Modeled Structure of a G-Protein-Coupled Receptor: The Cholecystokinin-1 Receptor

Elodie Archer-Lahlou,[†] Irina Tikhonova,[†] Chantal Escrieut,[†] Marlène Dufresne,[†] Catherine Seva,[†] Lucien Pradayrol,[†] Luis Moroder,[‡] Bernard Maigret,[§] and Daniel Fourmy*[,†]

INSERM U 531, Institut Louis Bugnard, CHU Rangueil, Bat. L3, 31403 Toulouse Cedex 4, France, Laboratoire de Chimie Théorique, Université de Nancy, 54506 Vandoeuvre les Nancy, and Max Planck Institut für Biochemie, 82143 Martinsried, Germany

Received February 6, 2004

The Cholecystokinin-1 receptor (CCK1R) mediates actions of CCK in areas of the central nervous system and of the gut. It is a potential target to treat a number of diseases. As for all G-protein-coupled receptors, docking of ligands into modeled CCK1R binding site should greatly help to understand intrinsic mechanisms of activation. Here, we describe the procedure we used to progressively build a structural model for the CCK1R, to integrated, and on the basis of site-directed mutagenesis data on its binding site. Reliability of the CCK1R model was confirmed by interaction networks that involved conserved and functionally crucial motifs in G-protein-coupled receptors, such as Glu/Asp-Arg-Tyr and Asn-Pro-Xaa-Xaa-Tyr motifs. In addition, the 3-D structure of CCK1R-bound CCK resembled that determined by NMR in a lipid environment. The derived computational model was also used for revealing binding modes of several nonpeptide ligands and for rationalizing ligand structure−activity relationships known from experiments. Our findings indeed support that our "validated CCK1R model" could be used to study the intrinsic mechanism of CCK1R activation and design new ligands.

Introduction

Cholecystokinin (CCK) is a neuropeptide that is widely distributed in the central nervous system as well as in the intestine.[1] In the body, CCK is composed of several molecular variants differing in length (CCK-58, -39, -33), the octapeptide (CCK-8, Asp-Tyr(SO_3H)-Met-Gly-Trp-Met-Asp-Phe-NH2) being the major fully active one.[1] Actions of CCK are mediated by two receptors, the CCK1 receptor (abbreviated CCK1R) and the CCK2 receptor (abbreviated CCK2R) which both are members of family I of the superfamily of G-protein-coupled receptors (GPCRs).[2,3] CCK1R activation by agonists initiates a cascade of signaling primarily resulting from CCK1R coupling to heterotrimeric protein Gq that activates phospholipase-C and subsequent signaling pathways.[4] CCK1R-mediated effects include control of gallbladder contraction, pancreatic growth and secretion, gastric emptying and gut motility, and satiety.[2,3] The wide spectrum of biological functions regulated by the CCK1R has made this receptor a candidate target for a therapeutic approach in a number of diseases related to nutrient assimilation.[5] As a consequence, academic and pharmaceutical research groups have generated a large variety of agonists and antagonists of the CCK1R, several of which are of therapeutic value and are in advanced clinical evaluation.[5]

Knowledge of the three-dimensional structure of GPCRs is essential for the understanding of their function, for the docking of existing ligands into their binding sites, and further for the rational design and/or optimization of drugs.[6,7] Structurally, GPCRs are believed to share seven predicted transmembrane (TM) helices connected by three extracellular and three intracellular loops. However, the structure of GPCRs at the atomic level are unknown, except for the light receptor rhodopsin.[8] Hence, use of a structure-based approach for drug design and the impact on the mechanisms controlling ligand binding, activation, signal transduction, and regulation of GPCRs are limited. Computer-assisted molecular modeling is an alternative means to access potential 3-D structures of GPCRs. The method relies on a valuable alignment between the target amino acid sequence and that of the template, the 3-D structure of which is available. For a long time the only available 3-D structure of a 7 TM segment protein was that of bacteriorhodopsin, which is not a GPCR and has a very low homology with GPCRs (less than 20%).[9] Homology modeling of GPCRs evolved until the proposal by Baldwin and co-workers of an Cα-template for the transmembrane helices in the rhodopsin family of GPCRs.[10] More recently, modeling software, such as the Viseur program, was developed in order to take into account structural data derived from biochemical and biophysical studies as well as those from site-directed mutagenesis-based investigations.[11] These new modeling modules have significantly contributed to the refinement of GPCR structures, providing constraints to be satisfied during the modeling process, as it was accepted that the models could be modified by translation and rotation of the transmembrane helices. Homology modeling of GPCR can now use the high-resolution 3-D structure of rhodopsin, which is a GPCR.[12]

At the beginning of our work aimed at delineation of the CCK1R binding site for CCK, we modeled the

* To whom correspondence should be addressed. Phone: 33 5 61 32 30 57. Fax: 33 5 61 32 24 03. E-mail: Daniel.Fourmy@toulouse.inserm.fr.
[†] Institut Louis Bugnard.
[‡] Max Planck Institut für Biochemie.
[§] Université de Nancy.

10.1021/jm049886y CCC: $30.25
Published on Web 12/10/2004

CCK1R using bacteriorhodopsin as a template since at this time, in 1996, the only available high-resolution structure of an heptahelical membrane protein was that of bacteriorhodopsin.[9] This work yielded a model for the CCK1R which was then used to dock CCK into its putative binding pocket. By doing so, a first set of candidate residues of the putative binding site were mutated and their involvement in binding and activation of the CCK1R was confirmed.[13] Experimentally validated interactions were then used as constraints to perform a second round of refinement of the liganded CCK1R. This led to finding new potential hits within the binding pocket.[14,15] The step-by-step forced refinement of the CCK1R model using experimental data regarding the binding site for CCK has yielded a so-called "refined model of the CCK1R". In recent work we compared this "refined model of the CCK1R" with another model of the CCK1R that we obtained directly by classical homology modeling using rhodopsin as a reference structure.[16] Notable structural divergences were observed between the two models, and the model that we built directly from the rhodopsin structure could not correctly accommodate CCK in its binding site.[16]

In the current work we describe, for the first time, the procedure used to progressively build and refine our CCK1R model. Furthermore, we evaluated the reliability of this "refined model of CCK1R.CCK complex". For this purpose, interaction networks in the CCK1R which involved conserved and functionally crucial motifs such as Glu/Asp-Arg-Tyr and Asn-Pro-Xaa-Xaa-Tyr motifs were compared to those found in rhodopsin, and the 3-D structure of bound CCK was compared with its NMR structure determined in a lipid environment. In addition, the docking of several nonpeptide ligands (SR-27,897; SR-146,131; L-364,718; L-365,260) within the refined CCK1R model was performed. New insights in the modeled structure of the CCK1R and the release of this refined model to the scientific community should serve to study molecular mechanisms of activation of this interesting receptor and to perform the design of new ligands.

Experimental Methods

General Strategy. Modeling of the human CCK1R was performed using a segmented approach where the transmembrane helices, loop regions, N-terminus, and C-terminus were modeled as separated molecular entities on separate individual templates. An unrefined model of the CCK1R was first generated. It was comprised of transmembrane helices and connecting loops. This unrefined model served for binding site studies and was refined on the basis of ligand docking. Then, in a third step, N-terminal and C-terminal regions, which were homology modeled using separated templates, were added to the refined model, yielding the final refined model of the CCK1R (Figure 1).

Strategy and Methods for Molecular Modeling the "Unrefined CCK1R". The transmembrane bundle of the CCK1R was modeled using the TM arrangement found in bacteriorhodopsin as the starting point since when the unrefined model of the CCK1R was built the only available high-resolution structure of a heptahelical membrane protein was that of bacteriorhodopsin.[9] However, because bacteriorhodopsin has, unfortunately, a poor degree of sequence homology with the CCK1R that does not allow direct use in the procedure of sequence alignment, intermediate steps were introduced to circumvent this difficulty. As the first step, rhodopsin, which is a GPCR exhibiting a higher degree of homology with any GPCRs than does bacteriorhodopsin and, on the other hand,

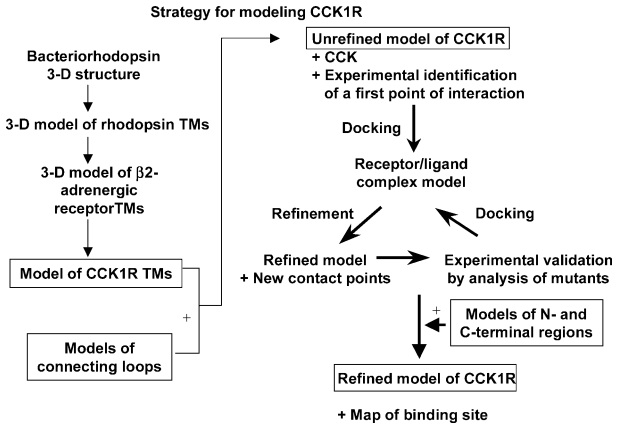

Figure 1. Strategy used to construct a refined model for the CCK1R and map its binding site.

shows functional similarities with bacteriorhodopsin, was modeled by homology on the basis of the 3.5 Å resolution structure of bacteriorhodopsin. Sequence alignment of transmembrane helices was performed as previously described.[17] Briefly, alignment used conserved functional residues found in the two proteins, as, for instance, Lys residue in TM VII that forms a Schiff base with all-*trans*- and 11-*cis*-retinal, respectively, and the acidic counterion of that Schiff base in TM III, namely, Asp85 in bacteriorhodopsin and Glu113 in rhodopsin. For TM VI alignment, conserved Trp in the middle of the helix was taken while for other TM amino acids surrounding retinal ligand were used. In a second step the hamster β2-adrenergic receptor was modeled using the previously built structure of rhodopsin as the template. The choice of hamster β2-adrenergic receptor as the intermediate target in the modeling procedure was dictated by its membership to family I in the superfamily of GPCRs and by availability of functional and pharmacological data which validated the 3-D model of this GPCR and, therefore, the choice for TM alignments. For alignment between hamster β2-adrenergic receptor and human rhodopsin, conserved residues in transmembrane helices of all GPCRs as well as amino acids and motifs which are crucial for activation, such as the Glu/Asp-Arg-Tyr motif at the bottom of TM III and the Asn-Pro-Xaa-Xaa-Tyr motif in TM VII, were taken into account.

For homology modeling of rhodopsin using bacteriorhodopsin as the template, the coordinates of backbone atoms of bacteriorhodopsin were transferred to those of the corresponding backbone atoms of rhodopsin in accordance to the sequence alignment between them, while the side chains were mutated using the Homology module of Insight II (Acelerys, San Diego, CA). The same procedure was applied to obtain a model of the seven-transmembrane helix core of the β2-adrenergic receptor from the rhodopsin-modeled structure and, finally, the unrefined model of the CCK1R from the β2-adrenergic receptor model. At each step newly constructed models were refined using energy minimizations. During the energy refinements the Cα backbone of each seven-transmembrane helix bundle were constrained to their initial positions and gradually relaxed. This starting unrefined CCK1R model was then successively modified in order to take into account data coming from improvements of inactive rhodopsin three-dimensional structures.[18,19] The relative arrangement of the TM helices, in particular TM III and TM IV, was modified in order to reproduce the projection structure of rhodopsin. The "in/out" concept associated with the VISEUR program was used to rotate and translate the helices in the appropriate way.[11]

The extracellular and intracellular loops connecting the TM helices were built using the HOMOLOGY loops module of Insight II, which uses a library of protein templates from the PDB (Protein Data Bank) database. In the procedure a set of high-resolution X-ray structures was searched for protein fragments with a similar length to the CCK1R loop to be modeled and a distance between the Cα atoms of the residues delimiting the loop window, which allows connection to trans-

membrane ends. Modeled loops were then inserted into the unrefined CCK1R bundle at appropriate positions, and the obtained structure was optimized using simulated annealing procedures. A cooling procedure going from 1000 to 300 K with steps of 100 K was used. This annealing procedure was applied only to the loops connecting the helices, the Cα-backbone of which was kept fixed. After minimizing the final annealing conformation removing all the constraints, the entire system was finally relaxed and submitted to a 1 ns molecular dynamics simulations with possible translational and rotational movements of individual TM helices taken into account.

Strategy and Methods for Obtaining the "Refined CCK1R" Model. Improvement of the "unrefined CCK1R model" built as described above was performed step-by-step on the basis of site-directed mutagenesis data regarding the binding site of the agonist ligand CCK. The general strategy for this refinement is shown in Figure 1. As indicated, the first step in the procedure was the docking of CCK in the CCK1R binding groove. For this docking the 3-D structure of CCK was computer generated as previously described.[20] Then we took advantage of the identification of a first contact point between the N-terminal moiety of CCK and two residues Trp39/Gln40 located at the top of TM I.[13] Briefly, this contact point was previously demonstrated by combination of photoaffinity labeling and two-dimensional site-directed mutagenesis. The best evidence for the first contact point was given by experimental results showing that the two receptor mutants resulting from exchange of Trp39 for a Phe and of Gln40 for an Asn demonstrated parallel decreases in both binding affinity for intact CCK and potency to induce inositol phosphate formation while the same mutants bound and responded to N-terminally truncated CCK as did the wild-type CCK1R.[13]

Thanks to this first experimental contact point, CCK was docked manually into the CCK1R groove by placing the C-terminal moiety of the CCK that composes most of the pharmacophore of the peptide into the cavity formed by receptor transmembrane segments, while the N-terminal part was maintained at the entrance of the pocket close to Trp39/Gln40. Manual positioning of CCK into the receptor was guided by inspection of potential electrostatic interactions between ligand and receptor side chains. By doing so, a first series of residues approximated, in particular, the sulfated tyrosine Tyr(SO_3H) of CCK and Met195/Arg197 of the CCK1R. To satisfy these predicted interactions, the position of the second extracellular loop was forced. The obtained structure for the ligand/receptor complex was further refined. Since site-directed mutagenesis studies with mutated receptors at positions Met195 and Arg197 validated the predicted interactions, a third round of ligand docking was achieved.[14,15] It allowed positioning of the C-terminal half of CCK at the bottom of the cavity formed by transmembrane segments. The resulting structure was once more refined. Docking of the C-terminal half of CCK generated multiple networks of interactions between CCK and the CCK1R. Site-directed mutagenesis of amino acids corresponding to these new hits further validated docking of CCK in the CCK1R model.[21,22]

At each step of ligand docking the structure was refined by submitting the receptor−ligand complex to molecular dynamics followed by energy minimization. We first used interaction constraints between amino acids of the receptor and chemical functions of the ligand, the backbone of the receptor being fixed. Then all the constraints were removed except hydrogen bonds within helices, and finally, the whole receptor−ligand complex was submitted to energy minimization.

In a third step the extracellular N-terminal and intracellular C-terminal regions of the CCK1R were built by homology modeling based on structural data from the PDB (Protein Data Bank) database.

The whole modeling was carried out with Accelrys software Insight II modules (Homology, Discover and Biopolymer; Accelrys, San Diego, CA) on a Silicon Graphics O_2 station.

Docking of Nonpeptide Ligands. Autodock3.0 was employed to carry out the docking calculations.[23] The protein model from refined CCK1R.CCK complex was used for docking studies. The charges were assigned for the protein using CVFF residue library of Insight II. The 3-D structures of the ligands were built using Insight II; partial atomic charges were calculated using the MOPAC module of InsightII (AM1-ESP) (Accelrys, 2001, San Diego, CA.). The rotatable bonds in the ligands were defined using AutoTors. The grid maps were calculated using AutoGrid. As used in the GRID calculations, the dimensions of the grid box was 70 × 70 × 70 Å but the grid spacing was set to 0.375 Å. Lennard−Jones parameters 12−10 and 12−6 were used for modeling H-bonds and van der Waals interactions, respectively. The distance-dependent dielectric function was used for calculation of the electrostatic grid map. Docking was performed using the Lamarckian genetic algorithm (LGA), and the pseudo-Solis and Wets methods were applied for the local search. Parameters for the docking experiments were as follows: initial population size of 4000, random starting position and conformation, maximal mutation of 0.1 Å in translation and 5° in orientation and rotation, and maximal iteration per local search of 300. Simulations were performed with a maximum of 270 000 generations. The best position of ligands was chosen on the basis of scoring function and graphical browsing. Then the complexes were subjected to dynamics simulation (1 ns) and following minimization using the Discover3 module of Insight II. During the simulation the Cα trace of the protein was restrained, while it was free in minimization.

Pharmacological and Functional Characterization of CCK1R Mutants. Most of the mutants regarding amino acids of the binding site were previously described. A summary of the functional properties of these mutants is presented in Table 1. Moreover, analysis of our "refined CCK1R model" and its comparison with existing structures led us to construct new CCK1R mutants and analyze their properties as follows.

Site-Directed Mutagenesis and Transfection of COS-7 Cells. All mutant receptor cDNAs were constructed by oligonucleotide-directed mutagenesis (Quick-Change site-directed mutagenesis kit, Stratagene, France) using the human CCK1R cDNAs cloned into pRFENeo vector as template. The presence of the desired and absence of undesired mutations were confirmed by automated sequencing of both cDNA strands (Applied Biosystem).

COS-7 cells (1.5×10^6) were plated onto 10-cm culture dishes and grown in Dulbelcco's Modified Eagle's Medium containing 5% fetal calf serum (complete medium) in a 5% CO_2 atmosphere at 37 °C. After overnight incubation cells were transfected with 2 μg/plate of pRFENeo vectors containing the cDNA for the wild-type or mutated CCK1 receptors using a modified DEAE-dextran method. Cells were transferred to 24-well plates at a density of 20 000−80 000 cells/well 24 h after transfection, depending on transfected mutant and experiment to be performed.

Receptor Binding Assay. Approximately 24 h after the transfer of transfected cells to 24-well plates, the cells were washed with phosphate-buffered saline pH 6.95, 0.1% BSA, and then incubated for 60 min at 37 °C in 0.5 mL of Dublelcco's Modified Eagle's Medium, 0.1% BSA with either 71 pM ^{125}I-BH-(Thr, Nle)CCK-9 (prepared as in ref 24) or 1.83 nM [^{3}H]-SR-27,897 (tritiated 1-[2-(4-(2-chlorophenyl)thiazol-2-yl)aminocarbonyl-indoyl]acetic acid, specific activity 31 Ci/mmol, Sanofi-Synthelabo, Toulouse, France) in the presence or absence of competing agonists or antagonists. The cells were washed twice with cold phosphate-buffered saline pH 6.95 containing 2% BSA, and cell-associated radioligand was collected with NaOH, 0.1 N, added to each well. The radioactivity was directly counted in a gamma counter (Auto-Gamma, Packard, Downers Grove, IL) or added to scintillant and counted for the tritiated radioligand.

Inositol Phosphate Assay. Approximately 24 h after the transfer to 24-well plates and following overnight incubation in complete medium containing 2 μCi/mL of myo-2-[3H]inositol (Amersham biosciences, Les Ulis, France), the transfected cells were washed with Dubelcco's Modified Eagle's Medium and then incubated for 30 min in 1 mL/well Dubelcco's Modified

Table 1. Summary of Mutagenesis Data Used for Refinement of CCK1R Model

CCK1R region	CCK1R aa	CCK aa in interaction	CCK1R mutation	effect of mutation ↓↑	ref
N-term	Trp39	N-term, Arg1, Asp2	Phe	↓ 10 -fold affinity and potency	13
N-term	Gln40		Asn	↓ 10 -fold affinity and potency	13
TM I	Leu50	Met/Nle	Ala	no effect	22
TM I	Ile51	Met/Nle	Ala	↓ 2-fold affinity ↓ 40%potency	22
TM I	Leu53	Met/Nle	Ala	no effect	22
TM I	Cys94	Met/Nle	Leu	↓ 62-fold affinity ↓ 28-fold potency ↓ 55% efficacy	22
EL1	Lys105		Leu	↓ 4-fold affinity ↓ 16-fold potency	15
EL1	Asp106	Arg1	Ala	↓ 4-fold affinity and potency	13
EL1	Phe107	Trp	Ala	↓ affinity[a] ↓ 4500-fold potency	c
EL 1	Lys115		Leu	no effect	d
TM III	Thr117		Ala	no effect	d
TM III	Thr118		Ala	no effect	d
TM III	Met121	Met/Nle	Val	↓ 16-fold affinity ↓ 100% efficacy	22
TM III	Ser124		Ala	↑ efficacy	c
TM III	Val125	Phe	Ala	↑ 2-fold affinity ↓ 2-fold potency	22
TM IV	Phe170		Ala	no effect	d
EL 2	Met195	Tyr	Leu	↓ 30-fold affinity[b] ↓ 54-fold potency	14
EL 2	Arg197	SO3H (Tyr)	Met	↓ affinity[a] ↓ 3150-fold potency	15
TM V	Phe218	Phe	Ala	↓ 2-fold affinity ↓ 65%efficacy	22
TM VI	Trp326	Phe	Ala	↑ 5-fold affinity ↓ 4-fold potency	22
TM VI	Ile329	Phe	Ala	↓ 8-fold affinity[b] ↓ 38-fold potency	22
TM VI	Phe330	Phe	Ala	↓ 10-fold affinity ↓ 3-fold potency ↓ 40% efficacy	22
TM VI	Asn333	C-term amide	Ala	↓ affinitya ↓1300-fold potency ↓ 40% efficacy	21
TM VI	Arg336	Asp8	Met	↓ affinity[a] ↓ 9300-fold potency	21
EL 3	Ser342		Ala	no effect	d
EL 3	Ser348	(Met) C═O	Ala	↓ affinity[a] ↓ 65-fold potency	*c,d*
TM VII	Ile352	Met/Nle, Phe	Ala	↓ affinity[a] ↓ 200-fold potency	22
TM VII	Leu356	Met/Nle, Phe	Ala	↓ 8-fold affinity[b] ↓ 30-fold potency	22
TM VII	Tyr360	Met/Nle, Phe	Phe	↓ 3-fold affinity ↓ 32-fold potency	22

[a] Affinity could not be evaluated precisely. [b] Seen on low-affinity state of the CCK1R. [c] Results from the current study. [d] Unpublished results performed in the course of CCK1R binding site mapping.

Eagle's Medium containing 20 mM LiCl at 37 °C. The cells were washed with PI buffer at pH 7.45: phosphate-buffered saline containing 135 mM NaCl, 20 mM HEPES, 2 mM $CaCl_2$, 1.2 mM $MgSO_4$, 1 mM EGTA,10 mM LiCl, 11.1 mM glucose, and 0.5% BSA. The cells were then incubated for 60 min at 37 °C in 0.3 mL of PI buffer with or without ligands at various concentrations. The reaction was stopped by adding 1 mL of methanol/hydrochloric acid to each well, and the content was transferred to a column (Dowex AG 1-X8, formate form, Bio-Rad, Hercules, CA) for determination of inositol phosphates. The columns were washed twice with 3 mL of distilled water and twice more with 2 mL of 5 mM sodium tetraborate/60 mM sodium formate. The content of each column was eluted by addition of 2.5 mL of 1 M ammonium formate/100 mM formic acid. A 0.5 mL amount of the eluted fraction was added to scintillant, and beta-radioactivity was counted.

Results

Unrefined Model of the CCK1R. The step-by-step alignment between bacteriorhodopsin and human rhodopsin (bR:HumR), human rhodopsin and the hamster β2-adrenergic receptor (HumR:Ham-Beta2), and the β2-adrenergic receptor and human CCK1R (Ham-Beta2:CCK1R) which yielded the bacteriorhodopsin-derived 3-D models of two intermediate GPCRs and subsequently an unrefined model of the CCK1R is shown in Figure 2a−c. Due to the low homology between sequences of rhodopsin and bacteriorhodopsin, alignment was carried out mainly on the basis of similarity in functioning of the two receptors which involved a set of conserved amino acids within the binding site of their ligand. These amino acids are depicted as bold letters. On the basis of this alignment, a 3-D structure of human rhodopsin was obtained (not shown).

To align the hamster β2-adrenergic receptor and the human rhodopsin sequences, homologies between the transmembrane regions of all GPCRs[17] and, more specifically, conserved residues and motifs shown on Figure 2b were used. In the 3-D structure (not shown) built on the basis of this alignment, residues Asp113 (TM III), Phe 290 (TM VI), Ser207, and Ser 204 (TM V) which have been recognized to be crucial for binding of catecholamine agonists were correctly positioned in the receptor groove.[25] The last alignment, between the hamster β2-adrenergic receptor and the human CCK1R, was facilitated by reasonable degrees of homology within transmembrane helices. They ranged from 19% (TM IV) to 46% (TM I, VI) (Figure 2c). A crude model of the CCK1R transmembrane helices was thus obtained (not shown).

Loops connecting the TM helices were modeled on the basis of protein templates from the PDB (Protein Data Bank) database as described in the Experimental Section. They were then inserted at appropriate positions (not shown). Positioning of the second extracellular loop was done in order to allow formation of a disulfide bond between Cys196 (ECL2) and Cys114 at the top of transmembrane III.

Refined Model of the CCK1R. Features of the Refined CCK1R.CCK Complex. The unrefined CCK1R structure composed of transmembrane helices and connecting loops was step-by-step refined on the basis of contact points which were identified between the CCK1R and CCK. These contact points were used as constraints in the refinement methodology (Figures 1 and 3). While experimental results demonstrating the first contact point between the N-terminal moiety of CCK and Trp39/Gln40 could be easily satisfied without affecting the structure of the unrefined model of the CCK1R, search for additional major interactions in the course of CCK docking into the receptor groove required several important structural modifications of the unrefined model.

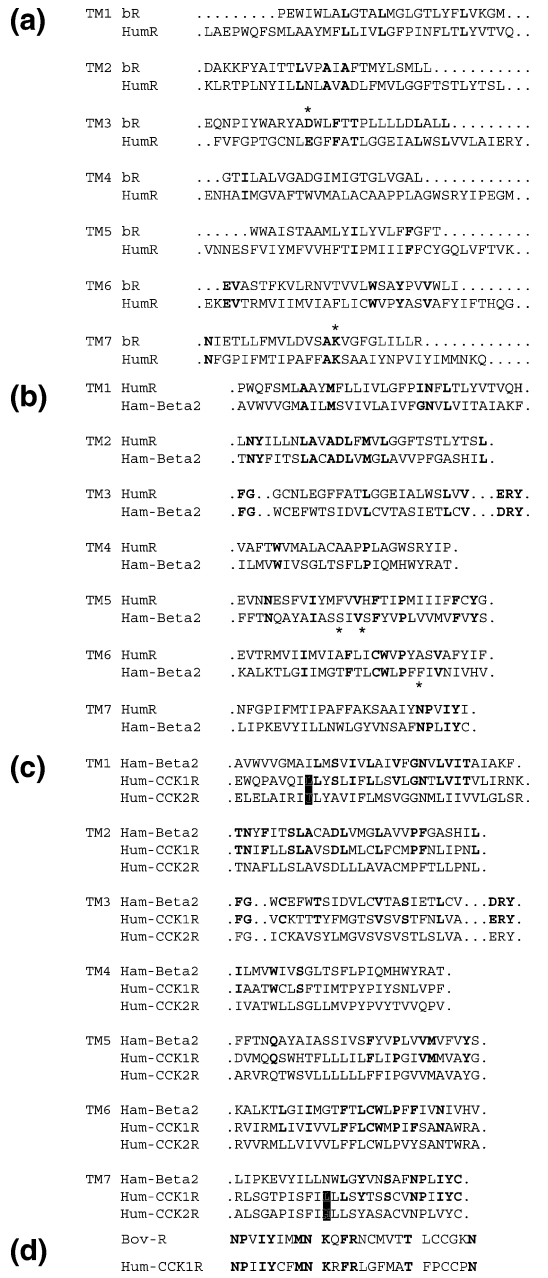

Figure 2. Sequence alignments between transmembrane helices of bacteriorhodopsin and human rhodopsin (bR:HumR, **2a**), human rhodopsin and hamster β2-adrenergic receptor (HumR:Ham-Beta2, **2b**), β2-adrenergic receptor and human CCK1R/CCK2R (Ham-Beta2:CCK1R/CCK2R, **2c**), and between C-terminal regions of bovine rhodopsin and human CCK1R (Bov-R:HumCCK1R) (**2d**). Residues which serve for alignment are in bold letters. Some functionally important residues are labeled with an asterisk (*): Lys (TM VII); Asp85/Glu113 (TM III) in bacteriorhodopsin/rhodopsin; Ser204, Ser 207 (TM V), and Phe 290 (TM VI) in the β2-adrenergic receptor. For clarity, homologous CCK1/CCK2 receptor residues mentioned in the discussion are highlighted.

The first important modification was that of the second extracellular loop of the receptor. The structure of this loop was constrained at its extremities due to connections with helices IV and V. Furthermore, this loop was flattened against the top part of helix III due to the disulfide bond linking Cys196 and Cys114. Although spatial orientation of Arg197/Met195 side chains toward the interior of the transmembrane cavity was proper for strong interactions with the Tyr(SO3H) moiety of

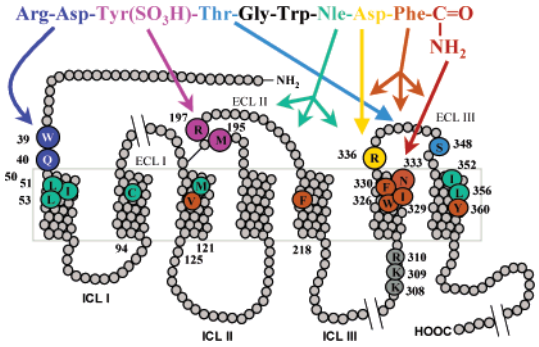

Figure 3. Serpentine representation of the human CCK1R. A code of colors is used to indicate the amino acids involved in binding and activity of the receptor and their corresponding interacting residues in CCK (taken from refs 13, 15, 21, 22, 32, and 35).

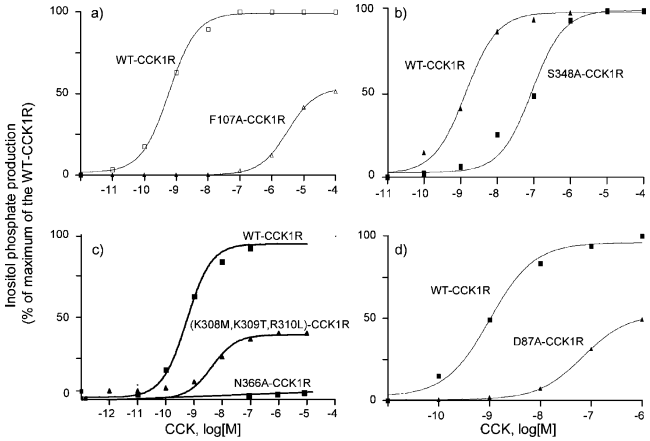

Figure 4. Effects of CCK1R mutations on CCK-induced inositol phosphate production. Inositol production assays were conducted as described in the Experimental Section. Results from all the CCK1R mutants are expressed as percent of maximal inositol phosphate production obtained in COS-7 expressing the wild-type CCK1R after stimulation by CCK. Potency (D_{50}) and efficacy (E_{max}) for the different mutants were as follows: F107A-CCK1R: D_{50} 3140 $\pm$ 232 nM, E_{max} 50; S348A-CCK1R: D_{50} 89.3 $\pm$ 3.3 nM, E_{max} 100; (K308M, K309T, R310L)-CCK1R: D_{50} 4.4 $\pm$ 0.4 nM, E_{max} 35; D87A-CCK1R: D_{50} 74.9 $\pm$ 14 nM, E_{max} 45%.

CCK, distances between the putative interacting partners were not optimal. Reduction of this distance to 3 Å during the modeling procedure required a slight displacement of the loop both laterally toward the receptor cavity and vertically toward the putative membrane plane. Concerning the first extracellular loop, its positioning in the unrefined model did not allow any interaction with CCK. This brought it nearer the Trp side chain of CCK in order to fulfill pharmacological findings showing the importance of Phe107 for CCK binding and action (Figure 4a). This step of refinement was carried out after docking of the C-terminal amided part of CCK into the proximity of Arg336 and Asn333 had been experimentally validated (see later). Initial positioning of the third extracellular loop allowed the hydroxyl group of Ser348 in this loop to directly approximate the backbone carbonyl of Met/Nle in CCK, a result which fully agrees with the 100−200-fold decreased affinity and potency of (S348A)-CCK1R mutant, as shown in Figure 4b.

The intracellular loops, the structures of which were generated by homology modeling using available se-

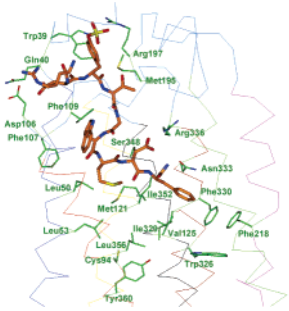

Figure 5. Side view of the CCK1R binding site for CCK. This side view shows the three-dimensional refined model of the active high-affinity CCK1R binding site. The model was built as described in the Experimental Section using the program package from Acelerys (San Diego, CA). For clarity, the detailed view shows only identified amino acid side chains of the CCK1R (in green) in interaction with CCK (in orange) in the phospholipase-C-coupled high-affinity CCK1R.CCK complex.

quences of the PDB library, were maintained in their initial conformations. It is obvious that such structures remain only hypothetical in the absence of a lipid environment and of interacting proteins such as heterotrimeric G proteins, regulatory G proteins (RGS), and G-protein-coupled receptor kinases which in the cellular context most likely affect the real conformation of intracellular regions of any GPCRs.

An important network of intramolecular interactions is seen in the refined model of CCK1R.CCK complex. These include hydrophobic interactions between amino acids of the transmembrane helices and ionic and hydrogen bonds between extracellular and intracellular regions (not shown). Furthermore, strong bonds are found between the two partners of the complex. Amino acids of the CCK1R involved in the CCK binding site are shown in Table 1 and Figures 3 and 5. Briefly, the N-terminal residues of [Thr-Nle]CCK-9 are connected to the receptor through a strong hydrogen bond and salt-bridge network. Asp106 forms a salt bridge between its side chain and the Arg-1 side chain of [Thr-Nle]CCK-9. Gln40 forms a H-bond between its NH amide side chain and the Asp-2 side chain of [Thr-Nle]CCK-9 and between its carboxyl side-chain group and the N-terminal ammonium of the ligand. Trp39 is hydrogen-bonded by its backbone carbonyl to the ligand N-terminal ammonium and by its backbone NH to the Tyr-3 sulfate oxygen of [Thr-Nle]CCK-9.

Besides interactions involving Trp39 and Gln40 residues and the N-terminal part of CCK,[13] two key interactions which account for the 500−1000-fold selectivity of the CCK1R for sulfated versus nonsulfated CCK are seen. A first key interaction involves the sulfur atom of Met195, located in the second extracellular loop and the aromatic ring of the sulfated tyrosine of CCK. In this interaction the sulfur atom of Met195 is positioned toward the center of the aromatic ring of the tyrosine at a distance of about 6 Å and the methyl group attached to the sulfur atom is pointed away from the aromatic plane. Such a positioning is typical of quadrupole/quadrupole interactions between a sulfur atom and the π-electron cloud of an aromatic ring.[26,27] The second key interaction in the same region involves the positively charged guanidium of Arg197 and negatively charged sulfate of CCK, the two interacting partners being at a distance of 4 Å. The last bond appears to be stabilized by Met195/Tyr(CCK) quadrupole−quadrupole interactions as well as by a T-shaped interaction between Trp39 of the CCK1R and Tyr(CCK).[14,15] While the N-terminal moiety of CCK is tightly linked to extracellular residues of the CCK1R, the C-terminal tetrapeptide of CCK appears as embedded between TM helices. In this context, several crucial networks of interactions must be mentioned. A first one concerns positively charged guanidium of Arg336 located at the top of TM VI and negatively charged carboxylate of the penultimate Asp of CCK. A second one involves the side chain of Asn333 located at the top of TM VI, one helix-turn down of Arg336, and the C-terminal amide of CCK. Hydrophobic chemical moieties within the C-terminal part of CCK, namely, Nle/Met and Phe side chains, which are also essential for binding and biological activities of CCK, are strongly engaged in two interdependent hydrophobic clusters. A first hydrophobic pocket surrounding Met/Nle residues of CCK is formed by the amino acids Met121 (TM III), Ile352, Leu356, Phe360 (TM VII), Leu50, Leu53 (TM I), and Cys94 (TM II). A second pocket surrounding the Phe aromatic of CCK is composed of Ile352, Leu356, Tyr360 (TM VII), Val125 (TM III), Phe218 (TM V), Trp326, Ile329, and Phe330 (TM VI) (Figure 5).

Model for the Extracellular N-Terminal and the Intracellular C-Terminal Regions. The N-terminal and C-terminal regions of the CCK1R were modeled on the basis of NMR data and rhodopsin structure, respectively, to be connected to the refined structure of the CCK1R. For the N-terminal region modeling was carried out on the basis of the high-resolution NMR structure of the biomolecular complex between the octapeptide of cholecystokinin, CCK-8, and the N-terminus of the CCK-1 receptor$^{1-47}$ (PDB ID: 1D6G). After homology modeling, sequence 1−31 was added to the previously refined CCK1R. In this structure there is a disulfide bond between Cys18 and Cys29, in agreement with data previously reported for the CCK1R.[28]

For the C-terminal region of the CCK1R, since part of this region of the CCK1R (369−390) has a reasonable degree of homology with rhodopsin it was modeled based on the X-ray diffraction structure of the bovine rhodopsin. To do this the alignment between the C-terminus of CCK1R and the C-terminus of rhodopsin was achieved using the common Asn-Pro-Xaa-Xaa-Tyr motif of helix VII as a reference point (Figure 2d). As found in rhodopsin crystal, an eighth helix was obtained. For sequence 391−409 of the CCK1R, a homotrimer of collagen V (PDB ID, 1A89) was chosen as the template since it has the highest identity and smallest number of gap residues. Finally, for region 410−428 of the CCK1R, a reovirus core (PDB ID, 1EJ6) was chosen for the template since it has 75% of identity.

Interaction Network between Functional Motifs in the Modeled CCK1R.CCK Complex. In the final model of the CCK1R (not shown), we examined amino acids within transmembrane domains and motifs that are highly conserved in the group I receptors of the superfamily of G-protein-coupled receptors. These included residues in TM I, II, and III, the Glu/Asp-Arg-Tyr motif at the end of TM III, and the Asn-Pro-Xaa-Xaa-Tyr motif in TM VII. The network of interactions

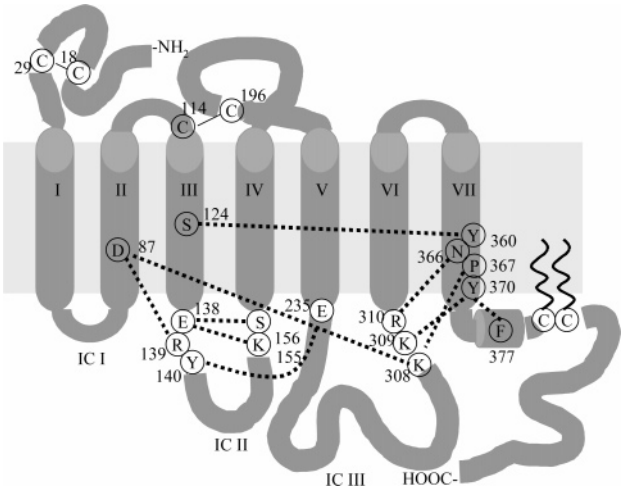

Figure 6. Serpentine representation of the active CCK1R.CCK complex showing interacting residues involving functionally crucial motifs in GPCRs.

between these residues have been demonstrated to undergo modifications during GPCR activation.[29]

In TM II of inactive rhodopsin an acidic residue (Asp83) interacts with the side chain of Asn55 of TM I. In the refined CCK1R.CCK complex the corresponding residue Asp87 is double-linked to Arg139 of the Glu/Asp-Arg-Tyr motif and Lys308 (Figure 6). Of note, Lys308 is one of the three basic amino acids (Lys308-Lys309-Arg310) in the C-terminus of the third intracellular loop of the CCK1R. This triple-basic motif is required for full phospholipase-C activation through Gαq coupling as illustrated on Figure 4c. Furthermore, involvement of Asp87 of the CCK1R in phospholipase-C activation agrees with our experimental results, showing that exchange of Asp87 for an Ala caused parallel drops in affinity and potency of the CCK1R which are most likely due to a defect in the activation/coupling process (Figure 4d). In rhodopsin Glu113 (TM III) plays the role of counterion of Lys296 (TM VII), which represents the site of attachment of retinal. In the β2-adrenergic receptor Asp113 side chain is believed, on the basis of mutagenesis data, to participate to agonist binding. There is no such acidic residue in the TM III of the CCK1R. The corresponding residue is Thr117, the mutation of which does not affect the binding and activation properties of the CCK1R (Table 1). However, two helix-turns down in TM III, Ser124 forms an hydrogen bond with Tyr360 of TM VII. This interaction, if it really exists, does not seem essential for CCK1R functioning since mutation of Ser124 and Tyr360 only affect weakly CCK-induced production of inositol phosphates (Table 1 and ref 22).

Regarding the Glu/Asp-Arg-Tyr motif, the following network of interactions is found in rhodopsin:[8,30] opposite-charged Asp and Arg residues point to the interior of the heptahelical bundle and form a salt bridge, while Tyr side chain is orientated in the opposite direction; furthermore, Arg is interacting with Glu247 located in the C-terminus of the third intracellular loop (Glu-Lys-Glu motif); the Glu/Asp-Arg-Tyr motif of rhodopsin is surrounded with several hydrophobic amino acids. A picture of the Glu/Asp-Arg-Tyr motif (Glu138-Arg139-Tyr140) of the refined CCK1R.CCK complex is shown in Figure 7a. Like in rhodopsin, the Glu/Asp-Arg-Tyr

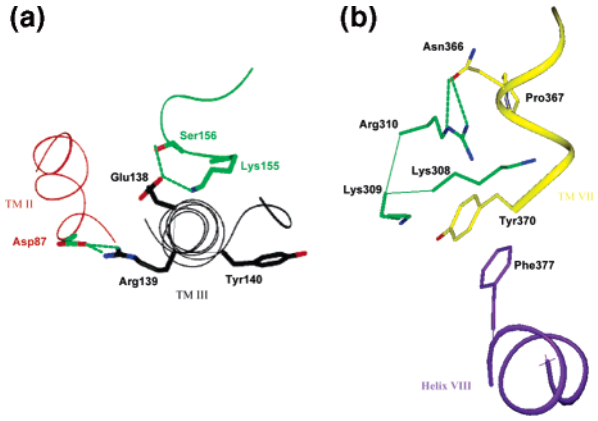

Figure 7. Detailed views of (a) Glu/Asp-Arg-Tyr region and (b) Asn-Pro-Xaa-Xaa-Tyr region of liganded CCK1R. To locate residues in the CCK1R structure, see Figure 6. Interactions are described in the Results section.

motif of the CCK1R appears in a cluster of several hydrophobic residues composed of Phe79, Leu83 (TM II), Ile135 (TM III), Ala228, Leu231 (TM V), Met314, Leu315 (TM VI), Pro367 (TM VII), Ile143 (ICL2), Leu236, Val311 (ICL3) (not illustrated). No bond is seen between Glu138 and Arg139, although the side chains of the two residues are pointing in the same direction to the interior of the receptor groove. Interestingly, Tyr140 hydroxyl is at an interaction distance of Glu235 which is located at the interface between the third intracellular loop and TM V, Glu138 is interacting with Ser156 and Lys155 at the bottom of TM IV, and Arg139 is salt-bridged to Asp87 of TM II.

A last crucial motif for GPCR activation is the Asn-Pro-Xaa-Xaa-Tyr motif of TM VII. In the CCK1R, mutation of Asn366 abolishes receptor activation (Figure 4c). In the refined model of the CCK1R.CCK complex, Asn366 side chain approximates Arg310 which belongs to basic motif Lys308-Lys309-Arg310 at the C-terminus of the third intracellular loop (Figure 7b). In addition, Pro367 and Tyr370 form hydrogen bonds with Lys308 and Lys309, respectively. Finally, like in the rhodopsin structure, the Asn-Pro-Xaa-Xaa-Tyr motif is close to helix VIII because of several interactions such as ring−ring interactions between Tyr370 and Phe377 (Figure 6).

3-D Structure of Receptor-Bound CCK. To perform docking of CCK into the CCK1R binding site, previously modeled CCK was taken.[20] However, the structure of CCK evolved in the course of its binding to the CCK1R to adopt a constrained 3-D structure. As shown in Figure 8, the CCK ligand adopts a helical conformation in the receptor-bound state. We then compared the structure of the bound ligand with that determined by NMR in a membrane-mimetic environment.[31,32] Indeed, studies with N-terminal lipo-derivatized CCK analogues have supported a membrane-bound pathway for the binding of CCK to its receptors.[33] Interestingly, the receptor-bound CCK displays structural characteristics very similar to those previously found for CCK in the presence of dodecylphosphocholine micelles (Figure 8c).[31,32] The rmsd between the two structures is 2.42. Interestingly, the C-terminal region of CCK adopted a folded conformation in the presence of dodecylphosphocholine micelles, whereas this region is in an extended conformation when bound to CCK1R.

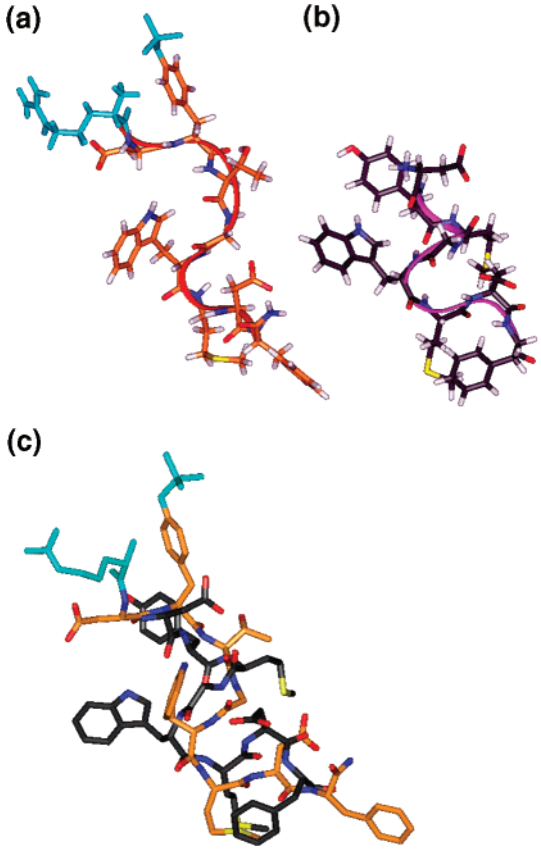

Figure 8. Structure of CCK bound to the CCK1R in the refined model of CCK1R.CCK complex (a), and NMR structure of CCK in DPC micelles taken from PDB[31] (b). Superimposition of the two structures (rmsd = 2.42) (c). For clarity, hydrogens were removed from c.

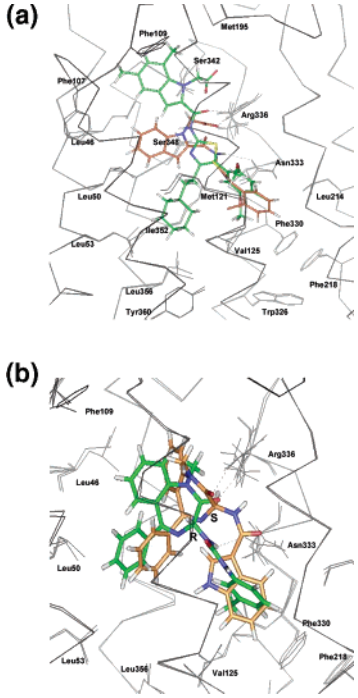

Figure 9. Binding mode of nonpeptide ligands to the refined model of the CCK1R. This figure shows superimposition of complexes formed between the CCK1R and antagonist SR-27897 (in yellow) and between the CCK1R and agonist SR-146131 (in green) (a); superimposition of complexes formed between the CCK1R and antagonist L-364718 (in yellow) and between the CCK1R and L-365260 (in green) (b). For clarity, the detailed view shows only amino acids side chains of the CCK1R in interaction with the nonpeptide ligands.

Binding Modes of Nonpeptide Ligands. The docking of nonpeptide ligands revealed that the optimal orientation of the selective CCK1 antagonist SR-27897 could be achieved when the carboxyl group of ligand is allowed to form a salt bridge with Arg336; the carbonyl and hydrogen of amide group are hydrogen-bonded with Ser348 and Asn333, respectively. The phenyl moiety provides hydrophobic interactions with Phe330, Val125, and Leu214, and the indole moiety lies in the hydrophobic pocket formed by Leu46, Leu50, and Ile352 (Figure 9). In the case of docking experiments with selective CCK1 agonist SR-146131, the orientation of molecule is similar. However, with this compound the cyclohexane ring forms hydrophobic interactions with Leu356, Leu125, Leu53, Ile352, and Met121 and the indole moiety approximates Phe107 and Phe109 side chains as observed with the Trp moiety of CCK. The docking of selective CCK1 antagonist L 364718 and selective CCK2 antagonist L 365260 shows that the binding positions of these ligands are similar, namely, the O=C group of the benzodiazepine ring forms hydrogen bonds with Arg336 and H−N< of the amide group interacts with Asn333 (Figure 9).

Discussion

G-protein-coupled receptors (GPCRs) are widely expressed in the body and play a fundamental role in physiology and pathophysiology. As such, they are potential targets for therapeutic interventions in many, if not most, diseases. Of the ~500 currently sold drugs, more than 30% are modulators of GPCR function.[34] In this context, elucidation of the structure, mechanisms of activation, and regulation of GPCRs is of prime importance. However, crystallization of GPCRs remains extremely difficult to achieve. Computer-assisted molecular modeling is an apparently easy mean to access 3-D structural models of GPCRs. However, one major issue concerning molecular models of GPCRs thus obtained is the question of their accuracy.

The CCK1R is an important GPCR specimen. It mediates actions of the neuropeptide CCK in some areas of the central nervous system and in many organs of the digestive apparatus, making it a potential target to treat a number of diseases affecting food intake, nutriment digestion, and energy homeostasis.[2,3] Furthermore, besides its natural ligand, this receptor possesses a wide variety of synthetic ligands including peptidic and nonpeptidic agonists and antagonists, the synthesis of which has generated abundant structure−activity data. The docking of the most relevant ligands into their binding site should greatly help to understand intrinsic molecular mechanisms which govern CCK1R activation. In this perspective, we generated a structural model for the CCK1R. Unlike models of GPCRs often presented in the literature, our model of the CCK1R was built progressively in order to integrate, and on the basis of, site-directed mutagenesis data regarding its binding site.

The "refined CCK1R model" was recently compared with a model directly obtained by classical homology modeling using the high-resolution structure of rhodopsin as the template.[16] In this study the rhodopsin-

derived model could not correctly accommodate CCK in its binding pocket, thus showing limitation of homology modeling and reinforcing the interest of including agonist docking and site-directed mutagenesis results in the procedure of model refinement. At this stage of our study it was however important to evaluate the degree of accuracy of the "refined CCK1R model" before its release to the scientific community. For this purpose interaction networks between amino acids of the CCK1R itself and with the agonist CCK were analyzed in light of the current knowledge regarding GPCRs. Moreover, the 3-D structure of CCK taken from the refined model of CCK1R.CCK complex was compared with that of CCK in a lipid environment.

The binding site of CCK involves a number of residues located both at the extracellular surface and in the upper third of transmembrane helix.[13−15,21,22,35] In this respect, the CCK1R has strong similarities with peptide GPCRs but differs with GPCRs for amines as well as with the light receptor rhodopsin, which have most, if not all, determinants of its binding site buried in transmembrane helix.[25,36,37] In common with all GPCRs of family I, TM III, TM VI, and TM VII of the CCK1R have major anchoring points of CCK. This common feature is important since spectroscopic studies on rhodopsin and β_2-adrenergic receptor and transfer of a metal-ion binding site by site-directed mutagenesis to peptide GPCRs have demonstrated the importance of transmembrane helix III, VI, and VII in the activation process.[29,38−40]

In family I of GPCRs the acidic residue of the Glu/Asp-Arg-Tyr motif is believed to undergo protonation during activation. In accordance with this view, exchange of the acidic residue Glu/Asp for an uncharged residue such as an Ala results is constitutive activation of the receptors.[41,42] In the refined model of CCK1R.CCK complex, Glu138, instead of being protonated, is linked to the facing residues Ser156 and Lys155 of the second intracellular loop while Arg139 is salt-bridged to Asp87 of TM II. Salt bridging the Arg residue of the Glu/Asp-Arg-Tyr motif to the conserved Asp is in agreement with the activation mechanism initially proposed by Ballesteros et al. The authors suggested that in the inactive state the side chains of Glu/Asp and Arg residues of the conserved Glu/Asp-Arg-Tyr motif form an electrostatic interaction and that during receptor activation Glu/Asp becomes protonated, causing the Arg residue to interact with Asp of TM II.[43] This representation of the activated state of the CCK1R fits experimental data from the current study, indicating that Asp87 of TM II is required for full biological activity. Similar data were reported for several other GPCRs.[29] Moreover, it agrees with studies on rhodopsin showing the proximity of the Glu134 and Arg135 side chains in the crystal of the inactive receptor and with spectroscopic results indicating that Asp83 of TM II is more strongly hydrogen-bonded upon activation, a fact that is consistent with its potential interaction with another residue in the active state of the receptor.[44] One must however mention the alternative mechanism proposed by Scheer et al. in which the Arg residue of the Glu/Asp-Arg-Tyr motif would be associated to the conserved Asp of TM II in the inactive state of the receptor; this interaction would be interrupted upon activation.[41]

Several other important features of the refined model of the CCK1R.CCK complex are in good agreement with the experimental results on this receptor and with data on other GPCRs. Indeed, in the model a network of intramolecular interactions involving, notably, residues Phe218, Met121, Phe330, and Trp326 is found. The role of residues Met121, Cys94, Phe218, Phe330, and Trp326 in the conformational stability of CCK1R was previously documented[22] (and unpublished data). Indeed, pharmacological data indicated that exchange of Met121 for a Val or an Ala converted the whole CCK1R population into a single relatively high- affinity state as did mutation of residues Cys94, Phe218, and Phe330. These were interpreted by considering the prevailing model for G-protein-coupled receptor activation, which is the allosteric ternary complex formed between the receptor R, the agonist L, and G protein(s).[45,46] According to this model, in the absence of any agonist stimulation, R is believed to undergo spontaneous conformational changes, however, with the inactive conformation being energetically the most stable. Binding of an agonist would either induce or stabilize active receptor species (R*) or both.[45,46] Accordingly, Met121, Phe330, Cys94, and Phe218 would represent key residues allowing the receptor to be stabilized in an inactive conformation in absence of ligand and undergo proper conformational changes for G-protein(s) coupling and phospholipase-C activation in the presence of the agonist.

Moreover, the bond between TM III and VII through residues Ser124 and Tyr360 is in line with the slight role of these residues for CCK1R activation and can be compared with the interaction between Glu113 and Lys296 of rhodopsin TM III and VII. Finally, interactions between residues of the Asn-Pro-Xaa-Xaa-Tyr motif and Lys308-Lys309-Arg310 of the C-terminal part of the third intracellular loop through hydrogen bonds is in line with results showing that both Asn366 and the three basic residues within the C-terminal part of the third intracellular loop are required for full biological activity of the CCK1R. The same findings were reported for other GPCRs.[29,47]

The last bit of interesting information given by the refined model of the CCK1R.CCK complex concerns the helical structure of CCK bound to the CCK1R which resembles that determined by NMR in the presence of DPC micelles known to mimic the water−phospholipid bilayer interface. The observation that the C-terminal Phe of CCK adopts distinct conformations in micelles and while bound to the CCK1R might have functional significance because the C-terminal amidated Phe of this peptide is crucial for full biological activity. This result is similar to that from NMR studies which demonstrated that conformations of an analogue of PACAP (the pituitary adenylate cyclase activating peptide) in the presence of DPC micelles and while bound to its receptor are very similar.[48] Functional studies involving the use of N-terminal lipo-derivatized CCK analogues have provided evidence to support a membrane-bound pathway for interaction of CCK with its cognate receptors.[33] In this mechanism adsorption of the ligand to the cell membrane is followed by its lateral diffusion whereby the ligand binds to and activates the receptor.[49] Thus, the conformation of the ligand in its cell membrane-bound state is believed to resemble the

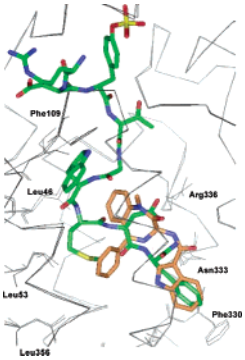

Figure 10. Superimposition of CCK (in green) and the antagonist L364718 (in yellow) in the binding site of the CCK1R.

active conformation. Such a mechanism for ligand binding to GPCRs allows reduction of the entropic penalty associated with ligand−receptor recognition.[49] The NMR structure of another neuropeptide, neurotensin, bound to its receptor was recently determined by NMR. Obtained results revealed that receptor-bound neurotensin adopts a linear structure which differs with the disordered state of the peptide in absence of receptor.[50]

Finally, automated docking of nonpeptide ligands allowed us to describe the binding site of agonists and antagonists within the CCK1R. The importance of residues that form agonist and antagonist binding sites is supported by site-directed mutagenesis data.[21,22] According to docking results, the binding sites of nonpeptide agonists and antagonists most likely occupy a region of CCK1R (Figure 10) which interacts with the C-terminal amidated tripeptide of CCK, i.e., Met-Asp-Phe-NH_2, and moreover, there is an overlap between agonist and antagonist binding sites. Indeed, both nonpeptide agonists and antagonist interact with Arg336, Asn333 and form hydrophobic contacts with two hydrophobic pockets composed of Phe330, Ile329, Val125 and Ile352, Leu50, Met121, Leu53, respectively (Figure 9). In the case of agonist SR-146131, there is a hydrophobic interaction with Leu356 which is absent in the binding mode of the antagonist SR-27897. This finding is in agreement to the mutagenesis study by Gouldson et al.[51] On the other hand, the antagonist binding site of SR-27897 in our CCK1R refined model is not totally in accordance with that proposed by Gouldson et al.[51] Indeed, these authors stated that antagonist SR-27897 interacts with Lys115 and Lys187 as mutation of these residues into alanine reduced by 56- and 68-fold the binding affinity of SR-27897, respectively. We could not confirm such results but observed that mutation of Lys115 dramatically affected CCK1R expression at the cell surface (unpublished results). Considering the positions of Lys115 and Lys187 in the receptor, respectively, the top of TM III and the second extracellular loop, one can assume that mutation into alanine can influence receptor conformation due to loss of interaction with negatively charged lipids of the cell membrane.

According to the docking results into our refined model of CCK1R, selective CCK1 antagonist L-364718 and selective CCK2 antagonist L-365260 were positioned in the same way into the binding pocket. However, in the case of the CCK2R ligand, L-365260, which is in the *R*-conformation, Leu46 side chain is very close to the 1,4-benzodiazepine moiety. Evidently this can cause steric hindrance, explaining in part the low affinity of this compound for the CCK1R. Interestingly, in the CCK2R there is a threonine (Thr59) in place of Leu46 of the CCK1R (Figure 2c). Detailed analysis of the L-365260 binding site also made it possible to explain structure−activity relationships in this family of compounds. For instance, addition of hydrophilic substituents in place of the methyl moiety on the phenyl ring of L-365260 decreases the affinity for the CCK1 receptor, whereas addition hydrogen-bondable moieties to this one increases the affinity for the CCK2 receptor.[5] The preference for additional hydrogen-bondable moieties for CCK2R binding deals with close position of nonconservative residue His376 (Leu356 in the CCK1 receptor, Figure 2c), which has been shown to be important for L-365260 binding selectivity to the CCK2R.[52]

Apart from Gouldson's model of the CCK1R, several other models were reported in the literature, although their coordinates were not released, except for a theoretical model (PDB: 1PB2). Comparison of these other models with our refined model is somewhat difficult and can be done only with respect to data on binding sites since no description of other functionally critical regions has been reported. A first model was built by homology with bacteriorhodopsin by Van der Bent.[53] The obtained 3-D model was used for virtual identification of the binding site of small-size antagonists that matches the structural and conformational characteristics of the ligands. This model did not predict correctly determinants of the binding site for CCK; however, it allowed residues which, according to our site-directed mutagenesis and modeling data, are close to or compose the binding site of the antagonist L-364718 to be highlighted. These are Met121, Trp326, Ile329, Asn333, and Leu356. Miller's group modeled the CCK1R and reported several refined versions of a model on the basis of photoaffinity labeling data.[35] Photoaffinity labeling results included the following: labeling of Glu345 by a photoreactive benzophenone analogue of CCK (Gly substitution); labeling of Trp39 by C-terminal photoreactive analogues of CCK.[54,55] Accordingly, a model for binding of CCK to the CCK1R was proposed in which the C-terminus of CCK is in interaction with Trp39, Tyr-(SO_3H) pairs with Arg197, and the N-terminal moiety of CCK is in contact with the third extracellular loop of the CCK1R. In our refined model of the CCK1R and on the basis of site-directed mutagenesis data, Trp39 belongs to the CCK1R binding site but interacts with the N-terminus of CCK.[13] On the other hand, in our refined model of the CCK1R the third extracellular loop, which contains Glu345, is positioned close to the center of the Cα backbone of CCK, which agrees with photoaffinity labeling results.[54] Thus, although the 3-D model of the CCK1R proposed by Miller's group may not be so far from ours, the structure of CCK while bound in their modeled CCK1R contains a β-turn centered to the middle of the peptide which is not consistent with the NMR structure of CCK in a lipid environment.[32,35] More importantly, positioning of the C-terminus of CCK in interaction with Trp39 (top of TM I) is inconsistent with structure−activity data on CCK. Indeed, deletion studies demonstrated that the amidated C-terminal Phe

residue of CCK highly contributes to CCK binding and activity.[56] If Trp39 was the residue in interaction with C-terminal Phe of CCK, mutation of Trp39 should cause an equivalent drop in the affinity of the receptor for CCK. This is not the case since exchange of Trp39 for a Phe or an Ala induced 10−20-fold-decreases in affinity of the receptor for CCK.[13,57]

To summarize, analysis of our refined model of the CCK1R and its release to the scientific community represent an additional basis for future investigation aimed at understanding the activation process of this important GPCR and to design or optimize ligands. These are major challenges for us in the near future.

Acknowledgment. This study was supported by grants from the Association pour la Recherche sur le Cancer 5481 and 4430.

References

(1) Rehfeld, J. F. Four basic characteristics of the gastrin-cholecystokinin system. *Am. J. Physiol.* **1981**, *240*, G255−266.
(2) Wank, S. A. G protein-coupled receptors in gastrointestinal physiology. I. CCK receptors: an exemplary family. *Am. J. Physiol.* **1998**, *274*, G607−613.
(3) Silvente-Poirot, S.; Dufresne, M.; Vaysse, N.; Fourmy, D. The peripheral cholecystokinin receptors. *Eur. J. Biochem.* **1993**, *215*, 513−529.
(4) Williams, J. A. Intracellular signaling mechanisms activated by cholecystokinin-regulating synthesis and secretion of digestive enzymes in pancreatic acinar cells. *Annu. Rev. Physiol.* **2001**, *63*, 77−97.
(5) Herranz, R. Cholecystokinin antagonists: pharmacological and therapeutic potential. *Med. Res. Rev.* **2003**, *23*, 559−605.
(6) Becker, O. M.; Shacham, S.; Marantz, Y.; Noiman, S. Modeling the 3D structure of GPCRs: advances and application to drug discovery. *Curr. Opin. Drug Discov. Devel.* **2003**, *6*, 353−361.
(7) Rognan, D. Application of three-dimensional (3D) models of G protein-coupled receptors (GPCRs) to drug discovery. *Curr. Opin. Drug Discov. Devel.* **2003**, *6*, 434.
(8) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H. et al. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289*, 739−745.
(9) Henderson, R.; Baldwin, J. M.; Ceska, T. A.; Zemlin, F.; Beckmann, E.; et al. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* **1990**, *213*, 899−929.
(10) Baldwin, J. M. The probable arrangement of the helices in G protein-coupled receptors. *EMBO J.* **1993**, *12*, 1693−1703.
(11) Campagne, F.; Jestin, R.; Reversat, J. L.; Bernassau, J. M.; Maigret, B. Visualisation and integration of G protein-coupled receptor related information help the modelling: description and applications of the Viseur program. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 625−643.
(12) Teller, D. C.; Okada, T.; Behnke, C. A.; Palczewski, K.; Stenkamp, R. E. Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry* **2001**, *40*, 7761−7772.
(13) Kennedy, K.; Gigoux, V.; Escrieut, C.; Maigret, B.; Martinez, J.; et al. Identification of two amino acids of the human cholecystokinin-A receptor that interact with the N-terminal moiety of cholecystokinin. *J. Biol. Chem.* **1997**, *272*, 2920−2926.
(14) Gigoux, V.; Escrieut, C.; Silvente-Poirot, S.; Maigret, B.; Gouilleux, L.; et al. Met-195 of the cholecystokinin-A receptor interacts with the sulfated tyrosine of cholecystokinin and is crucial for receptor transition to high affinity state. *J. Biol. Chem.* **1998**, *273*, 14380−14386.
(15) Gigoux, V.; Maigret, B.; Escrieut, C.; Silvente-Poirot, S.; Bouisson, M.; et al. Arginine 197 of the cholecystokinin-A receptor binding site interacts with the sulfate of the peptide agonist cholecystokinin. *Protein Sci.* **1999**, *8*, 2347−2354.
(16) Archer, E.; Maigret, B.; Escrieut, C.; Pradayrol, L.; Fourmy, D. Rhodopsin crystal: new template yielding realistic models of G-protein-coupled receptors? *Trends Pharmacol. Sci.* **2003**, *24*, 36−40.
(17) Cserzo, M.; Bernassau, J. M.; Simon, I.; Maigret, B. New alignment strategy for transmembrane proteins. *J. Mol. Biol.* **1994**, *243*, 388−396.
(18) Unger, V. M.; Hargrave, P. A.; Baldwin, J. M.; Schertler, G. F. Arrangement of rhodopsin transmembrane alpha-helices. *Nature* **1997**, *389*, 203−206.
(19) Krebs, A.; Villa, C.; Edwards, P. C.; Schertler, G. F. Characterisation of an improved two-dimensional p22121 crystal from bovine rhodopsin. *J. Mol. Biol.* **1998**, *282*, 991−1003.
(20) Pattou, D.; Maigret, B.; Fournie-Zaluski, M. C.; Roques, B. P. Computational analysis of conformational behavior of cholecystokinin fragments. I−CCK4, CCK5, CCK6, and CCK7 molecules. *Int. J. Pept. Protein Res.* **1991**, *37*, 440−450.
(21) Gigoux, V.; Escrieut, C.; Fehrentz, J. A.; Poirot, S.; Maigret, B.; et al. Arginine 336 and asparagine 333 of the human cholecystokinin-A receptor binding site interact with the penultimate aspartic acid and the C-terminal amide of cholecystokinin. *J. Biol. Chem.* **1999**, *274*, 20457−20464.
(22) Escrieut, C.; Gigoux, V.; Archer, E.; Verrier, S.; Maigret, B.; et al. The biologically crucial C terminus of cholecystokinin and the non-peptide agonist SR-146131 share a common binding site in the human CCK1 receptor. Evidence for a crucial role of Met-121 in the activation process. *J. Biol. Chem.* **2002**, *277*, 7546−7555.
(23) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; et al. Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.
(24) Fourmy, D.; Lopez, P.; Poirot, S.; Jimenez, J.; Dufresne, M.; et al. A new probe for affinity labelling pancreatic cholecystokinin receptor with minor modification of its structure. *Eur. J. Biochem.* **1989**, *185*, 397−403.
(25) Fong, T. M.; Strader, C. D. Functional mapping of the ligand binding sites of G-protein coupled receptors. *Med. Res. Rev.* **1994**, *14*, 387−399.
(26) Auld, D. S.; Young, G. B.; Saunders, A. J.; Doyle, D. F.; Betz, S. F.; et al. Probing weakly polar interactions in cytochrome *c*. *Protein Sci.* **1993**, *2*, 2187−2197.
(27) Burley, S. K.; Petsko, G. A. Weakly polar interactions in proteins. *Adv. Protein Chem.* **1988**, *39*, 125−189.
(28) Ding, X. Q.; Dolu, V.; Hadac, E. M.; Schuetz, M.; Miller, L. J. Disulfide bond structure and accessibility of cysteines in the ectodomain of the cholecystokinin receptor: specific monoreactive receptor constructs examine charge-sensitivity of loop regions. *Receptors Channels* **2003**, *9*, 83−91.
(29) Gether, U. Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocr. Rev.* **2000**, *21*, 90−113.
(30) Okada, T.; Ernst, O. P.; Palczewski, K.; Hofmann, K. P. Activation of rhodopsin: new insights from structural and biochemical studies. *Trends Biochem. Sci.* **2001**, *26*, 318−324.
(31) Pellegrini, M.; Mierke, D. F. Molecular complex of cholecystokinin-8 and N-terminus of the cholecystokinin A receptor by NMR spectroscopy. *Biochemistry* **1999**, *38*, 14775−14783.
(32) Giragossian, C.; Stone, S.; Papini, A. M.; Moroder, L.; Mierke, D. F. Conformational and molecular modeling studies of sulfated cholecystokinin-15. *Biochem. Biophys. Res. Commun.* **2002**, *293*, 1053−1059.
(33) Moroder, L.; Romano, R.; Guba, W.; Mierke, D. F.; Kessler, H.; et al. New evidence for a membrane-bound pathway in hormone receptor binding. *Biochemistry* **1993**, *32*, 13551−13559.
(34) Wise, A.; Gearing, K.; Rees, S. Target validation of G-protein coupled receptors. *Drug Discov. Today* **2002**, *7*, 235−246.
(35) Ding, X. Q.; Pinon, D. I.; Furse, K. E.; Lybrand, T. P.; Miller, L. J. Refinement of the conformation of a critical region of charge-charge interaction between cholecystokinin and its receptor. *Mol. Pharmacol.* **2002**, *61*, 1041−1052.
(36) Okada, T.; Palczewski, K. Crystal structure of rhodopsin: implications for vision and beyond. *Curr. Opin. Struct. Biol.* **2001**, *11*, 420−426.
(37) Schwartz, T. W.; Rosenkilde, M. M. Is there a 'lock' for all agonist 'keys' in 7TM receptors? *Trends Pharmacol. Sci.* **1996**, *17*, 213-216.
(38) Farrens, D. L.; Altenbach, C.; Yang, K.; Hubbell, W. L.; Khorana, H. G. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **1996**, *274*, 768−770.
(39) Elling, C. E.; Thirstrup, K.; Nielsen, S. M.; Hjorth, S. A.; Schwartz, T. W. Engineering of metal-ion sites as distance constraints in structural and functional analysis of 7TM receptors. *Fold Des.* **1997**, *2*, S76−80.
(40) Gether, U.; Lin, S.; Ghanouni, P.; Ballesteros, J. A.; Weinstein, H.; et al. Agonists induce conformational changes in transmembrane domains III and VI of the beta2 adrenoceptor. *EMBO J.* **1997**, *16*, 6737−6747.
(41) Scheer, A.; Fanelli, F.; Costa, T.; De Benedetti, P. G.; Cotecchia, S. Constitutively active mutants of the alpha 1B-adrenergic receptor: role of highly conserved polar amino acids in receptor activation. *EMBO J.* **1996**, *15*, 3566−3578.
(42) Gales, C.; Sanchez, D.; Poirot, M.; Pyronnet, S.; Buscail, L.; et al. High tumorigenic potential of a constitutively active mutant of the cholecystokinin 2 receptor. *Oncogene* **2003**, *22*, 6081−6089.
(43) Ballesteros, J.; Kitanovic, S.; Guarnieri, F.; Davies, P.; Fromme, B. J.; et al. Functional microdomains in G-protein-coupled receptors. The conserved arginine-cage motif in the gonadotropin-releasing hormone receptor. *J. Biol. Chem.* **1998**, *273*, 10445−10453.

(44) Rath, P.; DeCaluwe, L. L.; Bovee-Geurts, P. H.; DeGrip, W. J.; Rothschild, K. J. Fourier transform infrared difference spectroscopy of rhodopsin mutants: light activation of rhodopsin causes hydrogen-bonding change in residue aspartic acid-83 during meta II formation. *Biochemistry* **1993**, *32*, 10277−10282.
(45) Lefkowitz, R. J.; Cotecchia, S.; Samama, P.; Costa, T. Constitutive activity of receptors coupled to guanine nucleotide regulatory proteins. *Trends Pharmacol. Sci.* **1993**, *14*, 303−307.
(46) Kenakin, T. Pharmacological proteus? *Trends Pharmacol. Sci.* **1995**, *16*, 256−258.
(47) He, R.; Browning, D. D.; Ye, R. D. Differential roles of the NPXXY motif in formyl peptide receptor signaling. *J. Immunol.* **2001**, *166*, 4099−4105.
(48) Inooka, H.; Ohtaki, T.; Kitahara, O.; Ikegami, T.; Endo, S.; et al. Conformation of a peptide ligand bound to its G-protein coupled receptor. *Nat. Struct. Biol.* **2001**, *8*, 161−165.
(49) Schwyzer, R. In search of the 'bio-active conformation'—is it induced by the target cell membrane? *J. Mol. Recognit.* **1995**, *8*, 3−8.
(50) Luca, S.; White, J. F.; Sohal, A. K.; Filippov, D. V.; van Boom, J. H.; et al. The conformation of neurotensin bound to its G protein-coupled receptor. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 10706−10711.
(51) Gouldson, P.; Legoux, P.; Carillon, C.; Delpech, B.; Le Fur, G.; et al. Contrasting roles of leu(356) in the human CCK(1) receptor for antagonist SR 27897 and agonist SR 146131 binding. *Eur. J. Pharmacol.* **1999**, *383*, 339−346.
(52) Jagerschmidt, A.; Guillaume-Rousselet, N.; Vikland, M. L.; Goudreau, N.; Maigret, B.; et al. His381 of the rat CCKB receptor is essential for CCKB versus CCKA receptor antagonist selectivity. *Eur. J. Pharmacol.* **1996**, *296*, 97−106.
(53) Van Der Bent, A.; Ijzerman, A. P.; Soudijn, W. Molecular modelling of CCK-A receptors. *Drug Des. Discov.* **1994**, *12*, 129−148.
(54) Ding, X. Q.; Dolu, V.; Hadac, E. M.; Holicky, E. L.; Pinon, D. I.; et al. Refinement of the structure of the ligand-occupied cholecystokinin receptor using a photolabile amino-terminal probe. *J. Biol. Chem.* **2001**, *276*, 4236−4244.
(55) Ji, Z.; Hadac, E. M.; Henne, R. M.; Patel, S. A.; Lybrand, T. P.; et al. Direct identification of a distinct site of interaction between the carboxyl-terminal residue of cholecystokinin and the type A cholecystokinin receptor using photoaffinity labeling. *J. Biol. Chem.* **1997**, *272*, 24393−24401.
(56) Jensen, R. T.; Wank, S. A.; Rowley, W. H.; Sato, S.; Gardner, J. D. Interaction of CCK with pancreatic acinar cells. *Trends Pharmacol. Sci.* **1989**, *10*, 418−423.
(57) Gouldson, P.; Legoux, P.; Carillon, C.; Delpech, B.; Le Fur, G.; et al. The agonist SR 146131 and the antagonist SR 27897 occupy different sites on the human CCK(1) receptor. *Eur. J. Pharmacol.* **2000**, *400*, 185−194.

JM049886Y