ELSEVIER

Modeling of the three-dimensional structure of the human melanocortin 1 receptor, using an automated method and docking of a rigid cyclic melanocyte-stimulating hormone core peptide

Peteris Prusis,* Helgi B. Schiöth,* Ruta Muceniece,* Pawel Herzyk,† Mohammad Afshar,† Roderick E. Hubbard,† and Jarl E. S. Wikberg*

**Pharmaceutical Pharmacology, Uppsala University, Uppsala, Sweden*
†Department of Chemistry, University of York, Heslington, York, UK

A model is presented of the melanocortin 1 receptor (MC1R), constructed by use of an unbiased, objective method. The model is created directly from data derived from multiple sequence analysis, a low-resolution EM-projection map of rhodopsin, and the approximate membrane thickness. The model agrees well with available data concerning natural mutations of MC1Rs occurring in different species. A model is also presented of the most rigid ligand for this receptor, the cyclic pentapeptide cHFRWG, shown docked in the receptor model. The receptor–ligand complex model agrees well with available experimental data. The ligand is located between transmembrane region 1 (TM1), TM2, TM3, TM6, and TM7 of the receptor. Multiple interactions occur between ligand and receptor, including interactions with Leu-48 (TM1), Ser-52 (TM1), Glu-55 (TM1), Asn-91 (TM2), Glu-94 (TM2), Thr-95 (TM2) Ile-98 (TM2), Asp-121 (TM3), Thr-124 (TM3), Phe-257 (TM6), Phe-283 (TM7), Asn-290 (TM7), and Asp-294 (TM7) of the receptor. © *1998 by Elsevier Science Inc.*

The Color Plate for this article is on page 334.

Address reprint requests to: Jarl Wikberg, Farm Farm, Box 591, BMC, S-751 24 Uppsala, Sweden; e-mail: ⟨Jarl.Wikberg@farmbio.uu.se⟩.

Received 30 October 1997; revised 16 January 1998; accepted 19 January 1998.

Keywords: melanocortin receptor, G-protein coupled receptor, melanocyte stimulating hormone, 3D structure

INTRODUCTION

The melanocortin 1 receptor (MC1R) is a member of a family consisting of five subtypes of melanocortin receptor (MC1R–MC5R) that have been cloned.[1–5] Of these receptors, the MC2R is the adrenocorticotropic hormone (ACTH) receptor; the others are receptors for melanocyte-stimulating hormones (MSHs). The MC1R has been localized to the melanocytes of the skin, where it has a role in control of skin pigmentation, but it is also found in the periaqueductal grey area of the brain,[6] implying other functions as well. The MC2R is located in the adrenal cortex,[7] where it may respond to ACTH in the control of corticosteroid secretion. The other three melanocortin receptors are found in the central nervous system and in the periphery.[3–5,8] The exact functional roles of these receptors are presently under investigation, and the rapidly accumulating data indicate that they may be involved in control of many diverse processes such as eating, motor behaviour, body temperature, pain, and inflammation.

Knowledge about the three-dimensional (3D) structure of the melanocortin receptors is crucial for understanding the mechanisms underlying their functionality and would be useful for drug design. Melanocortin receptors belong to the class of

Journal of Molecular Graphics and Modelling 15, 307–317, 1997

655 Avenue of the Americas, New York, NY 10010

1093-3263/97/$17.00
PII S1093-3263(98)00004-7

membrane-bound proteins termed *G protein-coupled receptors* (GPCRs), which are all built from a single polypeptide chain and that all are presumed to have seven transmembrane (TM) regions with α-helical conformation (here termed TM1–TM7). These proteins have proved difficult to solubilise and crystallise and this has hindered the determination of their 3D structure using experimental methods such as crystallography and nuclear magnetic resonance (NMR).

We have previously proposed a model for the human MC1R, generated by homology modeling using bacteriorhodopsin as a template.[9] However, bacteriorhodopsin is not a G protein-coupled receptor and it does not show significant sequence homology with GPCRs, casting doubt on the validity of bacteriorhodopsin-based models. In addition, the docking of a ligand in our previous model was heavily influenced by the interpretation of some early mutagenesis data indicating that Asp-117 and His-260 of the MC1R could be located in the ligand-binding pocket of the receptor and participate directly in the binding of the MSH peptides.[10] More recent evidence indicates that these specific amino acids are unlikely to interact directly with the MSH peptides.[11]

In this study we propose a new model of the transmembrane domain of the human MC1R in complex with the ligand, cyclic pentapeptide cHFRWG.[12] To arrange the receptor TM helices into a bundle, we used the previously described automated computer methodology that takes into account both experimental and theoretical information as guiding forces in the modelling procedure.[13] Such information is derived from the 2D electron density projection map at 9-Å resolution available for another member of the GPCR superfamily, bovine rhodopsin,[14] as well as restraints derived from sequence analysis of the family of melanocortin (MC) receptors. The final model of the ligand–receptor complex, generated by use of a conformational search technique, shows good agreement with available experimental data.

METHODS

Software

Sequence analysis was performed with the GCG package.[15] All other calculations used the PANDA,[13] XPLOR,[16] and CHARMM[17] software. Display and examination of templates and refined models were performed with the programs SQUID[18] and QUANTA.[19] Figures showing receptor, ligand and receptor–ligand complex models were created using the program MOLSCRIPT.[20]

Summary of computational strategy

The computational strategy used in this work consists of several stages:

Sequence analysis The most conserved and variable residues in the family of MC receptors were identified using multiple sequence alignment. This information was then used to generate orientational restraints used in the receptor modeling procedure.

Receptor modeling A Monte Carlo simulated annealing (MCSA) procedure was used to generate an initial model template for the seven transmembrane helices of the MC1 receptor, using restraints derived from multiple sequence alignment and the 2D projection map for bovine rhodopsin.[14] The initial template was then refined by a molecular dynamics simulated annealing (MDSA) procedure to give a full-atom representation.

Ligand modeling A model of the cyclic pentapeptide, cHFRWG, was generated using a MDSA procedure.

Ligand–receptor complex modeling The ligand model was docked manually into the receptor template in order to identify key interactions between the receptor and the ligand. They were converted into protein–ligand distance restraints and used to generate and refine a model of the complex.

Sequence analysis

Sequences from all different species and subtypes of MC receptors found in the EMBL database[21] were considered. MC2 receptor sequences were excluded, because we have earlier shown that the MC2R is pharmacologically distinct from the other four MC receptor subtypes, as it binds only ACTH peptides but not MSH peptides.[22] Partial receptor sequences were also excluded from the analysis. The PILEUP program of the GCG package[15] was used to produce a multiple sequence alignment of the remaining sequences, which were then analyzed to identify conserved and varying residues. A residue is considered to be conserved if at a given position at least 90% of the amino acids in the alignment are identical. In addition, if the amino acids are identical within each group of receptor subtypes, but vary between the subtypes, the amino acid is also considered to be conserved. This ensures that sequence variability is not confused with family-specific sequence conservation. A residue is considered to be varying if at a given position four or more amino acids are different among all the receptors included in the alignment. A gap is considered a 'different' amino acid.

Template generation

The method used for generating templates for seven-helix receptors has been described in detail elsewhere and implemented in the computer program PANDA.[13] In summary, it involves aggregation of rigid, idealized helices into a bundle by optimisation of a function penalizing violations of experimental and theoretical restraints as well as steric overlaps. A simplified representation of the protein is used so that each residue in the helix is represented by one C_α atom and one virtual side-chain C_β atom, whose size and position depend on the size and topology of the side chain. In each helix three dummy atoms represent the helix centroid DC, extracellular end DE, and intracellular end DI. These can be used as anchor points for certain restraints but do not take part in the steric overlap calculation.

Throughout this study three kinds of structural restraints are used: *xy*-positional, *z*-positional, and orientational restraints. *xy-positional restraints* are imposed on *x* and *y* coordinates of DC, DE, and DI atoms in order to restrain their position with respect to the membrane plane (*xy* plane). They are inferred from the 2D projection map for bovine rhodopsin.[14] *z-position restraints* control the positions of helix ends with respect to boundaries of the membrane. These are deduced from the approximate thickness of the membrane.

Orientational restraints are used to restrain certain residues to be either pointing inside (IOR) or outside (OOR) the helix

bundle. Here, inside is defined as facing both the receptor pocket and interhelical space. In this context outside means 'not inside', which is equivalent to facing the lipid phase. These restraints are inferred from the analysis of the multiple sequence alignment of melanocortin receptors. In this analysis the first and the last four amino acids of each TM were ignored, as these might interact with the membrane surface, making their orientation difficult to define. For each orientational restraint a scaling constant was applied.

Inside orientational restraints were applied to the conserved amino acids. As several transmembrane regions show a high number of conserved residues, only those residues with side chains capable of forming hydrogen bonds were selected for further analysis. For the selected residues, an initial ORI score was assigned so that the charged amino acids (arginine, aspartate, glutamate, histidine, and lysine) were given a score of 2 and the rest (asparagine, cysteine, glutamine, serine, threonine, tryptophan, and tyrosine) were given a score of 1. The amino acids classified as varying were assigned the outside orientational restraints (OORs). The initial OOR score was assigned so that if four different amino acids are present at a given position in the alignment the initial OOR score is set to 1, and for five and six different amino acids the score is set to 2 and 3, respectively.

The penalty function, which measures how well the structural restraints are satisfied, is constructed as a sum of contributions from violations of each restraint, as described earlier.[13] The scaling of the inside orientational restraints for conserved residues was achieved by applying Eq. (1):

$$K_a = \frac{I_a \cdot 0.5}{\sum_i I_i} \quad (1)$$

where K_a is the scaling constant for conserved residue a, I_a is the initial IOR score for this residue, and $\Sigma_i I_i$ is the sum of all initial IOR scores over all conserved residues in a particular helix. Consequently, the sum of all scaling constants for IOR in each helix is 0.5 $Å^{-2}$.

Scaling of the outside orientational restraints was obtained in a manner similar to that used for the IOR, by applying Eq. (2):

$$K_a = \frac{O_a \cdot 0.5}{\sum_i O_i} \quad (2)$$

where K_a is the scaling constant for a varying residue a, O_a is the initial OOR score for this residue, and $\Sigma_i O_i$ the sum of the initial OOR scores over all varying residues in a particular helix. Consequently, the sum of all scaling constants for OOR in each helix is 0.5 $Å^{-2}$ and is equal to the sum of scaling constants for the IOR.

Global optimisation of the penalty function in the configurational space of 42 degrees of freedom (three rotations and three translations per helix) is achieved by the MCSA technique.[23] A number of MCSA trajectories are calculated, starting with different randomised initial configurations. The final configurations are assessed by their penalty function value and a few with higher values are discarded. The mean configuration is generated by averaging the accepted configurations and the final template is created by superposition of the idealised helices on the mean structure.[13]

Refinement of the receptor template

The receptor template generated within the first stage is further refined using a molecular dynamics simulated annealing (MDSA) protocol in order to (1) extend it into a full-atom representation, and (2) optimise the geometry of the system. The MDSA calculations, performed using the program XPLOR,[16] are similar to the protocol of Nilges and Brünger.[24] The refinement procedure has been described;[25] by its use the transmembrane domain of the ion-channel phospholamban was modeled from comprehensive mutagenesis data.

Throughout the MDSA calculation the protein is represented with a united-atom and explicit polar hydrogen topology and parameter sets TOPH19 and PARAM19.[17,26] All bond, angle, and improper energy terms have uniform constants of 500 kcal/mol · $Å^2$, 500 kcal/mol · rad^2, and 200 kcal/mol · rad^2, respectively, and all planes and chiral centres are maintained using improper dihedrals, in the same way as they are defined for NMR structure determination.[16] The charges on helix N and C termini were neutralised.

Three groups of distance restraints are used in the MDSA calculations, using a uniform energy constant of 2.5 kcal/mol · $Å^2$. *Intrahelical distance restraints* are imposed on intrahelical hydrogen bond lengths. The distance between O(k) and N(k + 4) is kept between 2.4 and 3.2 Å and the distance between O(k) and HN(k + 4) is kept between 1.7 and 2.3 Å. If the position k + 4 is occupied by proline these restraints are not applied and restraints between residues k + 1 and k + 5 are modified so that the distance between O(k + 1) and N(k + 5) is kept between 3.0 and 5.8 Å, whereas the distance between O(k + 1) and H(k + 5) is kept between 2.1 and 4.9 Å.[27] *Interhelical distance restraints* are imposed between the ends and also between the centroids of neighbouring helices and every second of the helices. The centroid positions are the mean coordinate of the middle seven C_α atoms in each helix, whereas the terminal positions are the mean coordinate of the first seven or last seven C_α atoms in each helix. The target distances are taken from the receptor template and the tolerances are set to 1.5 Å. *Side-chain distance restraints* are imposed between the centroids of the side chains and the position representing the virtual C_β atoms from the template. The upper limits of these restraints are different for different side chains and reflect the evaluation of an error of the approximation of the side-chain centroid position by the reduced representation used in the receptor template calculations.

The MDSA trajectories are calculated by starting from different random starting points for the side-chain atom positions. Each initial structure consists of the C_α atoms whose coordinates are taken from the receptor template, with the coordinates of all other atoms randomly placed within a sphere of 1 Å diameter centered on the appropriate C_α atom. This structure is then subjected to 10.3 ps (1 fs time step) of restrained molecular dynamics simulation at 1 000 K temperature, keeping C_α atom positions fixed, as described.[24] This protocol consists of several stages at which different potential energy terms are introduced one by one (first bond and angle, then impropers, then repulsive nonbonded potential). The energy constants are initially scaled down and then increased gradually up to their maximum values. The intrahelical and side-chain restraints are used at this stage of calculation. The system is then annealed to 300 K for 2.9 ps by reducing the temperature of the heat bath by 25 K every 0.1 ps, followed by energy minimisation. During

this stage C_α atoms are no longer fixed but harmonically restrained to their initial positions, with an initial energy constant of 50 kcal/mol · Å^2 reduced every 0.1 ps by a factor of 0.8. Consequently, the side-chain restraints are no longer used and the Lennard–Jones and electrostatic potentials are used instead of simplified nonbonded potentials. A switching function is applied to both potentials between 8 and 12 Å, and a nonbonded neighbour list generated within a radius of 13 Å. A distance-dependent dielectric constant ($\varepsilon ij = rij$, in angstroms) is used. During this stage both intrahelical and interhelical restraints are imposed on the system. Final structures with unusually high potential energy values are discarded and the remaining structures averaged to produce the mean structure. This structure is then minimised, using no restraints.

Ligand modeling

To find the most probable model for the peptide ligand cHFRWG, MDSA calculations were carried out on a number of ligand structures generated by randomising all ϕ and ψ backbone torsion angles. After a short, initial energy minimisation, 5 ps of molecular dynamics simulation was performed at 1 000 K, using a time integration step of 0.5 fs. The peptide was then annealed to 300 K by reducing the temperature of the heat bath by 10 K every 0.5 ps, followed by energy minimisation with the Powell minimiser for 10 000 steps. Initial random structures were generated using the CHARMM program,[17] whereas the rest of the calculations were performed using the XPLOR program.[16] The peptide was represented with united-atom and explicit polar hydrogen topology and parameter sets TOPH19 and PARAM19.[17,26] Throughout these calculations the nonbonded potential was treated in the same way as during MDSA refinement of the receptor model.

Ligand–receptor complex modeling

Modeling of the ligand–receptor complex involved two main stages. The first stage involved visual analysis of the receptor model, to find an approximate location and orientation on the receptor surface where the ligand could make appropriate interactions. This crude docking was performed manually by positioning ligand C_α atoms against the receptor template, minimising the steric overlap between ligand and receptor. From this crude model a number of key interactions between the receptor and ligand could be identified, which were added to the restraint list used in subsequent calculations. The same MDSA refinement procedure was then used with all atoms for the amino acids of the ligand randomly generated from their C_α atoms, which were not fixed during the simulation. A smaller time integration step of 0.2 fs was used during dynamics and simulated annealing. Restraints between ligand and receptor were applied during the initial stages of the refinement protocol, but these were later released during annealing. Final structures with high potential energy values were discarded, and the remaining structures were averaged to produce a mean structure. This structure was then minimised using no restraints.

RESULTS

Sequence analysis

The multiple sequence alignment of the 13 melanocortin receptor sequences is shown in Figure 1. The relative numbers of conserved and varying residues for the TM regions, as well as for all the extracellular loops (EL1–EL3) and the N terminus (NT), were calculated. These results are summarized in Figure 2. The analysis showed that the least conserved regions are NT, EL1, and EL2. For the TM regions the least conserved are TM4 and TM5. Furthermore, when the varying residues are considered, the most varied regions are NT, EL1, EL2, and TM4. It seems probable that these least conserved regions do not play an important role in receptor function and ligand binding. The most conserved regions are TM3, TM6, and TM7, which also show the smallest number of varying amino acids. This suggests that these regions are important in receptor structure and might be involved in the direct binding of the MSH peptides. TM1 and TM2 also show a high degree of conservation, although these regions also show a large number of varying amino acids. For the extracellular loops only EL3, which is located between TM6 and TM7, shows an appreciable degree of conservation and a low level of amino acid variation.

Receptor modeling

In this study the TM helix indices were taken from the Baldwin assignment,[28] with some minor adjustments. The most important deviation from the Baldwin prediction was an extension of the length of the second extracellular loop (EL2) in the MC1 receptor, from one to five amino acids. This change is necessary, as the length of EL2 would otherwise be too short for the transmembrane segments to join (Figure 1).

To explore how well the method for selection of orientationally restrained amino acids worked, a helical wheel plot was made for each TM region (Figure 3). As can be seen in Figure 3, the rule-based selection of orientational restraints places most of the 'inside' residues on the opposite side of the 'outside' residues. The only exception was TM3, for which inside residues are almost randomly scattered around this helix, consistent with this helix being at the centre of the TM helical bundle. Moreover, the only varying residue (Ile-123) of this region is facing the opposite side of a single charged conserved residue (Asp-121). The higher scaling constant achieved for these residues by our method would mainly determine the orientation of the third transmembrane region.

The receptor template model was generated with a combination of the orientational restraints with the *xy*-positional and *z*-positional restraints previously described.[13] Templates were generated for 50 initial structures. Penalty function values for the 50 final templates ranged from 17.6 to 26.3. The 3 structures with the highest values (25.0–26.3) were discarded, and the remaining 47 were averaged to generate the final template. The convergence measured by average RMSD from the mean structure over the 47 selected structures is 2.7 Å. This value is caused predominantly by the uncertainty of helix *z* coordinates.

The final receptor model was then obtained by refinement of the template. Twenty-five MDSA trajectories were calculated from the receptor template with differently randomised non-C_α atoms. The 23 final molecules having negative energy were averaged. Here a high convergence of 0.7 Å is achieved on main-chain atoms. The energy of the final receptor model after minimization was −2159.6 kcal/mol, and its RMSD from the initial template is 1.1 Å on C_α atoms.

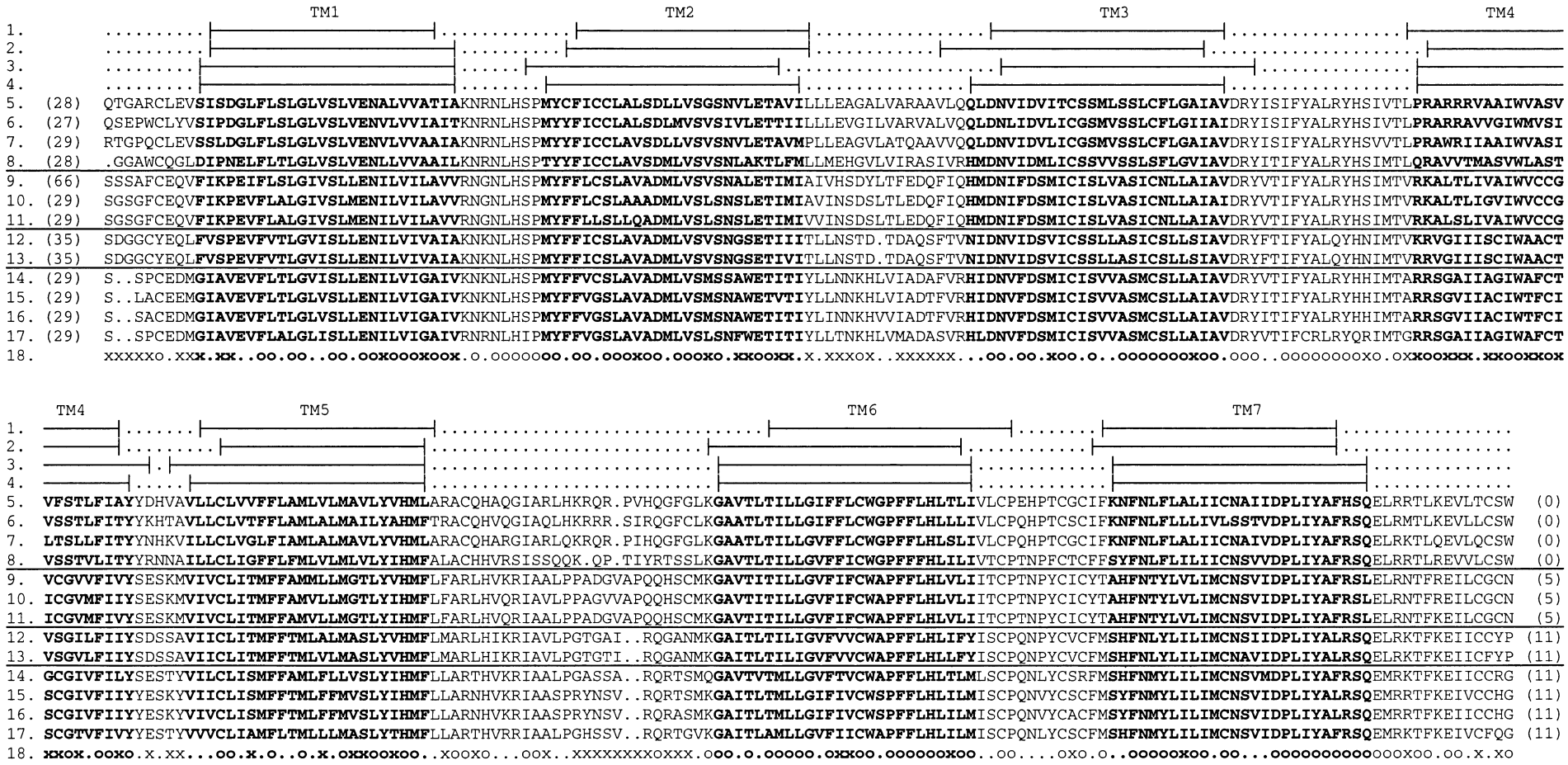

Figure 1. Multiple sequence alignment of melanocortin receptors and localisation of transmembrane regions. The bars in lines 1–4 show the transmembrane segments as assigned by (1) Chhajlani and Wikberg[1], (2) Prusis et al.[9], and (3) Baldwin[28], and (4) the assignment used for the present model. Sequences used in the analysis of conserved and varying amino acids are shown in lines 5–17 and are as follows (indicated within parentheses are the EMBL access codes): line 5, MC1 human (x67594); line 6, MC1 mouse (x65635); line 7, MC1 cattle (s71017); line 8, MC1 chicken (d78272); line 9, MC3 human (l06155); line 10, MC3 mouse (x74983); line 11, MC3 rat (x70667); line 12, MC4 human (l08603); line 13, MC4 rat (u67863); line 14, MC5 human (z25470); line 15, MC5 mouse (x76295); line 16, MC5 rat (l27081); line 17, MC5 ovary (z31369). In line 18 are indicated the conserved (○) and varying (×) residues found according to the analysis of the present study (see text for details).

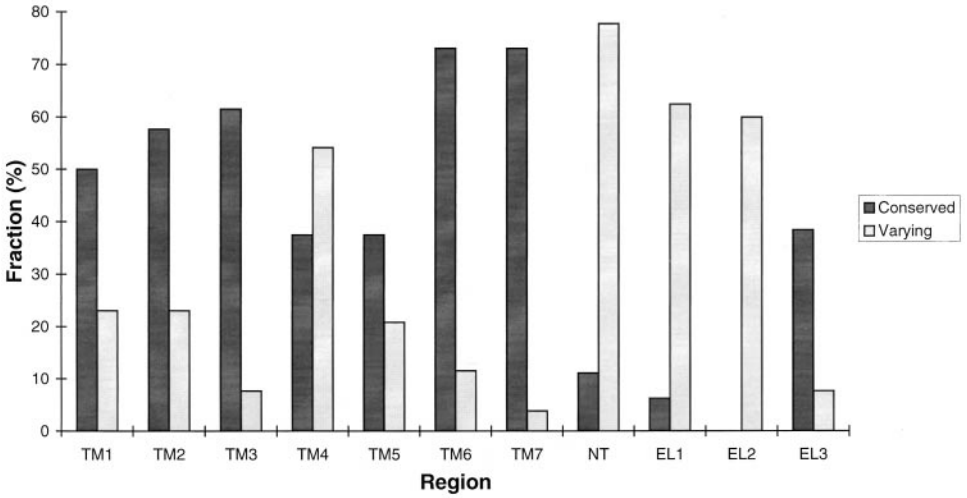

Figure 2. Proportion of conserved and varying residues in different regions of the MC receptor, calculated as described in Figure 1 and in further detail in text. Indicated are the transmembrane regions (TM1–7), the N terminus (NT), and the three extracellular loops (EL1–3).

Ligand modeling

The natural ligands for the melanocortin receptors are peptides more than 10 amino acids in length (Figure 4). A number of linear and cyclic MSH peptide analogs have been synthesised, most of which contain the core sequence HFRW, which appears to be a prerequisite for high-affinity binding. Unfortunately, the flexibility of these peptides makes their rational use in modelling difficult. In our earlier modeling study we used the smallest cyclic MSH analog known at that time (Figure 4), which is a peptide that still may show substantial flexibility. More recently, however, we described a novel series of small cyclic MSH peptides.[12] The smallest cyclic peptide from this series that was found to bind to the MC1 receptor was cHFRWG, and we therefore selected this peptide for our docking attempts in the present study.

One hundred initial randomised models of the peptide were generated, which after minimisation gave molecules with energies ranging from −42.44 to −71.08 kcal/mol. Figure 5 shows an overlay of the resulting molecules, which show an average RMSD from the mean structure of 0.7 and 2.8 Å on C_α and all atoms, respectively; these results show that the backbone of the cyclic peptide is quite rigid, while the side chains are rather flexible.

Thus, these calculations validated the use of the rigid backbone of the peptide for the initial orientation in the docking process. The more flexible side chains may of course change their conformation in order to achieve better interactions with the receptor. For the docking calculations, the structure with the lowest energy relative to all others in the set of 100 molecules was chosen.

Ligand–receptor complex modeling

Visual analysis of the receptor model suggested that a cluster of negatively charged and hydrophilic amino acids is located in the region between Glu-55 and Glu-94 (Figure 6). In fact, this region included not only four negatively charged residues (Glu-55, Glu-94, Asp-84, and Asp-294), but also several polar hydrophilic residues (Asn-56, Asn-91, and Asn-290). No other region in the receptor showed such richness of negatively charged and polar residues. This suggested that the most positively charged amino acid of the ligand (arginine) might bind there. Further analysis of the pharmacophore of the receptor model and positions of some conserved residues indicated that the tryptophan of the ligand might point toward Asp-121 (TM3) of the receptor and that the phenylalanine of the ligand points toward Phe-45 (TM1) of the receptor. Thus, the approximate initial docking of the C_α backbone of the ligand into a receptor template was able to achieve the close contacts between the C_α atoms of the arginine, tryptophan, and phenylalanine of the ligand and the receptor Glu-55/Glu-94 centroid, Asp-121, and Phe-45, respectively. The appropriate three restraints were kept throughout the initial high-temperature stage of the subsequent MDSA protocol. Twenty-five MDSA trajectories were calculated and 18 final structures selected on the basis of low potential energy value. These structures converged well and their average RMSD from the mean structure was 0.9 Å on main-chain atoms. The same value for ligand main-chain atoms only, while the structures are overlapped on the receptor main chain, is 3.0 Å, which shows that the final ligand position has substantial flexibility. The lower convergence seen for the ligand was to some extent expected, as the ligand C_α atoms were free to move throughout the calculation. However, the final ligand geometry did not change much, as the C_α backbone differs only by 0.5 Å of the RMSD from the mean ligand structure calculated during the ligand model generation stage. The energy of the final ligand–receptor complex model was −2677.1 kcal/mol (some 500 kcal/mol lower than the energy of the receptor without the ligand), and its RMSD from the initial ligand–receptor template is 2.1 Å on C_α atoms. The energy of the final model is lower than the energy of each of the 18 final MDSA structures after additional minimisation, which justifies the structure averaging.

In Color Plate 1 is shown an overview of the final model of the ligand–receptor complex. The ligand is located mainly between TM1, TM2, and TM7, although the tryptophan of the

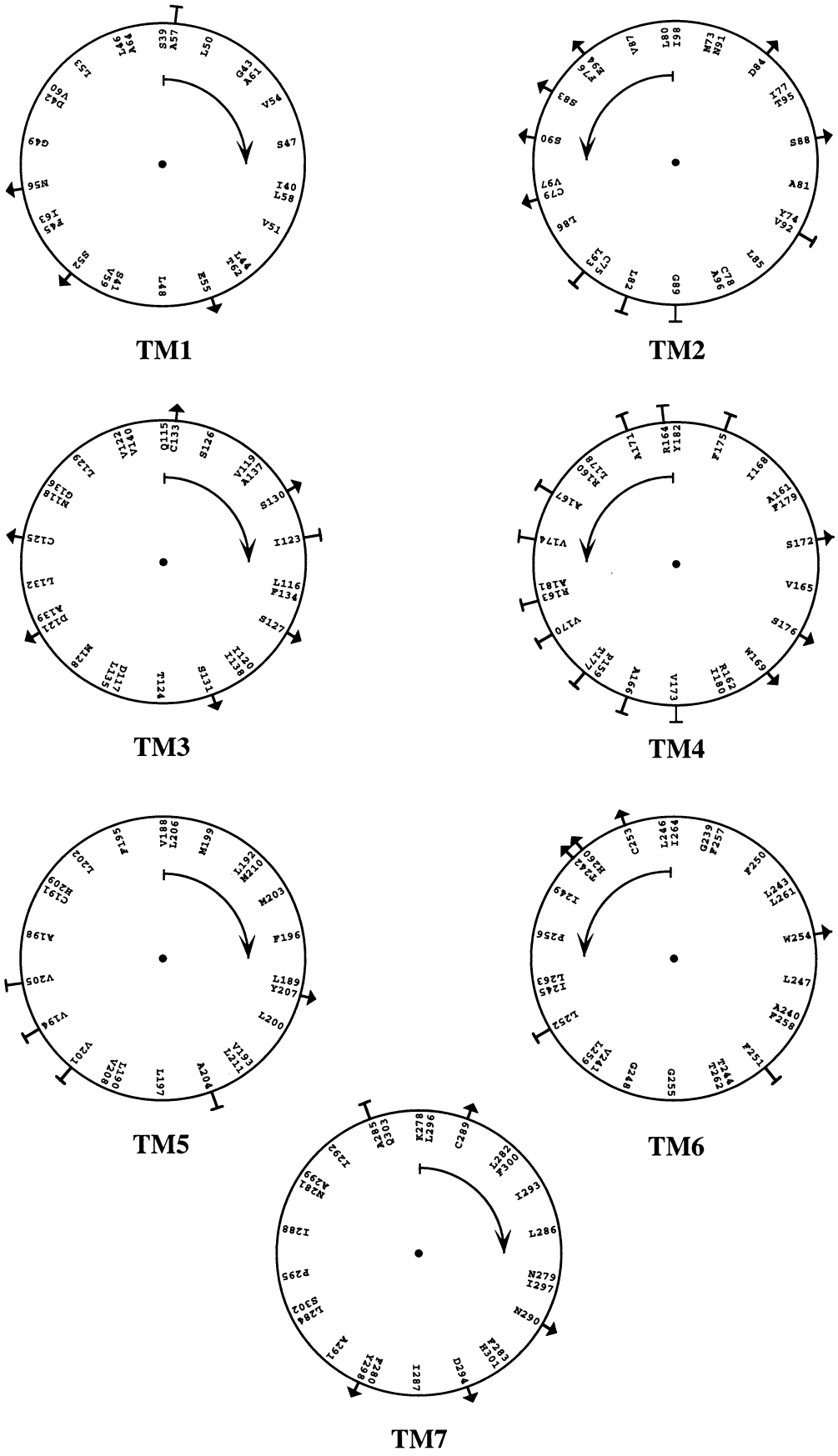

Figure 3. Helical wheels of all transmembrane regions of the MC1 receptor. Helices are viewed from the extracellular side. The first amino acid from the extracellular side is pointing upward. The amino acids selected for inside orientational restraints are marked with arrows and amino acids selected for outside orientational restraints are marked with bars.

α-MSH: Ser-Tyr-Ser-Met-Glu-**His- Phe-Arg-Trp**-Gly-Lys-Pro-Val

β-MSH: Ala-Glu-Lys-Lys-Asp-Glu-Gly-Pro-Tyr-Arg-Met-Glu-**His- Phe-Arg-Trp**-Gly-Ser-Pro-Pro-Lys-Asp

γ-MSH: Tyr-Val-Met-Gly-**His- Phe-Arg-Trp**-Asp-Arg-Phe

(4-10)D: Cys-Glu-**His-dPhe-Arg-Trp**-Cys (S—S)

cHFRWG: **His-dPhe-Arg-Trp**-Gly

Figure 4. Some ligands for melanocortin receptors. α-MSH, β-MSH, and γ-MSH are naturally occurring ligands for melanocortin receptors, (4–10)D is a peptide used in our previous studies[9], and cHFRWG is the peptide used in the present study.

ligand extends in the direction of TM3 and TM6. The phenylalanine of the ligand is located between TM1 and TM2. Interestingly, the arginine and phenylalanine of the ligand moved slightly away from their initial manually positioned interaction points, the arginine and phenylalanine being shifted approximately 4.4 and 8.4 Å deeper into the receptor, respectively. Moreover, the phenylalanine of the ligand does not show any interaction with Phe-45 of the receptor as it was tried initially (see above). Figure 7 summarises the interactions between the ligand and receptor. The arginine of the ligand is located within a network of hydrophilic and charged residues that includes Ser-52, Glu-55, Asn-290, and Asp-294. Tryptophan forms a possible hydrogen bond with Asp-121. Moreover, a number of hydrophobic interactions between tryptophan and Thr-124, Phe-257, Phe-283, and Leu-286 in the receptor are present. Phenylalanine is found in a cavity between TM1 and TM2, with some interactions with Leu-48, Thr-95, and Ile-98 being possible. The histidine of the ligand is pointing in the extracellular direction and interacts with Glu-94 and Ile-98. Asn-91 may interact with the backbone of the ligand.

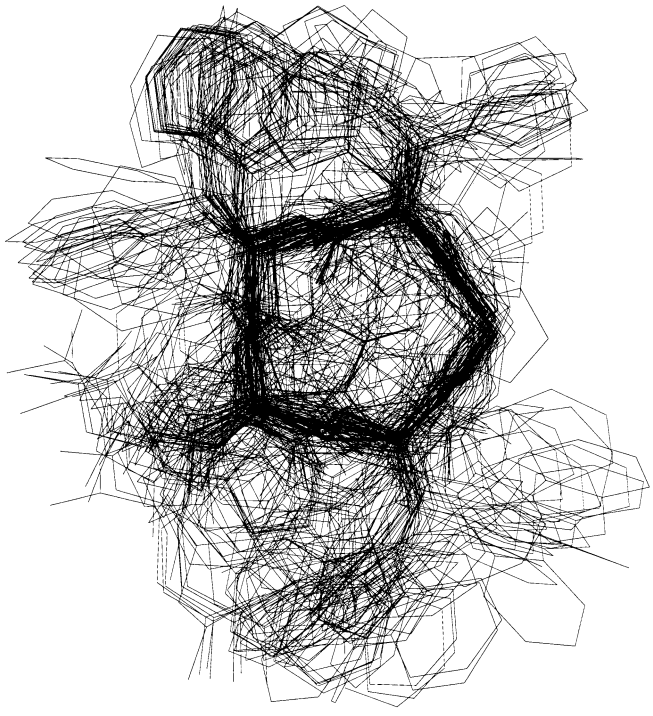

Figure 5. Alignment of 100 models of the cHFRWG peptide by RMS fit of the C_α atoms. [Figure made using the program MOLSCRIPT.[20]]

DISCUSSION

We have constructed a model of the MC1 receptor, using an unbiased, objective method and the rhodopsin footprint as a

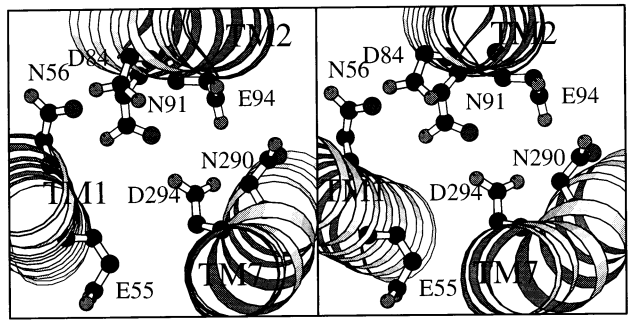

Figure 6. Wall-eyed stereoplot of the cluster of negatively charged and polar amino acids enclosing the point between Glu-55 and Glu-94 that was used for initial orientation of the ligand during docking. The view is from the intracellular side. [Figure made using the program MOLSCRIPT.[20]]

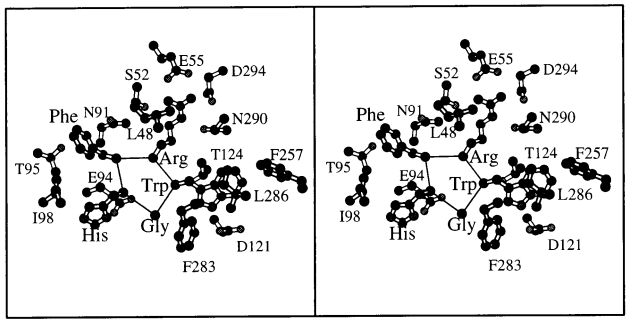

Figure 7. Wall-eyed stereoplot showing interactions between ligand and receptor in the model. Shown are the interacting amino acids and the trace of the ligand backbone. Amino acids of the ligand are indicated using three-letter notation, whereas amino acids of the receptor are given in single-letter notation. [Figure was made using the program MOLSCRIPT[20].]

template, and investigated how this model can be used to interpret the observed binding specificity of this receptor. The main source of information for the orientation of helices in the receptor model came from the analysis of the multiple sequence alignment of the subtypes of the cloned melanocortin receptors in different species and relied therefore on minimal prior information of the structural organization of the melanocortin receptors. The rationale behind using data from the sequences of all melanocortin receptor subtypes cloned in different species is that they are believed to share a similar structural organization. This seems highly probable as the MC1, MC3, MC4, and MC5 receptors bind the same ligands, albeit with some differences in affinities.[29,30] It seems reasonable to assume that highly conserved amino acids within the different MCRs are oriented inside the receptor pocket. However, there are a number of amino acids that are highly conserved only within each subtype among the different species. These amino acids potentially play important roles in the function of the receptor and were used to direct the modelling of the MC1 receptor. The selection of orientationally restrained amino acids was also supported by the helical wheel analysis. Moreover, the normalisation of orientational constraints allowed a balanced use of the information extracted from the sequence analysis. The present approach should also be applicable for other GPCR subfamilies, as well as for other clusters of membrane proteins with α-helical organization such as ion channels.

Allelic variants have been found for the MC1 receptor in mice, humans, cow, and fox.[31–36] The variation in MC1 receptor primary structure is associated with distinct variation in colour of hair, skin, and/or coat. In many cases these mutations seem to affect the activation mechanisms of the receptor. For example, the receptor may become constitutively activated and/or show reduced responsiveness to MSH peptides. The locations of natural mutations in the transmembrane regions of the MC1 receptor are shown in Figure 8 (excluding frameshift mutations). Although we did not take advantage of the information concerning these natural mutants in the modelling of the MC1 receptor, it is interesting to note that in our model virtually all of them occur in amino acids facing the receptor interior, which is to be expected if they cause changes in

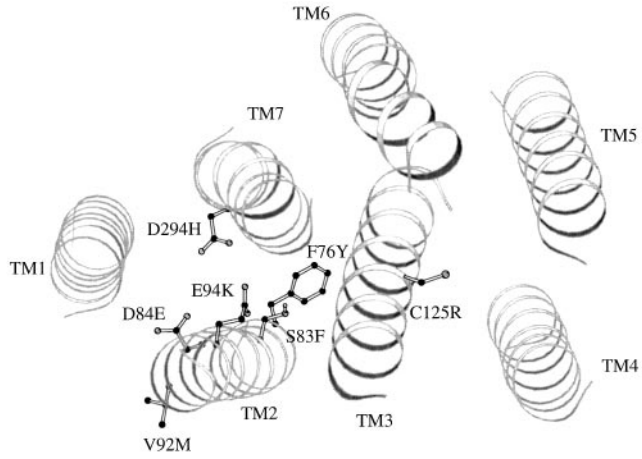

Figure 8. Natural mutations in transmembrane regions of MC1 receptors of different species. Amino acids are given in single-letter notation, along with sequence number according to the human MC1 receptor sequence, followed by the letter code for the mutation. Transmembrane regions of the receptor are shown using the C_α atom trace. The natural mutants were from references as follows: F76Y[33], S83F[34], D84E[33,35], V92M[33,35], E94K[31], C125R[36], and D294H.[33] [Figure was made using the program MOLSCRIPT[20].]

receptor functionality. However, an exception is the Val92Met mutation facing in the direction of the lipid bilayer. Still, this topology was expected because Val-92 is located on the opposite side of Glu-94, the latter having been found, when naturally mutated to lysine, to form a constitutive active receptor.[31] Moreover, we have found that mutating the highly conserved amino acid in the human MC3R (corresponding to Glu-94 in the MC1R) to either lysine or arginine leads to complete loss of ^{125}I-labeled NDP–MSH ligand binding.[37] In contrast, the Val92Met mutation of the MC1R is reported to cause only a minor effect on α-MSH binding[38] or MSH peptide activation of the receptor,[35] in support of our model.

Figure 9 shows 15 amino acids whose positions are conserved in more than 75% of G protein-coupled receptors.[21] All of them except for Pro-256, Pro-295, and Trp-169 face the bundle interior. However, conservation of prolines does not

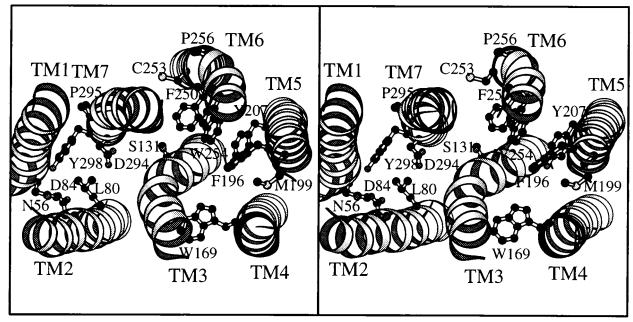

Figure 9. Wall-eyed stereoplot showing positions of the 15 most conserved amino acids found in GPCRs as viewed from the extracellular side. [Figure was made using the program MOLSCRIPT[20].]

necessarily imply their inside orientation, as they may be important for local perturbation of helical structure. Furthermore, although Trp-169 is technically oriented outside, its position suggests a possibility of involvement in interaction between TM4, TM3, and TM2. Thus, on the whole, the orientation of TM helices in the final MC1 receptor model is consistent with the conservation pattern in GPCRs.

Our analysis indicates that the receptor-binding pocket is located mainly between TM1, TM2, and TM7. This location of the ligand is supported by several experimental observations. According to the model, Glu-94 interacts with the backbone of the docked ligand and (as mentioned above) Glu94Lys and Glu94Arg mutations have great influence on ligand binding and receptor function. Moreover, in the model the arginine of the ligand interacts with Asp-294 of the receptor. The natural mutation of Asp-294 to histidine is found in individuals with red hair and skin that tans poorly, which is compatible with a failure or reduction in the ability of the receptor to become activated by the MSH hormone.[33] According to our model the change of the negatively charge aspartate to histidine is of course expected to influence MSH binding and receptor function.

Support for our ligand–receptor complex model comes also from the observations that alterations in TM4 or TM5 do not affect ligand binding. Thus, as we have shown earlier, the mutations Phe179Ala (TM4) and His209Ala (TM5) do not have an effect on either ^{125}I-labeled NDP-MSH[‡] or α-MSH ligand binding in the MC1 receptor.[10] Moreover, we have more recently shown that the exchange of TM4, TM5, and EL2 in the MC3 receptor with the corresponding parts of the MC1 receptor does not affect ligand binding.[39] Thus, assuming that the ligand-binding epitopes in the MC1 and MC3 receptors are located at similar positions, these data indicate that the ligand-binding pocket is not located in this region. Finally, the Cys125Arg mutation in TM3 (Figure 8), which is associated with constitutive activation of the receptor, is shown not to affect ligand binding,[36] in support of our model.

As mentioned above, by use of site-directed mutagenesis we have already presented evidence that the amino acids Asp-117 (TM3) and His-260 (TM6) are involved in the structural organisation of the receptor.[11] This evidence was based on the testing of a number of linear and cyclic MSH analogs on Asp117Ala and His260Ala mutants, as well as a mutant in which both these amino acids had been simultaneously changed to alanine. In our present model Asp-117 and His-260 are not directly interacting with the ligand. Instead, Asp-117 interacts with Asn-118 (TM3) whereas His-260 interacts with Asn-281 (TM7). Thus the subtle changes in the affinities for some ligands caused by mutations of Asp-117 and His-260[10,11] are in line with our model and are likely induced by slight changes in the overall organisation of the MC1 receptor.

ACKNOWLEDGMENTS

We thank Peter Gadd and Nils-Einar Eriksson at the BMC computing department for their invaluable advice in setting up the computer software and hardware that we used during our investigations. Support for this study was received from the Swedish MRC (04X-09575), BBSRC, British Council, CFN, Royal Swedish Academy of Sciences, and the Åke Wiberg and Groschinsky foundations.

[‡][^{125}I][Nle^4, D-Phe^7]αMSH

REFERENCES

1 Chhajlani, V., and Wikberg, J.E.S. Molecular cloning and expression of the human melanocyte stimulation hormone receptor cDNA. *FEBS Lett.* 1992, **309,** 417–420

2 Mountjoy, K.G., Robbins, L.S., Mortrud, M.T., and Cone, R.D. The cloning of a family of genes that encode the melanocortin receptors. *Science.* 1992, **257,** 1248–1251

3 Gantz, I., Konda, Y., Tashiro, T., Shimoto, Y., Miwa, H., Munzert, G., Watson, S.J., DelValle, J., and Yamada, T. Molecular cloning of a novel melanocortin receptor. *J. Biol. Chem.* 1993, **268,** 8246–8250

4 Gantz, I., Miwa, H., Konda, Y., Shimoto, Y., Tashiro, T., Watson, S.J., DelValle, J., and Yamada, T. Molecular cloning, expression, and gene localization of a fourth melanocortin receptor. *J. Biol. Chem.* 1993, **268,** 15174–15179

5 Chhajlani, V., Muceniece, R., and Wikberg, J.E.S. Molecular cloning of a novel melanocortin receptor. *Biochem. Biophys. Res. Commun.* 1993, **195,** 866–873

6 Xia, Y., Wikberg, J.E.S., and Chhajlani, V. Expression of melanocortin 1 receptor in periaqueductal gray matter. *NeuroReport.* 1995, **6,** 2193–2196

7 Xia, Y., and Wikberg, J.E.S. In situ hybridization histochemical localization of ACTH receptor mRNA in mouse adrenal gland. *Cell Tissue Res.* 1996, **286,** 63–68

8 Low, M.J., Simerly, R.B., and Cone, R.D. Receptors for the melanocortin peptides in the central nervous system. *Curr. Opin. Endocrinol. Diabetes* 1994, **1,** 79–88

9 Prusis, P., Frändberg, P.A., Muceniece, R., Kalvinsh, I., and Wikberg, J.E.S. A three dimensional model for the interaction of MSH with the melanocortin-1 receptor. *Biochem. Biophys. Res. Commun.* 1995, **210,** 205–210

10 Frändberg, P.A., Muceniece, R., Prusis, P., Wikberg, J., and Chhajlani, V. Evidence for alternate points of attachment for α-MSH and its stereoisomer [Nle4, D-Phe7]α-MSH at the melanocortin-1 receptor. *Biochem. Biophys. Res. Commun.* 1994, **202,** 1266–1271

11 Schiöth, H.B., Muceniece, R., Prusis, P., Szardenings, M., Lindeberg, G., Sharma, S.D., Hruby, V.J., and Wikberg, J.E.S. Characterization of D117A and H260A mutations in the melanocortin 1 receptor. *Mol. Cell. Endocrinol.* 1997, **126,** 213–219

12 Schiöth, B.H., Muceniece, R., Larsson, M., Mutulis, F., Szardenings, M., Prusis, P., Lindeberg, G., and Wikberg, J.E.S. Binding of cyclic and linear MSH core peptides to the melanocortin receptor subtypes. *Eur. J. Pharmacol.* 1997, **319,** 368–373

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

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

15 Devereux, J., Haeberli, P., and Smithies, O. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 1984, **12,** 387–395

16 Brünger A.T. *X-PLOR, Version 3.1: A System for X-Ray Crystallography and NMR.* Yale University Press, New Haven, Connecticut, 1992

17 Brooks, R.B., Bruccoleri, E.R., Olafson, D.B., States, J.D., Swaminathan, S., and Karplus, M. CHARMM: A program for macro-molecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 1983, **4,** 187–217
18 Oldfield, J.T., SQUID: A program for the analysis and display of data from crystallography and molecular dynamics. *J. Mol. Graphics* 1992, **10,** 247–252
19 MSI. *QUANTA96.* MSI, San Diego, California, 1996
20 Kraulis, J.P., MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 1991, **24,** 946–950
21 Oliviera, L., Paiva, A.C.M., and Vriend, G. A common motif in G-protein-coupled seven transmembrane receptors. *J. Comput. Aided Mol. Design* 1993, **7,** 649–658
22 Schiöth, B.S., Chhajlani, V., Muceniece, R., Klusa, V., and Wikberg, E.S.J. Major pharmacological distinction of the ACTH receptor from other melanocortin receptors. *Life Sci.* 1996, **59,** 797–801
23 Kirkpatrick, S., Gellat, C.D., and Vecchi, M.P. Optimization by simulated annealing. *Science* 1983, **220,** 671–680
24 Nilges, M., and Brünger, A.T. Automated modeling of coiled coils: Application to the GCN4 dimerization region. *Protein Eng.* 1991, **4,** 649–659
25 Herzyk, P., and Hubbard, R.E. Using experimental information to produce a model of the transmembrane domain of the ion channel phospholamban, *Biophys. J.* 1998, **74,** 1203–1214
26 Neria, E., Fisher, S., and Karplus, M. Simulation of activation free-energies in molecular-systems. *J. Chem. Phys.* 1996, **105,** 1902–1921
27 Sankararamakrishnan, R., and Vishveshwara, S. Geometry of proline-containing alpha-helices in proteins. *Int. J. Peptide Protein Res.* 1992, **39,** 365–363
28 Baldwin, M.J. The probable arrangement of the helices in G protein-coupled receptors. *EMBO J.* 1993, **12,** 1693–1703
29 Schiöth, B.H., Muceniece, R., Wikberg, J.E.S., and Chhajlani, V. Characterization of melanocortin receptor subtypes by radioligand binding analysis. *Eur. J. Pharmacol.* 1995, **288,** 311–317
30 Schiöth, B.H., Muceniece, R., and Wikberg, J.E.S. Characterization of the melanocortin 4 receptor by radioligand binding. *Pharmacol. Toxicol.* 1996, **79,** 161–165
31 Robbins, S.L., Nadeau, H.J., Johnson, R.K., Kelly, A.M., Roselli-Rehfuss, L., Baach, E., Mountjoy, G.K., and Cone, D.R. Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell* 1993, **72,** 827–834
32 Klungland, H., Våge, I.D., Gomez-Raya, L., Adalsteinsson, S., and Lien, S. The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. *Mamm. Genome* 1995, **6,** 636–639
33 Valverde, P., Healy, E., Jackson, I., Rees, L.J., and Thody, J.A. Variants of the melanocyte stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nature Genet.* 1995, **11,** 328–330
34 Marklund, L., Moller, J.M., Sandberg, K., and Andersson, L. A missence mutation in the gene for melanocyte-stimulating hormone receptor (MC1R) is associated with chestnut coat color in horses. *Mamm. Genome* 1996, **7,** 895–899
35 Koppula, V.S., Robbins, S.L., Lu, D., Baack, E., White, R.C., Swanson, A.N., and Cone, D.R. Identification of common polymorphisms in the coding sequence of the human MSH receptor (MC1R) with possible biological effects. *Hum. Mutat.* 1997, **9,** 30–36
36 Våge, I.D., Lu, D., Klungland, H., Lien, S., Adalsteinsson, S., and Cone, D.R. A non-epistatic interaction of agouti and extension in the fox, *Vulpes vulpes. Nature Genet.* 1997, **15,** 311–315
37 Schiöth, B.H., Muceniece, R., and Wikberg, J.E.S. Unpublished results, 1997
38 Xu, X., Thörnwall, M., Lundin, L.G., and Chhajlani, V. Val92Met variant of the melanocyte stimulating hormone receptor gene. *Nature Genet.* 1996, **14,** 384
39 Schiöth, B.H., Muceniece, R., Szardenings, M., Prusis, P., and Wikberg, E.S.J. Evidence indicating that the TM4, EL2 and TM5 of the melanocortin 3 receptor do not participate in ligand binding. *Biochem. Biophys. Res. Commun.* 1996, **229,** 687–692