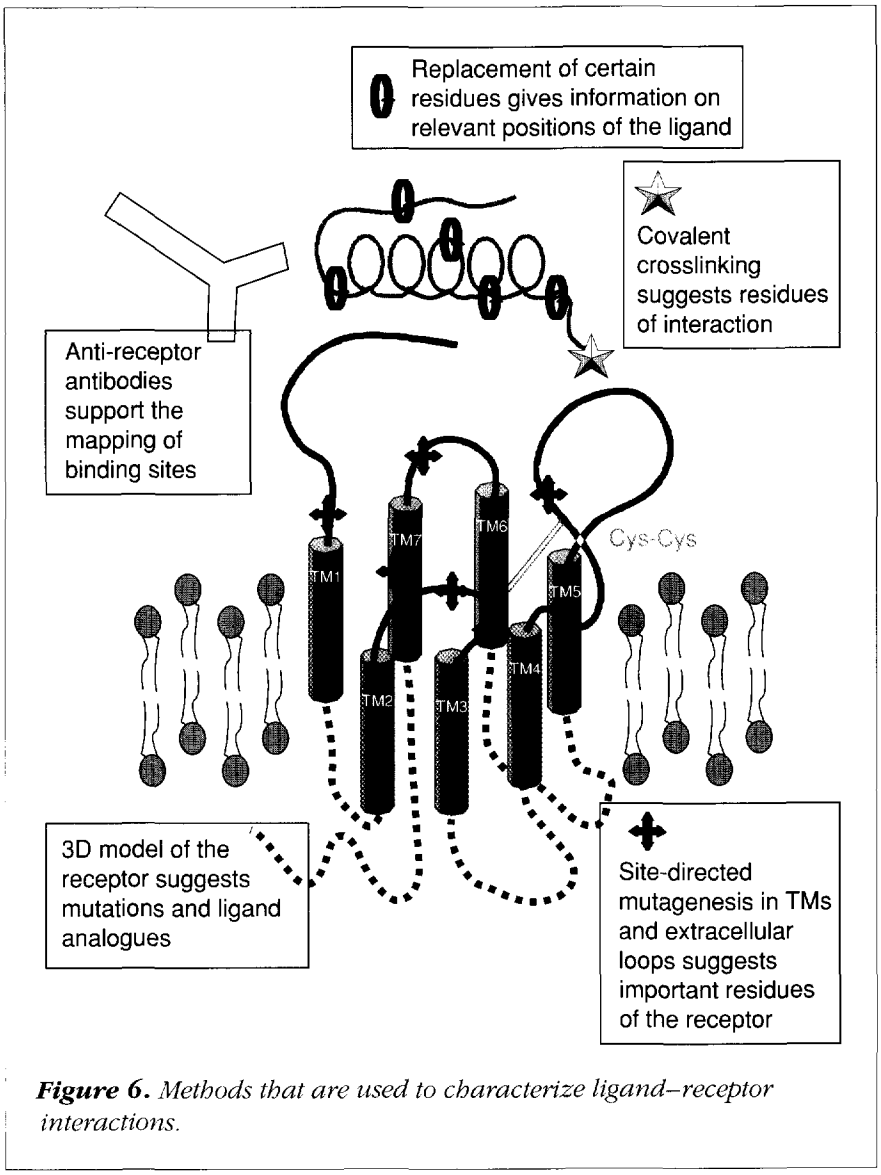

Figure 6. *Methods that are used to characterize ligand–receptor interactions.*

Recent strategies for the synthesis of new compounds include combinatorial chemistry, synthesis of peptide and nonpeptide libraries and solid-phase organic synthesis. In these protocols, mixtures of ligands are synthesized, and these are then tested for their ability to interact with the receptor of interest. The active compound is then identified either by testing less and less complex mixtures or by a direct method. In the latter case the polymer-bound peptide or oligonucleotide, which is either the ligand or is coding for an organic molecule that has been synthesized in parallel, is identified by sequencing[40]. However, high-throughput screening is still used very effectively to identify high-affinity ligands. Huge compound pools, which have been collected in each company over the years from various projects, are available as soon as a new project with a suitable testing system comes up. Those compound pools, in fact, are a major asset for companies, because the more diverse ligands are available, the more likely it is that a new lead is discovered. The major challenge is finding a strategy for selecting the compounds to be tested in the first step from the thousands of ligands stored.

Chimeric receptors and point mutations

One of the most frequently used tools to characterize the ligand–receptor interaction is the creation of special mutants. Studies of agonist and antagonist affinities to the mutant will suggest the relevance of a specific residue (Figure 6). However, most frequently mutants with loss of function/binding are generated and it becomes impossible to distinguish between indirect (conformational) and direct effects. So the effect of a single mutation is always questionable[6].

Either a number of ligands should be tested with the mutant to identify the important functional groups, or a set of mutations should be investigated. In particular, the mutation of a specific residue to Ala, which leads to a smaller side-chain, can be combined with a mutation to a larger side-chain in order to test the steric interactions of a residue and test the available space in the binding crevice. Using molecular modelling, the binding site that is suggested by the data obtained by mutagenesis can be hypothesized and used to forecast a next set of mutated receptors. These lead to a further optimization of the receptor models[22,25].

Crosslinking

Analyses of molecular interactions between ligands and their receptors are mainly performed indirectly by either synthetic modifications of the ligand or by receptor mutagenesis. However, whether the loss of affinity is due to a direct or an indirect interaction is difficult to determine. An alternative approach is to use the ligand to label the part of the receptor that it interacts with. Cross-linking of the ligand with the receptor via a linker molecule is an option (Figure 6). However, this method is not necessarily suited for identifying the binding site as bifunctional reagents will often cross-link the ligand with the receptor at 14–16 Å from the binding site.

An alternative method for the covalent binding of the ligand to its receptor is the use of a ligand with a built-in photoreactive amino acid that generates a highly reactive species, for example a carbene, a nitrene or a diradical. The exact knowledge of the ligand and the essential amino acids of peptide hormones, however, is of great significance[41]. Photocrosslinking using a ligand-bound photoactive amino acid has been used successfully to characterize hormone receptors recently. The central CCK receptor and the neurokinin 1 receptor were labelled using a *p*-benzoyl-Orn(propionyl) analogue and *p*-benzoylphenylalanine analogue of the native ligand, respectively, and also a photoactive analogue of the nonpeptide antagonist[42]. Identification of the first extracellular loop as part of the binding site of the V_2 vasopressin receptor was achieved with an analogue containing Lys(4-azidobenzimide)[43], and a photoactive β_2-receptor antagonist was used to label segments of transmembrane helices[44].

Anti-receptor antibodies

Various studies have shown that antibodies produced against hormone receptors are valuable tools. Beside immunization using purified or enriched receptors, in particular immunizations by the use of receptor fragments for obtaining antireceptor-antibodies have been described. Molecular mass determinations by SDS-PAGE and subsequent western blot, receptor purification by affinity chromatography on antibody-columns and investigations on the receptor localization, degradation, synthesis, turnover and covalent modifications are only a few applications of antireceptor-antibodies. Moreover, if the binding site of a monoclonal antibody is known, or if the antibodies are obtained against receptor fragments, then localization of the functional groups would be possible by competition with ligands[45] (Figure 6).

Future use of G-protein-coupled receptors in drug design

G-protein-coupled receptors transmit signals from a wide range of endogenous ligands, such as hormones, neurotransmitters and neurohormones. Many data have been obtained in recent years as a result of significant improvements in ligand-synthesis techniques, advances in homology expression, site-directed mutagenesis and protein-expression systems. However, many features of G-protein-coupled receptors, including their 3D structure and dynamics, the binding-site of many ligands and the pathway of signal transduction, all necessary for the successful rational design of agonists and antagonists, are not yet fully understood and new techniques are required to improve our understanding.

ACKNOWLEDGEMENTS

I thank Beate Rist, Peter Weber, Nik Ingenhoven, Christophe Eckard and Susanne Esser for their enthusiastic working. I appreciate the stimulating discussion and the collaboration with Thue W. Schwartz in the field of G-protein-coupled receptors. I gratefully acknowledge the support of the Swiss National Fund for grant No. 31-40860.94 and the Federal Institute of Technology for grant 41-2310.5.

REFERENCES

1 Schertler, G.F.X., Villa, C. and Henderson, R. (1993) *Nature* 362, 770–772
2 Unger, V.M. and Schertler, G.F.X. (1995) *Biophys. J.* 68, 1776–1786
3 Hubbell, W.L. and Altenbach, C. (1994) *Curr. Opin. Struct. Biol.* 4, 566–573
4 Thirstrup, K. *et al.* (1996) *J. Biol. Chem.* 271, 7875–7878
5 Macke, J.P. *et al.* (1993) *Am. J. Hum. Genet.* 53, 80–89
6 Schwartz T.W. (1994) *Curr. Opin. Biotechnol.* 5, 434–444
7 Hoflack, J., Trumpp-Kallmeyer, S. and Hibert, M. (1994) *Trends Pharmacol. Sci.* 15, 7–9

8 O'Hara, P.J. *et al.* (1993) *Neuron* 11, 41–52
9 Lefkowitz, R. *et al.* (1993) *Trends Pharmacol. Sci.* 14, 303–307
10 Leff, P. (1995) *Trends Pharmacol. Sci.* 16, 89–97
11 Chidiac, P. *et al.* (1994) *Mol. Pharmacol.* 45, 490–499
12 Barke, E.L. *et al.* (1994) *J. Biol. Chem.* 269, 11687–11690
13 Turner, P.R., Bambino, T. and Nissenson, R.A. (1996) *Mol. Endocrinol.* 10, 132–139
14 Ren, Q. *et al.* (1993) *J. Biol. Chem.* 268, 16483–16487
15 Westphal, R.S., Backstrom, J.R. and Sandersbush, E. (1995) *Mol. Pharmacol.* 48, 200–205
16 Abu Alla, S. *et al.* (1996) *J. Biol. Chem.* 271, 1748–1755
17 Kenakin, T. (1996) *Trends Pharmacol. Sci.* 17, 190–191
18 Hargrave, P.A. and McDowell, J.H. (1992) *FASEB J.* 6, 2323–2331
19 Strader, C.D., Sigal, I.S. and Dixon, R.A.F. (1989) *FASEB J.* 3, 1825–1832
20 Trumpp-Kallmeyer, S. *et al.* (1992) *J. Med. Chem.* 35, 3448–3462
21 Hunyady, L., Balla, T. and Catt, K.J. (1996) *Trends Pharmacol. Sci.* 17, 135–140
22 Trumpp-Kallmeyer, S. *et al.* (1995) *Pharm. Acta Helv.* 70, 255–262
23 Gether, U. *et al.* (1993) *Nature* 362, 345–348
24 Rosenkilde, M.M. *et al.* (1994) *J. Biol. Chem.* 269, 28160–28164
25 Sautel, M. *et al.* (1996) *Mol. Pharmacol.* 50, 285–292
26 Schwartz, T.W. and Rosenkilde, M.M. (1996) *Trends Pharmacol. Sci.* 17, 213–216
27 Horuk, R. (1994) *Trends Pharmacol. Sci.* 15, 159–165
28 Power, C.A. and Wells, T.N.C. (1996) *Trends Pharmacol. Sci.* 17, 209–212
29 Gerszten, R.E. *et al.* (1994) *Nature* 368, 648–651
30 Nanevicz, T. *et al.* (1996) *J. Biol. Chem.* 271, 702–706
31 Nystedt, S. *et al.* (1995) *Eur. J. Biochem.* 232, 84–89
32 Holtmann, M.H., Hadac, E.M. and Miller, L.J. (1995) *J. Biol. Chem.* 270, 14394–14398
33 Buggy, J.J. *et al.* (1995) *J. Biol. Chem.* 270, 7474–7478
34 Ohyama, K. *et al.* (1992) *Biochem. Biophys. Res. Commun.* 189, 2790–2795
35 Oliveira, L. *et al.* (1994) *Trends Pharmacol. Sci.* 15, 170–172
36 Scheer, A. *et al.* (1996) *EMBO J.* 15, 3566–3578
37 Baldwin, J.M. (1993) *EMBO J.* 12, 1693–1703
38 Jung, G. and Beck-Sickinger, A.G. (1992) *Angew. Chem.* 104, 375–391; *Angew. Chem., Int. Ed. Engl.* 31, 367–383
39 Beck-Sickinger, A.G. *et al.* (1994) *Eur. J. Biochem.* 225, 947–958
40 Früchtel and Jung, G. (1996) *Angew. Chem., Int. Ed. Engl.* 35, 17–42
41 Brunner, J. (1993) *Annu. Rev. Biochem.* 62, 483–514
42 MacDonald, D. *et al.* (1996) *Mol. Pharmacol.* 49, 808–813
43 Kojro, E. *et al.* (1993) *Biochemistry* 32, 13537–13544
44 Hockerman, G.H. *et al.* (1996) *Mol. Pharmacol.* 49, 1021–1032
45 Bahouth, S.W., Wang, H. and Malbon, C.C. (1991) *Trends Pharmacol. Sci.* 12, 338–343
46 Watson, S. and Arkinstall, S. (1994) *The G-Protein Linked Receptors* (FactsBook), Academic Press
47 Receptor & Ion Channel Nomenclature Supplement, 7th Edition (1996) *Trends Pharmacol. Sci.*, Elsevier

In short...

Scientists at the University of Virginia and Mitokor have identified a genetic defect in the mitochondrial DNA that may cause **Parkinson's disease** (PD). The defect was found to affect the production of complex I, a mitochondrial enzyme that serves as a starting point of the electron transport chain and a known factor in the onset on PD. Experiments with cytoplasmic hybrid cells containing mtDNA from PD patients showed that Complex I activity was 20% lower in PD hybrid cells than in the controls. Also, the PD hybrid cells were found to be more susceptible to MPTP, a parkinsonism-inducing toxin. The study is published in *Annals of Neurology* (1996) 40, 663–671. **Mitokor** (San Diego, CA, USA) is a biopharmaceutical company that focuses solely on the role of mitochondria in human disease.

SciClone Pharmaceuticals' lead product, thymosin α1 [tradename: Zadaxin(R)], achieved statistically significant results in the treatment of **chronic hepatitis B** in a study carried out by Dr P. Andreone and coworkers (University of Bologna, Italy). The randomized and controlled clinical trial involved 33 patients who were hepatitis B virus DNA (HVB DNA) and hepatitis B e antibody (HBeAb) positive. Seventeen patients were treated with thymosin α1 (twice weekly) and 16 patients with interferon α (three times weekly) for six months; 15 clinically similar patients were used as controls. Six months after completion of the treatment 41% of the thymosin α1 treatment group, 25% of the interferon α treatment group and 7% of the control group showed HBV DNA loss and normalization of alanine transaminase (ALT) levels. Results were reported in *Hepatology* (1996) 24, 774–777 and *Journal of Viral Hepatitis* (1996) 3, 191–196. Thymosin α1 is a synthetically produced, naturally occurring, 28-amino-acid peptide that is believed to stimulate the immune system.