Modeling and Docking the Endothelin G-Protein-Coupled Receptor

A. J. W. Orry and B. A. Wallace*
Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, United Kingdom

ABSTRACT A model of the endothelin G-protein-coupled receptor (ET_A) has been constructed using a segmented approach. The model was produced using a bovine rhodopsin model as a template for the seven transmembrane α-helices. The three cytoplasmic loop regions and the C-terminal region were modeled on NMR structures of corresponding segments from bovine rhodopsin. The three extracellular loops were modeled on homologous loop regions in other proteins of known structure. The N-terminal region was modeled as a three-helix domain based on its homology with a hydrolase protein. To test the model, the FTDOCK algorithm was used to predict the ligand-binding site for the crystal structure of human endothelin. The site of docking is consistent with mutational and biochemical data. The principal sites of interaction in the endothelin ligand all lie on one face of a helix that has been implicated by structure-activity relationship studies as being essential for binding. As further support for the model, attempts to dock bigET, an inactive precursor to endothelin that does not bind to the receptor, found no sites for tight binding. The model of the receptor-ligand complex produced forms a basis for rational drug design of agonists and antagonists for this G-protein-coupled receptor.

INTRODUCTION

Human endothelin-1 (ET-1) is a 21-amino acid polypeptide (Yanagisawa et al., 1988) that is cross-linked by two disulfide bonds and is the most potent vasoconstrictor yet characterized. The structure of ET-1 has been determined by x-ray crystallography (Janes et al., 1994). ET-1 is processed from a large precursor protein via the formation of a 38-residue intermediate called bigET-1 (Warner et al., 1994), which is biologically inactive. At least three different but highly homologous isoforms of human ET have been isolated.

The endothelins are important regulators of the vascular system and appear to be involved in the pathophysiological mechanism of a number of vascular conditions, including hypertension (Kohno et al., 1991), acute renal failure (Shibouta et al., 1990), and angina pectoris (Toyooka et al., 1991). They act via a G-protein-coupled receptor (GPCR)-mediated signal transduction pathway. The specific receptor to which ET-1 binds is a seven transmembrane helical protein (7TMS) called ET_A. This receptor is distinguished from the highly homologous ET_B receptor by its ability to discriminate between the various ET isoforms (Sakamoto et al., 1993).

GPCRs are integral membrane proteins that are characterized by the presence of seven hydrophobic domains, usually ranging in length from 20 to 25 residues, which represent the transmembrane regions of this large superfamily of proteins. Members of the superfamily are found in a wide range of organisms and are responsible for the transmission of a variety of signals to the interior of the cell. The signals can be activated initially by small peptides, amino acid derivatives, lipid analogs, or stimuli such as odor, light, and taste, depending on the receptor type. The activated cascade then passes information to the inside of the cell via interaction of the receptor with heterotrimeric G-proteins.

Knowledge of the three-dimensional structures of GPCRs is essential for our understanding of their function and for the rational design of drugs. Electron diffraction and electron microscopy (EM) studies have succeeded in determining the structure of a GPCR-like 7TMS protein, bacteriorhodopsin (BR), at medium resolution (Henderson et al., 1990; Gregorieff et al., 1996), and one GPCR, rhodopsin, at low resolution (Unger and Schertler, 1995; Unger et al., 1997). The x-ray crystallographic structure of BR has also been determined at high resolution from microcrystals grown in a cubic lipidic phase (Edman et al., 1999; Luecke et al., 1999). However, until recently there has been no high-resolution structural information on any GPCR (see Note added after submission). Consequently, a number of model-building studies have been undertaken to produce GPCR structures (for example, Findlay and Eliopoulos, 1990; Grotzinger et al., 1991; Hibert et al., 1991; Ijzerman et al., 1992; Lewell, 1992; Cronet et al., 1993; Sylte et al., 1993; Yamamoto et al., 1993; Zhang and Weinstein, 1993; Van Rhee et al., 1995; Pogozheva *et al.,* 1997; Baldwin et al., 1997; Herzyk and Hubbard, 1998; Konvicka et al., 1998).

Bacteriorhodopsin comprises seven transmembrane segments, all of which are helical in nature. The greatest sequence similarities between the GPCRs and bacteriorhodopsin are located in these regions, suggesting that the seven hydrophobic segments in GPCRs are also α-helical in nature (Findlay and Eliopoulos, 1990). Immunological mapping and protein digestion studies have provided supporting evidence for the topography of the receptors, consistent with seven α-helical regions, and have shown the C-terminal

Received for publication 7 September 1999 and in final form 5 September 2000.

Address reprint requests to Dr. B. A. Wallace, Department of Crystallography, Birkbeck College, University of London, Malet St., London WC1E 7HX, UK. Tel.: 44-(0)207-631-6857; Fax: 44-(0)207-631-6803; E-mail: ubcg91c@ccs.bbk.ac.uk.

© 2000 by the Biophysical Society

0006-3495/00/12/3083/12 $2.00

sequences to be intracellular and the N-terminal sequences to be extracellular (Dohlman et al., 1987). The N-terminal sequences are between 10 and several hundred amino acids in length in different members of the superfamily. The seven α-helical regions are connected by three sets of alternating intracellular and extracellular loops that are usually between 10 and 40 amino acids in length. The exception to this is the third intracellular loop, which in some members of the superfamily can be more than 150 amino acids in length. The intracellular C-terminal sequence can also vary in length from 50 to more than 150 amino acids.

Early models for GPCRs were based on the EM structure of bacteriorhodopsin (Findlay and Eliopoulos, 1990; Dahl et al., 1991; Grotzinger et al., 1991; Hibert et al., 1991; Ijzerman et al., 1992; Lewell, 1992; Cronet et al., 1993; Sylte et al., 1993; Yamamoto et al., 1993; Zhang and Weinstein, 1993). Most of these models assumed that the GPCRs had the same spatial arrangement for the seven helices and the six loops as bacteriorhodopsin. More recently it has been suggested that bacteriorhodopsin is not a suitable template for the construction of GPCR models because functionally bacteriorhodopsin is not a GPCR and because there is no overall significant sequence similarity between BR and the GPCRs (Attwood and Findlay, 1993). Pardo et al. (1992) suggested that the differences are largely due to exon shuffling, and that similarities between the helices in GPCRs and BR are observed if the sequential order of the helices is ignored. Alternatively, Taylor and Agarwal (1993) suggested that gene duplication occurred, which caused helices five, six, and seven to originate from helices one, two, and three. Because of the lack of one-to-one correspondence between the helices, it appears that the BR structure is not a suitable template for a GPCR model. Further evidence to this effect is found from the EM projection maps of bovine, frog, and squid rhodopsins (Schertler et al., 1993; Schertler and Hargrave, 1995; Davies et al., 1996; Krebs et al., 1998), which indicate that the arrangement of the helices in the rhodopsins is indeed different from that in bacteriorhodopsin.

Baldwin (1993) proposed two rules regarding the seven transmembrane helices, based on over 200 GPCR sequences and the EM structures. The first rule was that each helix must be positioned next to its neighbors in the sequence, and the second rule was that helices one, four, and five must be the ones most exposed to the lipid surrounding the receptor, with helix three being the least exposed. Donnelly et al. (1994) proposed a three-dimensional model of a GPCR based on Baldwin's rules and the EM projection map of bovine rhodopsin. This structure used as a basis the helical periodicity of amino acid substitution data (Donnelly et al., 1993). Other models have been produced using automated techniques based on detailed sequence analyses, multiple sequence alignments, and interactive graphics packages (Alkorta and Du, 1994; Taylor et al., 1994). Herzyk and Hubbard (1995, 1998) proposed a rule-based technique for packing helices in seven transmembrane helical proteins, using experimental and theoretical geometric constraint data. Pogozheva et al. (1997) produced a model that would maximize the number of interhelical hydrogen bonds for the entire GPCR superfamily.

Interestingly, although three recent models (Pogozheva et al., 1997; Baldwin et al., 1997; Herzyk and Hubbard, 1998) were produced by very different approaches, they exhibit three important similarities: 1) all have significant tilts in helices one, two, and three; 2) all model the third helix as being most deeply buried; and 3) the cytosolic end of helix three is inserted between helices four and five.

The diverse range of ligands that activate GPCRs appear to bind to their receptors in one of two ways: they either bind completely in the transmembrane region or they bind partly in this region and to part of the extracellular domain. For example, the adrenaline ligand appears to bind solely in the transmembrane region of the adrenoreceptor, probably because of its small size. Ligands of higher molecular weights (Fong et al., 1993), such as small peptides and glycolipid hormones, tend to interact with some of the transmembrane regions as well with an extracellular loop of the receptor. Within each G-protein-coupled receptor subfamily it is possible to detect conserved amino acids that have a role in agonist binding.

Rigid body docking algorithms for predicting protein-protein interactions tend to take into account shape complementarity between ligand and receptor, and hydrophobic effects from the change in the solvent-accessible surface (Cherfils et al., 1991), matching surfaces (Walls and Sternberg, 1992), or geometric complementarity (Norel et al., 1995). Electrostatic interactions are taken into account in some algorithms (Walls and Sternberg, 1992), while others employ Monte Carlo simulations to refine flexible side-chain positions (Totrov and Abagyan, 1994).

The aim of this study was to produce a model of the ET_A receptor and to use docking algorithms to examine the manner in which endothelin interacts with its receptor.

METHODS

The modeling described here utilizes a segmented approach where the C terminus, the N terminus, the transmembrane helices, and the loop regions were modeled on separate individual templates. The 1BOJ model for bovine rhodopsin (Pogozheva et al., 1997) was used as a template for the transmembrane segments. The COMPOSER (Srinivasan and Blundell, 1993) subroutine in SYBYL was used to identify sequence homologs whose structures were available in the Protein Data Base (PDB), as possible templates for the N-terminal (water-soluble) domain of the ET_A receptor. Homologs for the three extracellular loops were identified using a database of loop fragments (Jones and Thirup, 1986; Claessens et al., 1989) in SYBYL and the criterion of good geometric fit to the anchor region of the modeled protein. As a more stringent requirement, the loops identified also had to be located between two helices in the template structure. ClustalW (Thompson et al., 1994) was then used to determine which of the loop fragments that met this criterion had the highest homology with the ET_A loop sequences. PSI-BLAST (Altschul et al., 1997) was also used in an attempt to identify the loop regions; however, no hits were

found with an E-value greater than 0.01, probably because the sequences were too short. When looser constraints were used, the highest scoring sequences were not of loop regions. Hence, the SYBYL loop search results were used.

The sequence alignments were undertaken using the ClustalW site (located at http://www2.ebi.ac.uk/clustalw). The alignments of the ET_A and bovine rhodopsin transmembrane segments correspond to those identified in the GPCR data base (Horn *et al.,* 1998) (located at http://www.gpcr.org/7tm). The Blosum scoring matrix (Henikoff and Henikoff, 1992) was used for the BLAST and ClustalW alignments. In ClustalW the penalty value for closing or opening a gap was set at 10, and the penalty for extending a gap and separation of a gap was set at 0.05. The statistical threshold for reporting matches in BLAST was set at 10, and the corresponding cutoff scope for reporting high-scoring segment pairs was calculated from this.

The ends of the transmembrane (TM) segments were defined using the TMPred algorithm (Hofmann and Stoffel, 1993) (located at http://www.ch.embnet.org/software/TMPRED_form.html). Exceptions to predictions by this algorithm were that the starts of TM2, TM4, and TM6 and the end of TM3 were modified to mesh with the intracellular loop templates, and the end of TM6 and the start of TM7 were modified to correspond to the TM helices in the 1BOJ template.

Once a suitable alignment between the ET_A receptor sequence and the template sequence had been established, SYBYL V.6.4 was used to model the structure. Nonequivalent amino acids were mutated to produce the ET_A sequence. The loops were positioned with their N and C termini approximately equidistant from the two end residues of the helices to which they were to be connected (in all cases, the distance between the loop ends differed by $\leq$1 A from the distance between the target helices). A peptide bond was then formed with the "join chain" function in SYBYL, and the six residues surrounding the new bond were minimized. After all connections were made, the whole structure was energy-minimized using the Kollman (all-atom) minimization algorithm (Weiner et al., 1984, 1986).

The CONTACT algorithm (Collaborative Computational Project Number 4, 1994) was used to detect the interhelical hydrogen bonds, and MODEL (Driessen, unpublished observations) was used to determine the regions of close intermolecular approach of the polypeptide backbone atoms. The root mean square deviations were calculated using MOLMOL (Koradi et al., 1996). The modeled structure was validated with PROCHECK (Laskowski et al., 1993) and WHATIF (Vriend and Sander, 1993).

The FTDOCK (Gabb et al., 1997) algorithm was used to dock the receptor model with either the endothelin-1 crystal structure (Brookhaven Protein Data Bank ID, 1EDN) (Janes et al., 1994) or the bigET-1 model structure (Cronin and Wallace, 1999). The bigET-1 model structure used had been produced by a combination of homology modeling and threading, based on NMR and x-ray crystallographic data for ET-1 and a related toxin structure. FTDOCK was run with electrostatics on, a solvent accessibility of 10%, and a search over the complete binding space for both molecules. Negative FTDOCK scores indicate overlap/interpenetration of the ligand and receptor and are therefore not possible; a score of zero indicates that the ligand and receptor do not interact at all, and large positive scores denote complex formation with good surface complementarity. The NACCESS program (Hubbard and Thornton, 1993) was used to calculate the solvent accessibility of the modeled receptor and the ligands before and after docking.

RESULTS

Models for the seven transmembrane regions

The Brookhaven Protein Data Bank (PDB) currently contains two theoretical models of a GPCR. The two models are of bovine rhodopsin (entries 1BOJ and 1BOK, with *trans* and *cis* retinal ligands, respectively). These models had been constructed using an iterative distance geometry refinement with an evolving system of hydrogen bonds. Over 400 GPCRs were used to produce an average three-dimensional structure for the seven transmembrane helical region; in these models the numbers of buried polar side chains that formed hydrogen bonds were maximized for all of the members of the family simultaneously (Pogozheva et al., 1997). The resulting structure was consistent with Baldwin's for GPCR helix organization. The transretinal bovine rhodopsin model (PDB ID, 1BOJ) was used as a template for the seven transmembrane segments of the ET_A receptor. It has significant sequence identity in the transmembrane regions with the ET_A receptor. The sequence alignment of ET_A and bovine rhodopsin is shown in Fig. 1. The percentage identities (and similarity—identical plus conservative substitutions) for each of the seven transmembrane regions are TM1, 14% (50%); TM2, 23% (77%); TM3, 25% (67%); TM4, 18% (55%); TM5, 26% (63%); TM6, 29% (52%); TM7, 16% (52%).

Model for the N-terminal region

The ET_A N-terminal sequence of 80 amino acids was initially tested for sequence homologs of known structure, using a COMPOSER search. The two candidate template proteins with the highest sequence identity for the whole fragment (21%) were aligned with the ET_A sequence, using ClustalW, and the alignment with the smallest number of gap residues, a hydrolase (carboxylic esterase) (PDB ID, 1HPL) (Lombardo et al., 1989), was chosen as the template. Its alignment with the ET_A N terminus is shown in Fig. 2. The homologous part of the template protein consists of three long helices, the first and second of which run nearly antiparallel to each other. In addition there is a short single-turn helix between the first and second helices. It should be noted that the other candidate with 21% homology (PDB ID, 1OXY) also had a high helix content.

Models for the intracellular loop regions

The intracellular loop regions (Fig. 3 *A*) were modeled based on the structures of the intracellular loop regions of bovine rhodopsin, as determined by NMR spectroscopy (Yeagle et al., 1995b, 1997). The sequence identities (and similarities) between each of the intracellular loop regions of ET_A and bovine rhodopsin are as follows: loop 1, 24% (59%); loop 2, 14% (38%); loop 3, 13% (48%).

Model for the C-terminal region

The C terminus was modeled based on an NMR structure of the C terminus of bovine rhodopsin (Yeagle et al., 1995a). The alignment between the C terminus of ET_A and the C

```
TM1
SMLAAYMFLLIMLGFPINFLTL          ser38 to leu59      1boj
YINTVISCTIFIVGMVGNATLL          tyr81 to leu102     eta
 ...     ....*.  *   *

TM2
LNLAVADLFMVFGGFTTTLYTS          leu77  to ser98     1boj
ASLALGDLIYVVIDLPINVFKLL          ala120 to leu142    eta
 .**..**. *. ... ....

TM3
LEGFFATLGGEIALWSLVVLAIER         leu112 to arg135    1boj
LFPFLQKSSVGITVLNLCALSVDR         leu160 to arg183    eta
*  *. . .  *.. .* .*...*

TM4
VAFTWVMALACAAPPLVGWSRY          val157 to tyr178    1boj
IVSIWILSFILAIPEAIGFVMV          ile206 to val227    eta
..  *....  * *  .*.

TM5
IYMFVVHFIIPLIVIFFCY             ile205 to tyr223    1boj
WWLFGFYFCMPLVCTAIFYTLM           trp257 to met278    eta
 ..* ..* .**.   . *

TM6
MVIAFLICWLPYAGVAFYIFT           met257 to thr277    1boj
LVVIFALCWFPLHLSRILKKT           leu311 to thr331    eta
.*. * .**.*     .   *

TM7
PIFMTIPAFFAKTSAVYNPVIYIMM        pro285 to met309    1boj
LLMDYIGINLATMNSCINPIALYFV        leu348 to val372    eta
 ..  *   .*.  .  **.   ..
```

FIGURE 1 Sequence alignment of the seven transmembrane (TM) segments of bovine rhodopsin (1boj) with ET_A (eta). In all sequence alignment figures, an asterisk (*) indicates an identical amino acid; a "." indicates a "conserved" amino acid, which meets the criteria for either highly conservative substitutions or semiconservative substitutions, as defined by ClustalW, using the Gonnet Pam250 matrix.

terminus of human rhodopsin (Fig. 3 *B*) up to the site of palmitoylation (Cys^{387}) shows a sequence identity of 40%.

Models for the extracellular loop regions

The templates for each of the extracellular loops (EXLs) were identified according to the following criteria: 1) the sequence in the template protein had to be of a length comparable to that of the target loop, 2) the template structure had to be a loop structure, 3) the template structure had to form a loop between two antiparallel helical elements, 4) the distance between the ends of the loop in the template structure had to differ by no more than 1 Å from the distance between the TMS to be connected in the target protein, and 5) the sequence identity of the template protein with the target protein had to be ≥20%. For each loop only one potential template structure was identified in the PDB that met these stringent criteria. In addition, it was noted that, in general, the templates and the target sequences identities included "key residue" types (Martin and Thornton, 1996) such as glycines, prolines, cysteines, and charged and large hydrophobic residues. The templates used were as follows (Fig. 3 *C*): EXL1 was modeled on 3GLY, which is a hydrolase protein structure and has a sequence identity to the ET_A loop of 50% (and a similarity of 75%) (Aleshin et al., 1994). EXL2 was modeled on 1PNE, which is an actin-binding protein structure and has a sequence identity with the ET_A loop of 21% (similarity of 34%) (Cedergren-Zeppezauer et al., 1994). EXL3 was modeled on 1PII, which is an isomerase protein structure and has a sequence identity with the ET_A loop of 25% (similarity of 44%) (Wilmanns et al., 1992). All loop regions used were checked to confirm that the loops were well ordered (low B factors and full occupancies) in the template structure.

```
1HPL    LSTMC-QNMFKVESVNCICVDWKSGSRTAYS-QASQNVRIVGAEVAYLVGVLQ 137
ETA     METLCLRASFWLALVGCVISDNPERYSTNLSNHVDDFTTFRGTELSFLVTTHQ 53
        ..*.* .  * .  *.*.  *  .   *  * .... . . *.*...** . *

1HPL    SSFDYSPSNVHIIGHSLGSHAAGEAGRRT 167
ETA     PTNLVLPSNG--SMHNYCPQQTKITSAFK 80
        ..    ***     *.  .. .  .
```

FIGURE 2 Sequence alignment of the N terminus of ET_A with 1HPL. The three helical regions in the template and in the model are shown in bold.

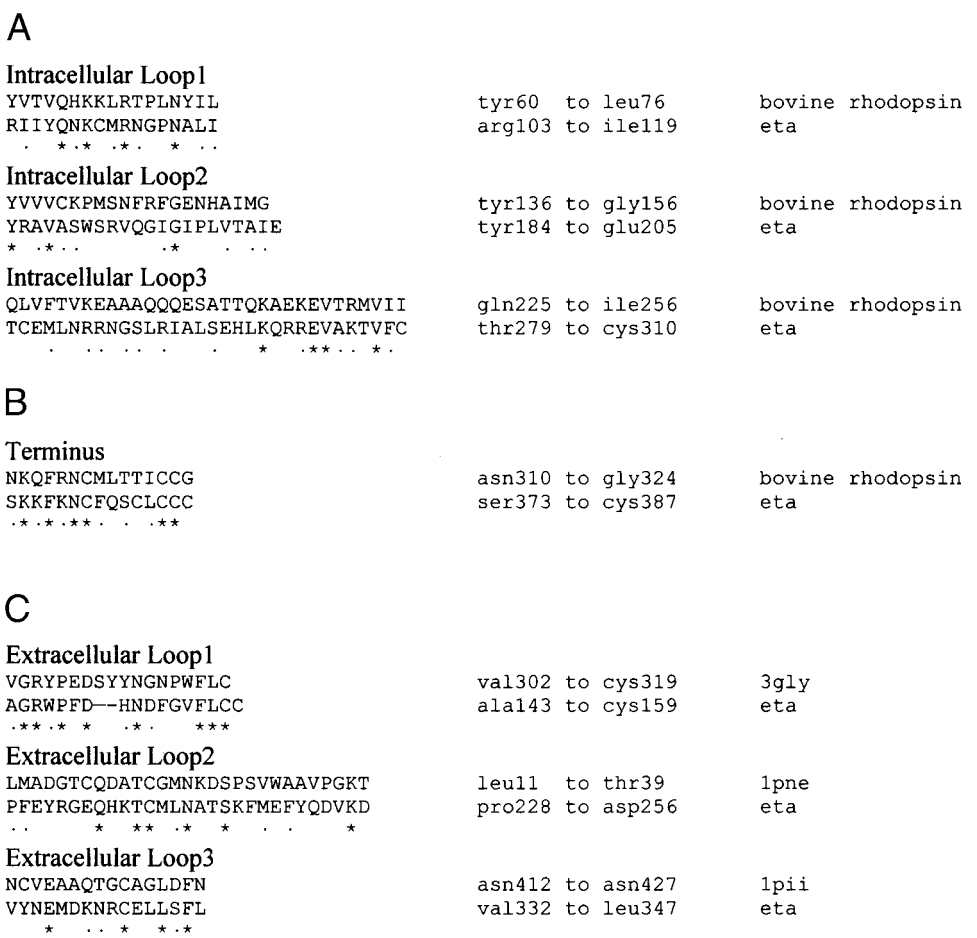

FIGURE 3 Sequence alignments of ET_A with (*A*) bovine rhodopsin intracellular loops, (*B*) C-terminus (up to Cys^{387}) and (*C*) loops of 3GLY, 1PNE, and 1PII.

Features of the modeled receptor structure

The structure of the modeled ET_A receptor is shown in Fig. 4. The extracellular N terminus contains three helices: Met^{1}-Phe^{10}; Leu^{30}-Thr^{55}; and Tyr^{68}-Ser^{77}, with a small, single-turn helix between residues 23 and 25. Preliminary NMR results by Yeagle and Albert (1999) on the isolated rhodopsin N terminus indicate that it also contains helical segments. Although the rhodopsin N terminus is only about half the length of the ET_A receptor N terminus, it also has one long helix near the beginning, followed by a short, single-turn helix. Hence some similar structural motifs may be present in the N termini of these two GPCRs.

The seven transmembrane helices were modeled in the following regions of the protein, using the sequence alignment shown in Fig. 1: TM1, Tyr^{81}-Leu^{102}; TM2, Ala^{120}-Leu^{142}; TM3, Leu^{160}-Arg^{183}; TM4, Ile^{206}-Val^{227}; TM5, Trp^{257}-Met^{278}; TM6, Leu^{311}-Thr^{331}; TM7, Leu^{348}-Val^{372}.

The fourth sixth and seventh transmembrane helices are approximately parallel to each other, whereas the other four helices are tilted. A proline kink is present at Pro^{267} in the fifth transmembrane helix (as it is in the corresponding part of the template protein). The root mean square deviations between backbone atoms in each of the TM segments of the rhodopsin template and the ET_A model ranged from 0.15 to 0.84 $Å^2$, with the largest variation being found in TM6. A preponderance of aromatic amino acids is located at or near the ends of many of the transmembrane segments (a feature previously noted in the crystal structures of many other membrane proteins; (Deisenhofer and Michel, 1989; Wallace and Janes, 1999)).

The intracellular C terminus was modeled as a compact globular structure. There is sequence conservation between the two rhodopsin cysteines (at positions 322 and 323) and the two ET_A cysteines (at positions 385 and 386). In rhodopsin these two cysteines can be acylated; however, in the NMR structure (Yeagle et al., 1995a), the C terminus was found to be a compact structure even in the absence of cysteine acylation. The last 40 amino acids of the C terminus of ET_A have not been modeled in this study because there was insufficient sequence homology to construct a model.

The three modeled cytoplasmic loops have β-turns in the middle, as in the rhodopsin templates on which they were modeled. The ends of all of the extracellular and intracellular loops were separated by distances that enabled them to be annealed to their corresponding helices.

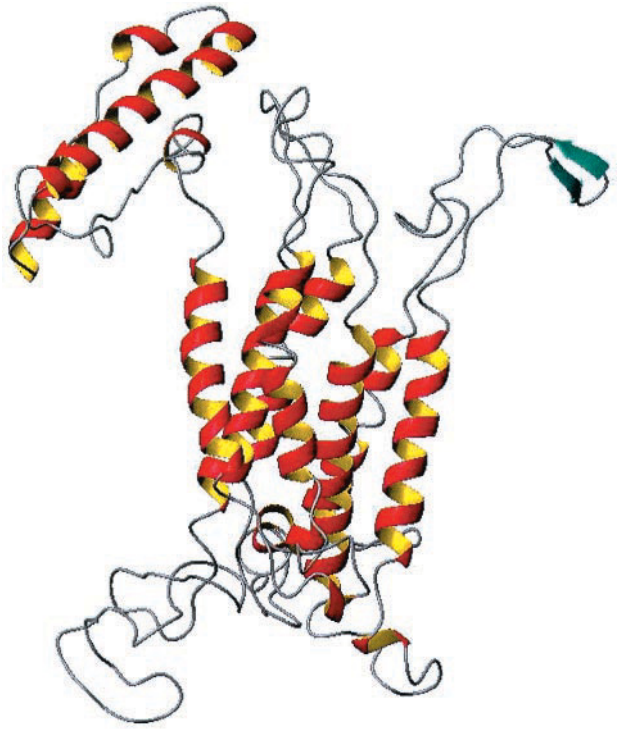

FIGURE 4 The modeled ET_A receptor, drawn as a ribbon diagram with MOLMOL (Koradi et al., 1996).

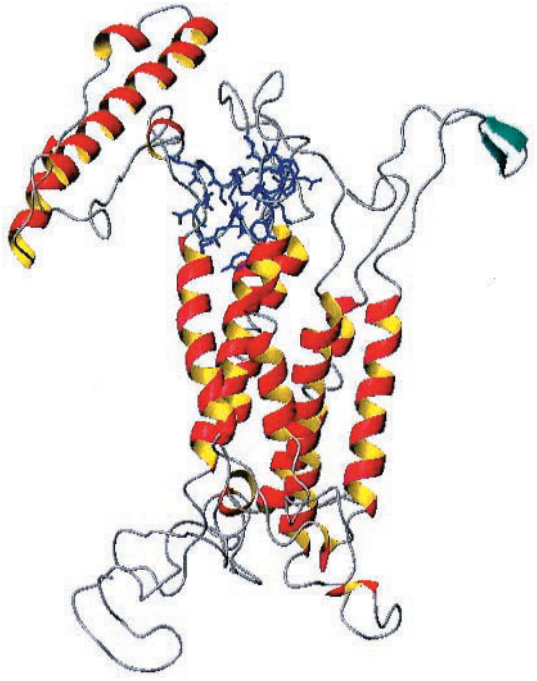

FIGURE 5 Predicted docking of the endothelin-1 crystal structure (drawn in *dark blue* in stick mode) to the modeled ET_A receptor (drawn as in Fig. 4).

The PROCHECK statistics showed that 87% of the residues in the ET_A model are in either the most favored or in the additionally allowed regions of the Ramachandran plot. The overall main-chain and side-chain parameters, as evaluated by PROCHECK are all very favorable. The WHATIF validation found loose rms Z-scores, which are typical of model, as opposed to experimental, structures.

Ligand docking

Changes in accessible surface area were used to demonstrate regions of binding. Although the magnitudes of the changes are susceptible to minor changes in the side-chain conformations, which are not reliably known for these modeled structures, this physical characteristic does clearly indicate which residues are in most intimate contact. Thus, as a conservative approach, we have only highlighted the residues that exhibit the largest changes in their surface areas as being regions of contact. Supporting this is the observation that similar groups of residues are indicated as being involved in the intermolecular interactions if other measurements of close approach (such as those involving intramolecular backbone distances of <5 Å) are used.

Fig. 5 shows the predicted docking of endothelin-1 into the modeled ET_A receptor. The highest positive correlation score structure ($= 176$) predicted by FTDOCK is the one shown. All other highly positively scored structures were also in this region. The change in accessible surface area of the ET_A receptor model with and without ET-1 docked is $\sim$1100 $Å^2$ or 4% of the total surface area, indicative of a substantial interaction. Furthermore, the majority of the changes are mapped (Tables 1–4) to the extracellular ends of the TM segments involved (1, 2, and 7), as well as EXL1, and to a lesser extent EXL2 and part of the N terminus. In the TM segments, only one face of each helix tends to be involved, thus defining their orientations. No interactions

TABLE 1 Accessible surface area of the first 10 residues on the extracellular side of TM1

Residue	Accessible surface area of free ET_A (A^2)	Accessible surface area of ET_A with bound ET-1 (A^2)
Tyr^{81}	121	117+
Ile^{82}	102	102
Asn^{83}	31	19+
Thr^{84}	35	35
Val^{85}	93	93
Ile^{86}	64	64
Ser^{87}	1	1
Cys^{88}	65	65
Thr^{89}	79	79
Ile^{90}	29	29

In Tables 1–5, * in the right-hand column denotes amino acids with the largest changes between the free receptor or ligand and the receptor complex. + indicates ones with smaller changes.

TABLE 2 Accessible surface area of the first 10 residues on the extracellular side of TM2

Residue	Accessible surface area of free ET_A (A^2)	Accessible surface area of ET_A with bound ET-1 (A^2)
Leu^{142}	144	43*
Leu^{141}	83	77*
Lys^{140}	40	37+
Phe^{139}	59	59
Val^{138}	95	95
Asn^{137}	5	5
Ile^{136}	12	12
Pro^{135}	37	37
Leu^{134}	66	66
Asp^{133}	0	0

were found with TM 3, 4, 5, or 6, any of the intracellular loops, or the C terminus.

Fig. 6 shows the highest correlation structure predicted for the docking of the endothelin precursor (bigET-1) and the ET_A receptor, which had a very small score of 28. That is, essentially no interaction between the receptor model and the modeled structure of bigET-1 is observed. There was no change in the accessible surface area of the modeled receptor when the proposed complex with bigET docked was analyzed. The "docked" structures with the next 20 highest scores also did not show binding to the receptor.

DISCUSSION

This study took a segmented approach to modeling a GPCR. Such a procedure is only justified if there is evidence that the individual segments in GPCRs can fold independently. For rhodopsin, it has been shown that the C-terminal region and the intracellular loops form compact structures that retain functional activity when examined as isolated polypeptides (Yeagle et al., 1995a,b, 1997). Furthermore, Martin et al. (1999) have recently shown that various TM segments of the α-factor receptor can fold and assemble into functional forms independent of the loop segments.

TABLE 3 Accessible surface area of the first 10 residues on the extracellular side of TM7

Residue	Accessible surface area of free ET_A (A^2)	Accessible surface area of ET_A with bound ET-1 (A^2)
Leu^{348}	93	10*
Leu^{349}	32	1*
Met^{350}	80	62+
Asp^{351}	47	47
Tyr^{352}	33	14+
Ile^{353}	88	38*
Gly^{354}	8	8
Ile^{355}	10	10
Asn^{356}	7	7
Leu^{357}	85	85

TABLE 4 Accessible surface area of EXL1 residues

Residue	Accessible surface area of free ET_A (A^2)	Accessible surface area of ET_A with bound ET-1 (A^2)
Ala^{143}	60	60
Gly^{144}	40	40
Arg^{145}	170	170
Trp^{146}	141	141
Pro^{147}	115	115
Phe^{148}	158	158
Asp^{149}	26	26
His^{150}	164	158+
Asn^{151}	102	101
Asp^{152}	112	110
Phe^{153}	142	40*
Gly^{154}	37	8*
Val^{155}	64	48+
Phe^{156}	205	118*
Leu^{157}	147	147
Cys^{158}	79	79

The TM segment template used in this study (Pogozheva et al., 1997) was constructed by a procedure that would produce the highest possible number of interhelical hydrogen bonds for the entire GPCR superfamily, although it did not necessarily maximize the number of hydrogen bonds in a individual member of the family. There are nine interhelical hydrogen bonds involving side chains between TM helices in the ET_A model structure; there were six such hydrogen bonds in the equivalent regions of the template structure. The model produced is consistent with preliminary circular dichroism data on refolded ET_A (Crawford and Wallace, unpublished results), which suggests that it has $<50\%$ helix content, in contrast to bacteriorhodopsin, which is $\geq 80\%$ helical (Wallace and Teeters, 1987). This lower overall helix content is obviously due to the larger nonhelical loops and termini found in the ET_A receptor.

The x-ray crystal structure of ET-1 (Janes et al., 1994) was used in the docking studies with the ET_A receptor. Structures of ET-1 and its analogs have also been determined by NMR spectroscopy (Endo et al., 1989; Saudek et al., 1989; Krystek et al., 1991; Munro et al., 1991; Reily and Dunbar, 1991; Tamaoki et al., 1991; Andersen et al., 1992). The NMR structures differ substantially, both from each other and from the x-ray structure (Wallace et al., 1995), especially in the all-important C terminus, a region that is disordered in many of the NMR structures. No NMR structures of ET-1 that include the C terminus have been deposited in the PDB, and hence none were available for docking studies. Justification for using the crystal structure, however, derives not from availability, but from the structure-activity relationship (SAR) studies that have been done on ET-1 (for a review, see Huggins et al., 1993), including alanine scans (Tam et al., 1994). When residues Glu^{10}, Phe^{14}, Asp^{18}, and Trp^{21} are mutated to almost any other amino acid, vasoactivity function (and hence receptor bind-

ing) is dramatically decreased. This pattern of significant substitutions separated by three or four residues is indicative of a helical secondary structure, which would place these amino acids on the same surface (Wallace and Janes, 1995). Such a helical secondary structure is found only in the x-ray crystal structure—hence its use in these studies. In the docked complex, the C terminus of the ligand shows a pattern of maximum changes in accessible surface area involving residues 10, 13, 14, 17, 18, and 21 (Table 5), the same sort of helical repeating pattern, which is nearly identical to the residues indicated to be the most important by SAR studies. N-terminal residues also clearly interact with the receptor, but this region appears to be associated with specificity of binding rather than binding affinity (Huggins et al., 1993).

Mutational studies of the ET_A receptor have suggested receptor residues that may be involved in ligand binding (for example, Adachi et al., 1994; Krystek et al., 1994; Breu et al., 1995; Rose et al., 1995; Webb et al., 1996). These are compiled in the tinyGRAP (Kristiansen et al., 1996; Edvardsen and Kristiansen, 1997) mutant database (Rel. 6.0). It is not expected that all of the mutations that cause binding effects will be at actual sites of contact in the complex. Many may be folding mutants that alter the receptor's overall conformation but not specifically its binding site. Indeed, many of the TinyGRAP sites in the TM segments appear to be buried and thus likely modulate the packing and ultimately the folding of the receptor itself. Furthermore, there is a very broad spatial distribution of the residues implicated in binding. Some clearly cannot involve the actual binding site, given their distance from each other and even their apparent locations on distal sides of the membrane. While the situation is complicated, those mutants that appear to specifically affect ET-1 binding (as opposed to other effects, such as antagonist binding or specificity for endothelin isoforms) seem to be concentrated in TM2, TM7, and EXL1. A number of the residues between 132 and 156 appear to be especially important for the binding function. In the receptor complex produced in these studies, as indicated by a decrease in the accessible surface areas of various

TABLE 5 Accessible surface area of the ET-1 ligand

Residue	Accessible surface area of free ET-1 (A^2)	Accessible surface area of ET-1 bound to ET_A (A^2)
Cys^1	114	84+
Ser^2	107	49*
Cys^3	75	34*
Ser^4	26	27
Ser^5	87	46*
Leu^6	164	13*
Met^7	116	59*
Asp^8	119	52*
Lys^9	170	95*
Glu^{10}	96	41*
Cys^{11}	4	4
Val^{12}	38	26+
Tyr^{13}	185	14*
Phe^{14}	90	31*
Cys^{15}	14	13
His^{16}	96	80+
Leu^{17}	95	23*
Asp^{18}	123	79*
Ile^{19}	122	124
Ile^{20}	101	74+
Trp^{21}	230	75*

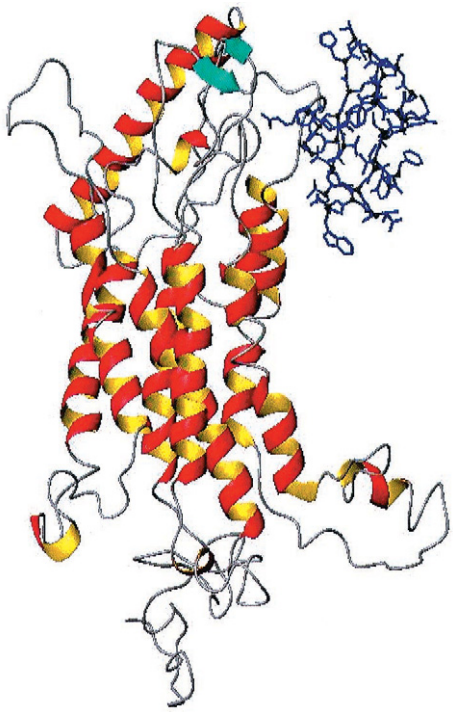

FIGURE 6 The "undocked" modeled bigET-1 structure with the ET_A receptor (drawn as in Fig. 5). The view in this figure is rotated about the *x* axis by ~70° relative to the views in Figs. 4 and 5, for the ligand "binding site" to be visible.

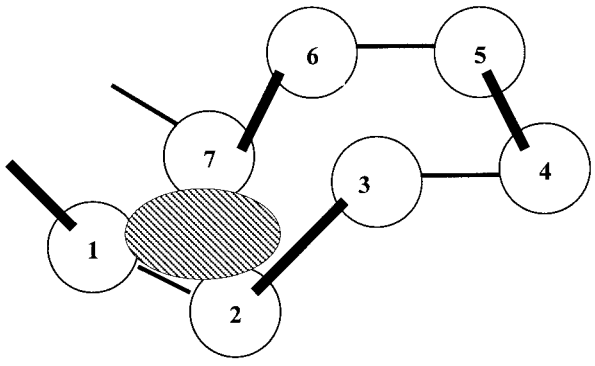

FIGURE 7 Schematic diagram showing where the ligand (indicated as a *hatched oval*) binds. This is an end-on view of the seven transmembrane helices of the receptor, with each helix drawn as a circle.

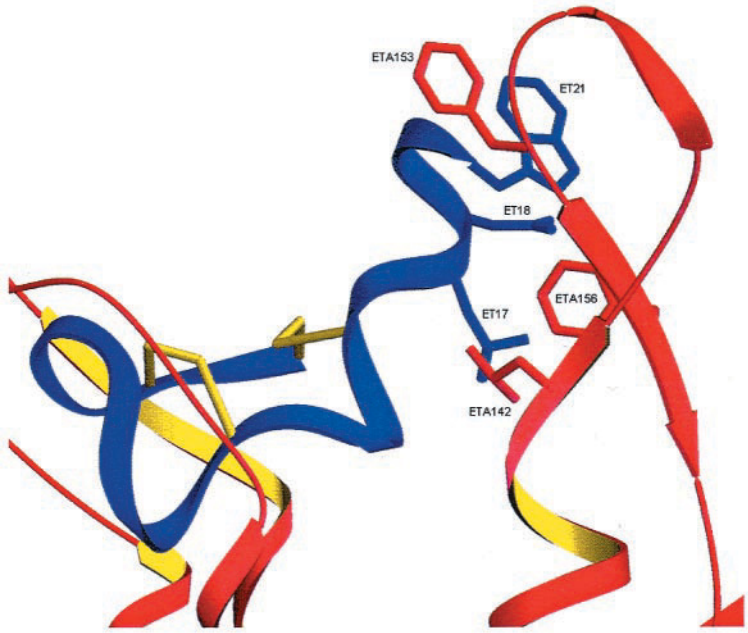

FIGURE 8 Detailed view showing the orientation of the ET-1 ligand (in *blue*) in the receptor (in *red*) binding site. The side chains of the ET_A and ET-1 residues that are involved in the closest contacts are shown in stick representation.

residues, the ligand docks principally at the extracellular ends of TM1, TM2, and TM7 of the receptor (Tables 1–3) as well as EXL1 (Table 4) and, to a lesser extent, EXL2 and part of the N terminus. Residues 140–156 are clearly involved in the binding interaction. The location of the binding site is shown schematically in Fig. 7, a view looking down onto the receptor from the extracellular side.

The proposed interactions of the ET-1 ligand in the ET_A receptor-binding pocket are shown in Fig. 8. The residues that are involved in the closest contacts and have been indicated to be important for ligand binding to the ET_A receptor (Huggins et al., 1993) are explicitly shown.

Fig. 6 shows the results obtained for the prediction of the docking of bigET and the ET_A receptor, as a control for the docking procedure used. Biochemical evidence (N′Diaye et al., 1997) indicates that bigET-1 exhibits negligible biological activity at the receptor, and hence it is believed not to be able to bind in the ligand-binding site. The in vivo pressor effect, an indirect measure of binding affinity, suggests 100-fold lower potency for bigET-1 relative to ET-1 (Kimura et al., 1989). Correspondingly, in this modeling study, essentially no interaction between the receptor and the modeled structure of bigET-1 was observed. The bigET structure is too bulky to interact with the receptor in the region occupied by ET-1 because of the size restraints caused by the extracellular loop regions of the ET_A receptor. Furthermore, the C-terminal precursor extension on bigET-1 (residues 22–38) tends to block a number of the helical residues that appear to be involved in the binding, namely residues 10, 17, 18, and 21.

This model now permits us to predict the binding of an endothelin analog, the precursor BigET(1–31). At present it has been speculated that this molecule might bind to the receptor, but there is no experimental evidence to support or refute this. Using the same type of procedure for docking as described above, it was found that this molecule could bind in approximately the same site as the mature ET-1, although the extent of interaction with ET_A was less (score = 63). This would predict binding, but perhaps with a lower affinity, a testable hypothesis.

In summary, now that we have a working model for the receptor-ligand complex, it should be possible to design biochemical and mutational studies to test the interactions proposed and to design other molecules that may fit in the binding pocket and act as agonists or antagonists.

Note added after submission of this paper

Since this paper was first submitted, the crystal structure of a GPCR, rhodopsin, was published (Palczewski et al., 2000). In general, the overall structures of that and the model in this work are similar, and the RMSDs between the backbone atoms of each of the TMS segments of that structure and our model were <2.0 Å. There are significant differences in the N termini. These parts of the molecules, however, are not expected to be very similar, because the length of the N terminus of rhodopsin is only ~1/2 of that of ET_A receptor, and there is little sequence homology between the two.

This work was supported by a grant from the British Heart Foundation. AO was the recipient of a Biotechnology and Biological Science Research Council studentship.

REFERENCES

Adachi, M., Y. Furuichi, and C. Miyamoto. 1994. Identification of a ligand-binding site of the human endothelin-A receptor and specific regions required for ligand selectivity. *Eur. J. Biochem.* 220:37–43.

Aleshin, A. E., C. L. Hoffman, M. Firsov, and R. B. Honzatko. 1994. Refined crystal structures of glucoamylase from *Aspergillus awamori* Var. X100. *J. Mol. Biol.* 238:575–591.

Alkorta, I., and P. Du. 1994. Sequence divergence analysis for the prediction of 7-helix membrane-protein structures. A 3-D model of human rhodopsin. *Protein Eng.* 7:1231–1238.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.

Andersen, N. H., C. P. Chen, T. M. Marschner, S. R. Krystek, and D. A. Bassolino. 1992. Conformational isomerism of endothelin in acidic aqueous-media—a quantitative NOESY analysis. *Biochemistry.* 31: 1280–1295.

Attwood, T. K., and J. B. C. Findlay. 1993. Design of a discriminating fingerprint for G-protein-coupled receptors. *Protein Eng.* 6:167–176.

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

Baldwin, J. M., G. F. X. Schertler, and V. M. Unger. 1997. An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors. *J. Mol. Biol.* 272:144–164.

Breu, V., K. Hashido, C. Broger, C. Miyamoto, Y. Furuichi, A. Hayes, B. Kalina, B. M. Loffler, H. Ramuz, and M. Clozel. 1995. Separable binding sites for the natural agonist endothelin-1 and the non-peptide antagonist bosentan on human endothelin-A receptors. *Eur. J. Biochem.* 231:266–270.

Cedergren-Zeppezauer, E. S., N. C. W. Goonesekere, M. D. Rozycki, J. C. Myslik, Z. Dauter, U. Lindberg, and C. E. Schutt. 1994. Crystallization and structure determination of bovine profilin at 2.0 angstroms resolution. *J. Mol. Biol.* 240:459–475.

Cherfils, J., S. Duquerroy, and J. Janin. 1991. Protein-protein recognition analyzed by docking simulation. *Proteins Struct. Funct. Genet.* 11: 271–280.

Claessens, M., E. Van Cutsem, I. Lasters, and S. Wodak. 1989. Modeling the polypeptide backbone with spare parts from known protein structures. *Protein Eng.* 2:335–345.

Cronet, P., C. Sander, and G. Vriend. 1993. Modeling of transmembrane 7 helix bundles. *Protein Eng.* 6:59–64.

Cronin, N. B., and B. A. Wallace. 1999. Do the structures of big ET-1 and big ET-3 adopt a similar overall fold? Consequences for endothelin converting enzyme specificity. *Biochemistry.* 38:1721–1726.

Collaborative Computational Project Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.* D50:760–763.

Dahl, S. G., O. Edvardsen, and I. Sylte. 1991. Molecular dynamics of dopamine at the D2 receptor. *Proc. Natl. Acad. Sci. USA.* 88: 8111–8115.

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

Deisenhofer, J., and H. Michel. 1989. Nobel lecture: The photosynthetic reaction centre from the purple bacterium *Rhodopseudomonas viridis. EMBO J.* 8:2149–2170.

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

Donnelly, D., J. B. C. Findlay, and T. L. Blundell. 1994. A model of the human visual pigment rhodopsin based upon the human beta2-adrenergic receptor. *Receptors Channels.* 2:61–78.

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

Edman, K., P. Nollert, A. Royant, H. Belrhali, E. Pebay-Peyroula, J. Hajdu, R. Neutze, and E. M. Landau. 1999. High resolution x-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature.* 401: 822–826.

Edvardsen, O., and K. Kristiansen. 1997. Computerization of mutant data: the tinyGRAP mutant database. *7TM J.* 6:1–6.

Endo, S., H. Inooka, Y. Ishibashi, C. Kitada, E. Mizuta, and M. Fujino. 1989. Solution conformation of endothelin determined by nuclear magnetic-resonance and distance geometry. *FEBS Lett.* 257:149–154.

Findlay, J., and E. Eliopoulos. 1990. Three-dimensional modelling of G protein-linked receptors. *Trends Pharmacol. Sci.* 12:492–499.

Fong, T. M., M. A. Cascieri, H. Yu, A. Bansal, C. Swain, and C. D. Strader. 1993. Amino aromatic interaction between histidine-197 of the neurokinin-1 receptor and CP-96345. *Nature.* 362:350–353.

Gabb, H. A., R. M. Jackson, and M. J. E. Sternberg. 1997. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J. Mol. Biol.* 272:106–120.

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

Grotzinger, J., M. Engels, E. Jacoby, A. Wollmer, and W. Strassburger. 1991. A model for the C5a receptor and for its interaction with the ligand. *Protein Eng.* 7:767–771.

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

Henikoff, S., and J. G. Henikoff. 1992. Amino-acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA.* 89:10915–10919.

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

Herzyk, P., and R. E. Hubbard. 1998. Combined biophysical and biochemical information confirms arrangement of transmembrane helices visible from the three-dimensional map of frog rhodopsin. *J. Mol. Biol.* 281: 741–754.

Hibert, M. F., S. Trumppkallmeyer, A. Bruinvels, and J. Hoflack. 1991. 3-dimensional models of neurotransmitter G-binding protein-coupled receptors. *Mol. Pharmacol.* 40:8–15.

Hofmann, K., and W. Stoffel. 1993. TMbase—a database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler.* 347:166.

Horn, F., J. Weare, M. W. Beukers, A. Bairoch, W. Chen, O. Edvardsen, F. Campagne, and G. Vriend. 1998. GPCRDB: an information system for G protein-coupled receptors. *Nucleic Acids Res.* 26:277–281.

Hubbard, S. J., and J. M. Thornton. 1993. NACCESS, Computer Program, Department of Biochemistry and Molecular Biology, University College London.

Huggins, J. P., J. T. Pelton, and R. C. Miller. 1993. The structure and specificity of endothelin receptors—their importance in physiology and medicine. *Pharmacol. Ther.* 59:55–123.

Ijzerman, A. P., J. Van Galen, and K. A. Jacobson. 1992. Molecular modeling of adenosine receptors. I. The ligand binding site on the A1 receptor. *Drug Des. Discov.* 9:49–67.

Janes, R. W., D. H. Peapus, and B. A. Wallace. 1994. The crystal structure of human endothelin. *Nature Struct. Biol.* 1:311–319.

Jones, T. A., and S. Thirup. 1986. Using known substructures in protein model-building and crystallography. *EMBO J.* 5:819–822.

Kimura, S., Y. Kasuya, T. Sawamura, O. Shinmi, Y. Sugita, M. Yanagisawa, K. Goto, and T. Masaki. 1989. Conversion of big endothelin-1 to 21-residue endothelin-1 is essential for expression of full vasoconstrictor activity: structure-activity relationships of big endothelin-1. *J. Cardiovasc. Pharmacol.* 13:S5–S7.

Kohno, M., K. Murakawa, T. Horio, K. Yokokawa, K. Yasunari, T. Fukui, and T. Takeda. 1991. Plasma-immunoreactive endothelin-1 in experimental malignant hypertension. *Hypertension.* 18:93–100.

Konvicka, K., F. Guarnieri, J. A. Ballesteros, and H. Weinstein. 1998. A proposed structure for transmembrane segment 7 of G protein-coupled receptors incorporating an Asn-Pro/Asp-Pro motif. *Biophys. J.* 75: 601–611.

Koradi, R., M. Billeter, and K. Wuthrich. 1996. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 14: 51–55.

Krebs, A., C. Villa, P. C. Edwards, and G. F. X. Schertler. 1998. Characterisation of an improved two dimensional $p22_12_1$ crystal from bovine rhodopsin. *J. Mol. Biol.* 282:991–1003.

Kristiansen, K., S. G. Dahl, and O. Edvardsen. 1996. A database of mutants and effects of site-directed mutagenesis experiments on G-protein coupled receptors. *Proteins Struct. Funct. Genet.* 26:81–94.

Krystek, S. R., D. A. Bassolino, J. Novotny, C. Chen, T. M. Marschner, and N. H. Andersen. 1991. Conformation of endothelin in aqueous ethylene glycol determined by H-1-NMR and molecular-dynamics simulations. *FEBS Lett.* 281:212–218.

Krystek, S. R., Jr., P. S. Patel, P. M. Rose, S. M. Fisher, B. K. Kienzle, D. A. Lach, E. C. Liu, J. S. Lynch, J. Novotny, and M. L. Webb. 1994. Mutation of the peptide binding site in transmembrane region of a G protein-coupled receptor accounts for endothelin receptor subtype selectivity. *J. Biol. Chem.* 269:12383–12386.

Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.

Lewell, X. Q. 1992. A model of the adrenergic beta-2 receptor and binding sites for agonist and antagonist. *Drug Des. Discov.* 9:29–48.

Lombardo, D., C. Chapus, Y. Bourne, and C. Cambillau. 1989. Crystallization and preliminary x-ray study of horse pancreatic lipase. *J. Mol. Biol.* 259–261.

Luecke, H., B. Schobert, H. T. Richter, J. P. Cartailler, and J. K. Lanyi. 1999. Structure of bacteriorhodopsin at 1.55 Angstrom resolution. *J. Mol. Biol.* 291:899–911.

Martin, A. C. R., and J. M. Thornton. 1996. Structural families in loops of homologous proteins: automatic classification, modelling, and application to antibodies. *J. Mol. Biol.* 263:800–815.

Martin, N. P., L. M. Leavitt, C. M. Sommers, and M. E. Dumont. 1999. Assembly of G protein-coupled receptors from fragments: identification of functional receptors with discontinuities in each of the loops connecting transmembrane segments. *Biochemistry.* 38:682–695.

Munro, S., D. Craik, C. McConville, J. Hall, M. Searle, W. Bicknell, D. Scanlon, and C. Chandler. 1991. Solution conformation of endothelin, a potent vaso-constricting bicyclic peptide—a combined use of H-1-NMR spectroscopy and distance geometry calculations. *FEBS Lett.* 278:9–13.

N′Diaye, N., M. E. Pueyo, T. Battle, C. Ossart, D. Guedin, and J. B. Michel. 1997. Conversion of big-endothelin-1 elicits an endothelin ETA receptor-mediated response in endothelial cell. *Eur. J. Pharmacol.* 321: 387–396.

Norel, R., S. L. Lin, H. L. Wolfson, and R. Nussinov. 1995. Molecular surface complementarity at protein/protein interfaces: the critical role played by surface normals at well placed sparse points in docking. *J. Mol. Biol.* 252:263–273.

Palczewski, K., T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. L. Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, and M. Miyano. 2000. Crystal structure of rhodopsin: a G protein-coupled receptor. *Science.* 289:739–745.

Pardo, L., J. A. Ballesteros, R. Osman, and H. Weinstein. 1992. On the use of the transmembrane domain of bacteriorhodopsin as a template for modeling the 3-dimensional structure of guanine nucleotide-binding regulatory protein-coupled receptors. *Proc. Natl. Acad. Sci. USA.* 89: 4009–4012.

Pogozheva, I. D., A. L. Lomize, and H. I. Mosberg. 1997. The transmembrane 7-alpha-bundle of rhodopsin: distance geometry calculations with hydrogen bonding constraints. *Biophys. J.* 72:1963–1985.

Reily, M. D., and J. B. Dunbar. 1991. The conformation of endothelin-1 in aqueous-solution—NMR-derived constraints combined with distance geometry and molecular- dynamics calculations. *Biochem. Biophys. Res. Commun.* 178:570–577.

Rose, P. M., S. R. Krystek, Jr., P. S. Patel, E. C. Liu, J. S. Lynch, D. A. Lach, S. M. Fisher, and M. L. Webb. 1995. Aspartate mutation distinguishes ET_A but not ET_B receptor subtype-selective ligand binding while abolishing phospholipase C activation in both receptors. *FEBS Lett.* 361:243–249.

Sakamoto, A., M. Yanagisawa, T. Sawamura, T. Enoki, T. Ohtani, T. Sakurai, K. Nakao, T. Toyoka, and T. Masaki. 1993. Distinct subdomains of human endothelin receptors determine their selectivity to endothelin A-selective antagonist and endothelin B-selective agonists. *J. Biol. Chem.* 268:8547–8553.

Saudek, V., J. Hoflack, and J. T. Pelton. 1989. H-1-NMR study of endothelin, sequence-specific assignment of the spectrum and a solution structure. *FEBS Lett.* 257:145–148.

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

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

Shibouta, Y., N. Suzuki, A. Shino, H. Matsumoto, Z. I. Terashita, K. Kondo, and K. Nishikawa. 1990. Pathophysiological role of endothelin in acute renal failure. *Life Sci.* 46:1611–1618.

Srinivasan, B. N., and T. L. Blundell. 1993. An evaluation of the performance of an automated procedure for comparative modelling of protein tertiary structure. *Protein Eng.* 6:501–502.

Sylte, I., O. Edvardsen, and S. G. Dahl. 1993. Molecular dynamics of the 5-HT1a receptor and ligands. *Protein Eng.* 6:691–700.

Tam, J. P., W. Liu, J.-W. Zhang, M. Galantino, F. Bertolero, C. Cristiani, F. Vaghi, and R. de Castiglione. 1994. Alanine scan of endothelin: importance of aromatic residues. *Peptides.* 15:703–708.

Tamaoki, H., Y. Kobayashi, S. Nishimura, T. Ohkubo, Y. Kyogoku, K. Nakajima, S. Kumagaye, T. Kimura, and S. Sakakibara. 1991. Solution conformation of endothelin determined by means of H-1-NMR spectroscopy and distance geometry calculations. *Protein Eng.* 4:509–518.

Taylor, E. W., and A. Agarwal. 1993. Sequence homology between bacteriorhodopsin and GPCRs—exon shuffling or evolution by duplication. *FEBS Lett.* 325:161–166.

Taylor, W. R., D. T. Jones, and N. M. Green. 1994. A method for alpha-helical integral membrane protein fold prediction. *Proteins.* 3:281–294.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.

Totrov, M., and R. Abagyan. 1994. Detailed ab initio prediction of lysozyme-antibody complex with 1.6 angstrom accuracy. *Nature Struct. Biol.* 1:259–263.

Toyooka, T., T. Aizawa, N. Suzuki, Y. Hirata, T. Miyauchi, W. S. Shin, M. Yanagisawa, T. Masaki, and T. Sugimoto. 1991. Increased plasma level of endothelin-1 and coronary spasm induction in patients with vasospastic angina pectoris. *Circulation.* 83:476–483.

Unger, V. M., P. A. Hargrave, J. M. Baldwin, and G. F. X. Schertler. 1997. Arrangement of rhodopsin transmembrane alpha-helices. *Nature.* 389: 203–206.

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

Van Rhee, A. M., B. Fischer, P. J. M. Van Galen, and K. A. Jacobson.1995. Modelling the P2y purinoceptor using rhodopsin template. *Drug Des. Discov.* 13:133–137.

Vriend, G., and C. Sander. 1993. Quality-control of protein models—directional atomic contact analysis. *J. Appl. Crystallogr.* 26:47–60.

Wallace, B. A., and R. W. Janes. 1995. The crystal structure of human endothelin-1 and how it relates to receptor binding. *J. Cardiovasc. Pharmacol.* 26:S250–S253.

Wallace, B. A., and R. W. Janes. 1999. Tryptophans in membrane proteins: x-ray crystallographic analyses. *Adv. Exp. Med. Biol.* 467:789–799.

Wallace, B. A., R. W. Janes, D. Bassolino, and S. Krystek. 1995. Comparison of x-ray and NMR structures of human endothelin-1. *Protein Sci.* 4:75–83.

Wallace, B. A., and C. L. Teeters. 1987. Differential absorption flattening optical effects are significant in the circular dichroism spectra of large membrane fragments. *Biochemistry.* 26:65–70.

Walls, P. H., and M. J. E. Sternberg. 1992. New algorithm to model protein recognition based on surface complementarity—applications to antibody antigen docking. *J. Mol. Biol.* 228:277–297.

Warner, T. D., B. Battistini, A. M. Doherty, and R. Corder. 1994. Endothelin receptor antagonists: actions and rationale for their development. *Biochem. Pharmacol.* 48:625–635.

Webb, M. L., P. S. Patel, P. M. Rose, E. C. Liu, P. D. Stein, J. Barrish, D. A. Lach, T. Stouch, S. M. Fisher, O. Hadjilambris, H. Lee, S. Skwish, K. E. Dickinson, and S. R. Krystek, Jr. 1996. Mutational analysis of the endothelin type A receptor (ET_A): interactions and model of selective ETA antagonist BMS-182874 with putative ETA receptor binding cavity. *Biochemistry.* 35:2548–2556.

Weiner, S. J., P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. 1984. A new force-field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106:765–784.

Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. 1986. An all atom force-field for simulations of proteins and nucleic acids. *J. Comp. Chem.* 7:230–252.

Wilmanns, M., J. P. Priestle, T. Niermann, and J. N. Jansonius. 1992. Three-dimensional structure of the bifunctional enzyme phosphoribosylanthranilate isomerase: indoleglycerolphosphate synthase from *Escherichia coli* refined at 2.0 angstroms resolution. *J. Mol. Biol.* 223: 477–507.

Yamamoto, Y., K. Kamiya, and S. Terao. 1993. Modeling of human thromboxane A2 receptor and analysis of the receptor-ligand interaction. *J. Med. Chem.* 36:820–825.

Yanagisawa, M., H. Kurihara, S. Kimura, Y. Tomobe, M. Kobayashi, Y. Mitsui, Y. Yazaki, K. Goto, and T. Masaki. 1988. A novel potent vasoconstrictor peptide produced by vascular endothelial cells. *Nature.* 332:411–415.

Yeagle, P. L., and A. D. Albert. 1999. The three dimensional structure of the G protein receptor rhodopsin: a domain approach. *Biophys. J.* 76: a277.

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

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

Yeagle, P. L., J. L. Alderfer, A. C. Salloum, L. Ali, and A. D. Albert. 1997. The first and second cytoplasmic loops of the g-protein receptor, rhodopsin, independently form beta-turns. *Biochemistry.* 36:3864–3869.

Zhang, D. Q., and H. Weinstein. 1993. Signal transduction by a 5-HT2 receptor—a mechanistic hypothesis from molecular-dynamics simulations of the 3-dimensional model of the receptor complexed to ligands. *J. Med. Chem.* 36:934–938.