# Direct Identification of the Agonist Binding Site in the Human Brain Cholecystokinin$_B$ Receptor†

Jonas Anders,‡,§ Martin Blüggel,|| Helmut E. Meyer,|| Ronald Kühne,⊥ Antonius M. ter Laak,⊥ Elzbieta Kojro,‡ and Falk Fahrenholz*,‡

*Institut für Biochemie, Johannes Gutenberg Universität Mainz, Becherweg 30, D-55099 Mainz, Germany, Institut für Immunologie, Medizinisches Proteom Center, Ruhr-Universität Bochum, Universitätsstrasse 150, D-44780 Bochum, Germany, and Forschungsinstitut für Molekulare Pharmakologie, Alfred Kowalke Strasse 4, D-10315 Berlin, Germany*

*Received February 4, 1999; Revised Manuscript Received March 9, 1999*

ABSTRACT: In investigating the agonist binding site of the human brain cholecystokinin$_B$ receptor (CCK$_B$R), we employed the direct protein chemical approach using a photoreactive tritiated analogue of sulfated cholecystokinin octapeptide, which contains the *p*-benzoylbenzoyl moiety at the N-terminus, followed by purification of the affinity-labeled receptor to homogeneity. This probe bound specifically, saturably, and with high affinity ($K_D$ = 1.2 nM) to the CCK$_B$R and has full agonistic activity. As the starting material for receptor purification, we used stably transfected HEK 293 cells overexpressing functional CCK$_B$R. Covalent labeling of the WGA−lectin-enriched receptor revealed a 70−80 kDa glycoprotein with a protein core of about 50 kDa. Identification of the agonist binding site was achieved by the application of subsequent chemical and enzymatical cleavage to the purified receptor. A radiolabeled peptide was identified by Edman degradation amino acid sequence analysis combined with MALDI-TOF mass spectrometry. The position of the radioactive probe within the identified peptide was determined using combined tandem electrospray mass spectrometry and peptide mapping. The probe was covalently attached within the sequence L$^{52}$ELAIRITLY$^{61}$ that represents the transition between the N-terminal domain and predicted transmembrane domain 1. Using this interaction as a constraint to orientate the ligand within the putative receptor binding site, a model of the CCK-8s-occupied CCK$_B$R was constructed. The hormone was found to be placed in a binding pocket built from both extracellular and transmembrane domains of CCK$_B$R with its N-terminus mainly interacting with residues Arg$^{57}$ and Tyr$^{61}$.

Cholecystokinin (CCK)[1] plays an important physiological role both in the gastrointestinal tract and in the central nervous system. So far, two CCK receptor subtypes have been described that could be distinguished by their affinities for gastric peptides and for different forms of CCK. The CCK$_A$ receptor (CCK$_A$R) exhibits high-affinity binding only for sulfated forms of CCK and weak affinity for gastric peptides and nonsulfated CCK. In contrast, the CCK$_B$ receptor (CCK$_B$R) has less extensive structural requirements for its ligands, since sulfated and nonsulfated forms of CCK and gastric peptides are bound with comparable high affinities (reviewed in ref *1*).

The CCK$_B$R is a particularly important drug target because of its widespread expression throughout the central nervous system and in certain peripheral tissues, including stomach, pancreas, and smooth muscle. In the central nervous system, the major form of CCK is the C-terminally sulfated cholecystokinin octapeptide (CCK-8s) and the CCK$_B$R is the predominant subtype with the highest receptor densities found in the cortex regions and nucleus caudatus (*2*). Binding of CCK to the CCK$_B$R in the central nervous system increases the rate of release of dopamine and causes interference with the action of opioids where CCK is suggested to be a specific opiate antagonist in the analgesia-mediating system implicated in the perception of pain (*3*−*5*). CCK is colocalized with the neurotransmitters serotonin and $\gamma$-aminobutyric acid, and accordingly, it mediates anxiety and panicogenic effects (*6*, *7*).

Both CCK$_A$R and CCK$_B$R have been cloned (*8*−*10*) and share approximately 50% amino acid identity. Despite their different agonist selectivities, the two receptors use the same signal transduction mechanism: activation of phospholipase C with release of inositol trisphosphate and diacylglycerol ultimately leading to an increase in the intracellular $Ca^{2+}$ concentration (*10*, *11*). Both CCK receptor subtypes are

† This work was supported by the Fonds der Chemischen Industrie and by the Naturwissenschaftlich-Medizinisches Forschungszentrum, Johannes Gutenberg Universität Mainz.

* To whom correspondence should be addressed. E-mail: ibc1950@mail.uni-mainz.de. Fax: +49-6131-395348. Phone: +49-6131-395833.

‡ Johannes Gutenberg Universität Mainz.

§ Present address: Division of Molecular Neurobiology, Department of Neuroscience, Karolinska Institute, Doktorsringen 12 B Plan 3, S-17177 Stockholm, Sweden.

|| Ruhr-Universität Bochum.

⊥ Forschungsinstitut für Molekulare Pharmakologie.

[1] Abbreviations: CCK, cholecystokinin; CCK-8s, sulfated cholecystokinin octapeptide; CCK$_A$R, cholecystokinin$_A$ receptor; CCK$_B$R, cholecystokinin$_B$ receptor; EDTA, ethylenediaminetetraacetic acid; EGTA, ethylene glycol bis($\beta$-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid; FPLC, fast protein liquid chromatography; HPLC, high-performance liquid chromatography; MALDI-TOF mass spectrometry, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid.

10.1021/bi990269z CCC: $18.00 
Published on Web 04/16/1999

members of the family of rhodopsin-like seven-transmembrane domain G-protein-coupled receptors (GPCRs). Within this broad group of GPCRs, it is known that residues belonging to the transmembrane domains of biogenic amine receptors are the primary determinants of agonist binding (*12*, *13*). For protein and peptide ligands, both transmembrane and extracellular domains of the receptors are reported to contribute to the ligand binding site (*14*−*17*). Understanding the molecular interactions involved in agonist binding provides insights that could be helpful in targeted drug design and in understanding the basic mechanisms of GPCR activation. The emphasis is on the question about the molecular basis for two receptor subtypes with different agonist binding selectivities that induce the same signal transduction mechanisms.

The current insights into the agonist binding domains of the $CCK_BR$ were obtained exclusively from mutagenesis studies, and multiple contributions of both transmembrane domains (TMs) and extracellular loops (ECLs) have been reported. Residues responsible for the high-affinity binding of gastric peptides to the $CCK_BR$ were found to be located in the second extracellular loop (*18*), and critical residues for high-affinity CCK-8s binding were identified in ECL 1 and on top of TM 1 (*19*) while mutations of transmembrane domain residues interfere only weakly with agonist binding (*20*−*23*). However, the data currently available from mutagenesis experiments represent mainly contributions of receptor regions to agonist selectivity or affinity; thus, these studies lack direct structure information about the agonist−receptor complex.

In this report, we present the protein chemical approach to gaining additional and direct structure information about the interaction between the $CCK_B$ receptor and its agonists. The agonist binding site of $CCK_BR$ was probed using a tritiated photoreactive CCK analogue [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s, which represents a full agonist on the recombinant $CCK_B$ receptor. The sequence $L^{52}$-$ELAIRITLY^{61}$ in $CCK_BR$ that forms the transition between the N-terminal domain and the predicted first transmembrane domain was clearly shown to be the contact site between the N-terminus of the probe and the $CCK_BR$. Using this interaction as a constraint to orientate the ligand within the putative receptor binding site, a model of the CCK-8s-occupied $CCK_BR$ was constructed, including additional information from site-directed mutagenesis studies.

## MATERIALS AND METHODS

### *Reagents*

The reagents used were from the following suppliers: fetal calf serum, PAA Laboratories; dodecyl $\beta$-D-maltopyranoside, Anatrace Inc.; *N*-hydroxysuccinimidyl [$^3$H]propionate, Amersham Buchler; *n*-octyl glucoside, Calbiochem; cholecystokinin octapeptide, Bachem; and Fura-2AM and cyanogen bromide, Fluka. All proteinases were from Boehringer Mannheim, and the photoreactive tritiated CCK analogue [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s was synthesized and purified as described previously (*24*).

### *Generation of Stably Transfected HEK 293 Cells*

A stably transfected HEK 293 cell line expressing the human brain $CCK_B$ receptor was constructed as described recently (*25*). In brief, the receptor cDNA was fused with an additional *c-myc* tag to allow immunodetection, cloned into the vector pCDNA3, and used for transfection of cells. Selection of stable transfectants was performed using 1 mg/mL G418 as a selecting antibiotic until a single clone could be isolated which was designated as HEK 293-$CCK_BR$. Cells were maintained in complete Dulbecco's modified Eagle's medium (DMEM) supplemented with penicillin, streptomycin, and 10% (w/v) fetal calf serum.

### *Membrane Preparation*

Stably transfected HEK 293 cells expressing the *c-myc*-tagged human brain $CCK_B$ receptor (*25*) were grown to confluence and then washed with ice-cold PBS/10 mM EDTA. After incubation for 5 min in PBS/10 mM EDTA, the cells were lifted mechanically and collected by sedimentation (800*g* at room temperature for 10 min). The cell pellet was washed (2 × 10 mL of PBS/10 mM EDTA) and finally resuspended at a cell density of $5 \times 10^7$ cells/mL in lysis buffer [10 mM Hepes/NaOH (pH 7.4)] and protease inhibitors 10 mg/mL benzamidine, 10 $\mu$g/mL leupeptin, 0.1 mg/mL Pefabloc, 10 $\mu$g/mL 1,10-phenanthroline, 1 mM EGTA, 0.1 mg/mL bacitracin, 10 $\mu$g/mL SBTI type I, and 5 $\mu$g/mL pepstatin A (Pefabloc from Boehringer Mannheim, all others from Sigma). After the mixture had been shaken for 15 min at 4 °C, cell homogenization was performed by a polytron ultrathorax tissue homogenizer at the maximum setting with three pulses of 15 s at 4 °C. Membranes were collected by centrifugation (32000*g* for 30 min at 4 °C), resuspended in binding buffer [10 mM Hepes/NaOH (pH 7.4), 120 mM NaCl, 5 mM $MgCl_2$, and 1 mM EGTA], and stored at −70 °C.

### *Receptor Binding Assays*

For saturation studies, membranes were incubated with increasing concentrations of radioligand in a total volume of 100 $\mu$L of binding buffer for 30 min at 30 °C. Then the membranes were washed three times with 3 mL of each filtration buffer [10 mM Hepes (pH 7.4)]. The filtration buffer for the solubilized fraction also contained 7% (w/v) polyethylene glycol (PEG) 8000. Prior to filtration, the solubilized proteins were precipitated by adding 0.075% (w/v) bovine $\gamma$-globulin and 7% (w/v) polyethylene glycol 8000 for 15 min on ice. The amount of bound radioactivity was separated from the amount of free radioactivity by rapid filtration over Whatman GF/C filters using a Brandel cell harvester. Filters were washed twice with the corresponding filtration buffers, placed in scintillation vials, and made transparent with 1.5 mL of ethylene glycol monomethyl ether. After 15 min, 10 mL of scintillation cocktail (Rotiszint eco plus) was added. The amount of radioactivity was measured in a Packard Tri-Carb 2100 TR liquid scintillation counter. Nonspecific binding was assessed in the presence of a 200-fold excess of unlabeled cholecystokinin octapeptide. Data analysis for saturation experiments was carried out using the LIGAND program (*26*).

### *Probing the Biological Activity of [$^3$H]-p-Benzoylbenzoyl-Orn(propionyl)-CCK-8s*

HEK 293 cells stably transfected with the *c-myc*-tagged $CCK_BR$ (*25*) were grown to 70% confluence and loaded with

2 $\mu$M Fura-2AM (Fluka) for 40 min at 37 °C. Then the cells were washed with HBS [10 mM Hepes/NaOH (pH 7.4), 140 mM NaCl, 5 mM KCl, 0.5 mM $MgCl_2$, and 2 mM $CaCl_2$] and were finally suspended in HBS with additional 10 mM glucose. Samples (3.0 mL) of the cell suspension equal to about $4 \times 10^5$ cells were placed in a stirred thermostated quartz cuvette in a PTI 810 spectrophotometer at 37 °C, and increasing concentrations (from $10^{-15}$ to $10^{-6}$ M) of either CCK-8s (Bachem) or *p*-benzoylbenzoyl-Orn-CCK-8s, the nonradioactive precursor of [$^3$H]-*p*-benzoylbenzoyl-Orn-(propionyl)-CCK-8s, were added. Fluorescence was measured by ratiometry at 500 nm after excitation at 340 and 380 nm. The intracellular $Ca^{2+}$ concentration was calculated as described (*27*) after calibration using 0.1% (w/v) Triton X-100 to permeabilize cells to determine the maximum fluorescence ratio. The minimum fluorescence ratio was determined after additionally adding 25 mM EGTA (pH 8.0) and 0.005 N NaOH. All data were corrected against the autofluorescence of unloaded cells.

*Photoaffinity Labeling*

The tritiated photoaffinity labeling probe [$^3$H]-*p*-benzoyl-benzoyl-Orn(propionyl)-CCK-8s was synthesized, and photoaffinity labeling of the $CCK_BR$ was performed as described previously (*24*). In brief, the probe was bound to equilibrium (40 min, 30 °C) to the proteins in binding buffer containing the following additional protease inhibitors: 10 mg/mL benzamidine, 10 $\mu$g/mL leupeptin, 0.1 mg/mL Pefabloc, 10 $\mu$g/mL 1,10-phenanthroline, 1 mM EGTA, 0.1 mg/mL bacitracin, 10 $\mu$g/mL SBTI type I, 5 $\mu$g/mL pepstatin A. Then the sample was photolyzed using a 200 W mercury lamp ($\lambda > 320$ nm) in a thermostated cuvette at 4 °C for 30 min. Nonspecific labeling was assessed in the presence of 1 $\mu$M CCK-8s. To subject the samples to analytical SDS−PAGE according to the method of Laemmli (*28*), proteins from photoaffinity labeling experiments were precipitated using the chloroform/methanol method (*29*). After electrophoresis, the gels were cut into 2 mm slices which were treated with 0.5 mL of lumasolve (Packard) each. The amount of radioactivity was determined by liquid scintillation counting using 4 mL of lipoluma (Packard) as the scintillator.

*Enzymatical Deglycosylation*

In the case of receptor deglycosylation, a total of 50 $\mu$g of WGA−lectin-enriched and photoaffinity-labeled membrane protein was collected by precipitation using the chloroform/methanol method (*29*). Pellets were dissolved in 10 $\mu$L of 10% (w/v) SDS by incubating them for 15 min at room temperature and finally diluted to a volume of 200 $\mu$L with deglycosylation buffer containing 50 mM sodium phosphate (pH 7.5), 1% (w/v) NP-40, and 1% (v/v) $\beta$-mercaptoethanol. In the case of cyanogen bromide, cleavage product samples after purification by reversed-phase HPLC were at first lyophilized to evaporate the excess of cyanogen bromide, and the residue was dissolved in deglycosylation buffer. The reaction was started by adding the indicated amounts of *N*-glycosidase F, and the mixture was incubated at 25 °C for 18 h. Then the samples were mixed with equal volumes of 2× SDS−polyacrylamide gel electrophoresis (PAGE) sample buffer [8% (w/v) SDS, 24% (v/v) glycerol, 50 mM Tris-HCl (pH 6.7), 4% (v/v) $\beta$-mercaptoethanol, and 0.01% (w/v) Serva Blue G]. Samples were subjected to 9% SDS−PAGE according to the method of Laemmli (*28*) in the case of the receptor and 16.5% T/6% C Tris/Tricine/urea SDS−PAGE (*30*) in the case of cyanogen bromide fragments. After electrophoresis, gels were sliced and the amount of radioactivity was determined as described above.

*Cyanogen Bromide Cleavage*

Proteins after gel filtration chromatography or HPEC were collected by the chloroform/methanol procedure (*29*), and the dried pellet was dissolved in 50 $\mu$L of 70% (v/v) formic acid and 5% (v/v) $\beta$-mercaptoethanol. After the addition of 5 mg of crystalline cyanogen bromide, the sample was incubated at 20 °C for 20 h under a nitrogen atmosphere in the dark. The labeled peptide was identified by subjecting the cleavage fragments to 16.5% T/6% C Tris/Tricine/urea SDS−PAGE (*30*) with or without prior enzymatical deglycosylation. After electrophoresis, the gels were cut into 2 mm slices which were treated with 0.5 mL of lumasolve (Packard) each. The amount of radioactivity was determined by liquid scintillation counting using lipoluma (Packard) as the scintillator.

*V8 Protease Peptide Mapping*

For peptide mapping, the radioactive photoaffinity-labeled $CCK_B$ receptor after gel filtration chromatography was subjected to cyanogen bromide cleavage as described above, and the radiolabeled peptide was purified by HPLC (see below) and lyophilized and the residue dissolved in 100 mM $(NH_4)_2CO_3$ and 1.0% (w/v) *n*-octyl glucoside. Analytical cleavage by V8 protease was performed by diluting portions of this sample 40-fold with 100 mM $(NH_4)_2CO_3$ to decrease the detergent concentration to 0.025% (w/v) and finally adding V8 protease (sequencing grade, Boehringer Mannheim) at a concentration of 100 ng/$\mu$L. Samples were then incubated for 24 h at 37 °C. Enzymatical reactions were stopped by adding the same volume of SDS sample buffer and subjecting the fragments to analytical Tris/Tricine/urea SDS−PAGE, gel slicing, and $\beta$-scintillation counting as described above.

*Purification of the Photoaffinity-Labeled $CCK_B$ Receptor*

Prior to purification, the $CCK_BR$ was solubilized using the nonionic detergent dodecyl $\beta$-D-maltopyranoside as described previously (*31*). Starting with portions of 400 mg protein in membranes of stably transfected HEK 293 cells (*25*), we obtained 150 mg of the solubilized protein.

*Lectin Affinity Chromatography.* The detergent soluble membrane proteins were at first separated by wheat germ agglutinin (WGA) affinity chromatography. A column (18 cm × 1.5 cm) was filled with WGA−agarose (Pharmacia) at hydrostatic pressure and was equilibrated with starting buffer [10 mM Hepes (pH 7.4), 120 mM NaCl, 5 mM $MgCl_2$, 10% (v/v) glycerol, and 0.1% (w/v) dodecyl $\beta$-D-maltopyranoside]. Solubilized membranes (150 mg of protein in 80 mL of buffer) were diluted 6-fold with starting buffer to obtain a detergent concentration of 0.1% (w/v) and were then subjected to the column at 4 °C, and a linear flow rate of 0.25 mL/min was applied. After extensive washing with starting buffer, receptor-containing fractions were eluted

using 0.6 M *N*-acetylglucosamine in starting buffer as a competitor. Peak fractions containing radioligand binding activity were pooled and subsequently subjected to photoaffinity labeling with [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s as described above.

*Gel Filtration Chromatography.* To separate the noncovalently bound excess of the probe from the photoaffinity-labeled $CCK_BR$, gel filtration chromatography was employed as a second purification step. A Superdex 200 column (2 cm × 60 cm) connected to an FPLC system (Pharmacia) was equilibrated with buffer [25 mM Tris-HCl (pH 7.0), 150 mM NaCl, 1 mM EGTA, and 2% (w/v) SDS] at room temperature, and proteins were eluted with 400 mL of buffer at a linear flow rate of 1 mL/min at room temperature. The column eluate was fractionated (5 mL fractions), and 50 $\mu$L portions were taken for scintillation counting. Peak fractions containing high-molecular mass radioactivity were pooled, and the photoaffinity-labeled $CCK_BR$ was identified by subsequent analytical SDS−PAGE, gel slicing, and $\beta$-scintillation counting as described above. The procedure starting with the solubilization of a portion of 400 mg of membrane protein in cell membranes was repeated twice, and both eluates after gel filtration chromatography were pooled prior to further purification of the photoaffinity-labeled $CCK_B$ receptor.

*High-Performance Electrophoresis Chromatography (HPEC).* HPEC was performed on an Applied Biosystems model 230A device. For each HPEC run, the protein (500 $\mu$g) from 5 mL of a protein solution obtained after gel filtration chromatography was collected by the chloroform/methanol method (*31*). The protein pellet was incubated for 30 min with 10 $\mu$L of 10% (w/v) SDS, and then 140 $\mu$L of sample buffer [15 mM Tris-phosphate (pH 7.5), 2% (w/v) SDS, 0.4% (v/v) $\beta$-mercaptoethanol, and 30% (v/v) glycerol] was added and the sample incubated for another 30 min at room temperature. Insoluble material was separated by centrifugation at 10000*g* for 10 min. One hundred fifty microliters of the supernatant from this centrifugation was applied on a calibrated acrylamide tube gel in the HPEC system. Protein separation was performed with a Tris-phosphate/SDS buffer system at pH 7.5 according to the protocol of Applied Biosystems; continuous 6% polyacrylamide tube gels (3.5 mm × 100 mm) were used. The gels were prerun for 60 min with a constant current of 0.5 mA and then for 240 min with a current of 2.5 mA. Calibrating runs with standard proteins and protein separation runs were performed for 30 min with a constant current of 0.5 mA and then for 600 min with a constant current of 2.5 mA. All electrophoresis runs were performed at 8 °C. Proteins were eluted from the gel at a flow rate of 10 $\mu$L/min and detected at 220 nm. Protein fractions were collected each 5 min at 4 °C; 1 $\mu$L portions were taken both for scintillation counting and for SDS−PAGE analysis for analysis of the protein composition (minigels and silver staining). Relevant peak fractions were stored at −20 °C and prepared for chemical and enzymatical cleavage.

### *Purification of the Agonist Binding Domain of the $CCK_B$ Receptor*

The purified affinity-labeled receptor from nine HPEC runs was collected by the chloroform/methanol procedure (*31*) and subjected to cyanogen bromide cleavage as described above. Cyanogen bromide fragments were separated by reversed-phase high-performance liquid chromatography prior to further tryptic cleavage. This was achieved using a Varian system equipped with a $C_4$ reversed-phase analytical column (Vydac 215TP5415, 4.6 mm × 250 mm) using a linear flow rate of 1.5 mL/min at room temperature and a gradient of increasing concentrations of 50% (v/v) acetonitrile in 2-propanol (buffer B) with a background of 0.1% (v/v) trifluoroacetic acid. Buffer A was water containing 0.1% (v/v) trifluoroacetic acid. The gradient consisted of the following: 20% buffer B over the course of 10 min, advancing to 100% buffer B over the course of 60 min, and holding at 100% buffer B for an additional 10 min. Peak radioactive fractions were collected, and the identity of the isolated peptide was again examined by enzymatical deglycosylation and SDS−PAGE as described above. Pooled eluate fractions containing radioactivity were dried in vacuo and then dissolved in 50 $\mu$L of 100 mM $(NH_4)CO_3$ containing 1% (w/v) *n*-octyl glucoside. Trypsin (from bovine pancreas, sequencing grade, Boehringer Mannheim) was added to a final concentration of 100 ng/$\mu$L, and the sample was incubated for 5 h at 37 °C. Separation of tryptic peptides was achieved by reversed-phase HPLC. One microliter of TFA was added to the digest prior to application on a $C_4$ reversed-phase microbore column (Vydac 214TP215, 2.1 mm × 150 mm). Tryptic peptides were separated using a linear flow rate of 200 $\mu$L/min at 25 °C and a gradient of increasing concentrations of 2-propanol (buffer B) with a background of 0.1% (v/v) trifluoroacetic acid. Buffer A was water containing 0.1% (v/v) trifluoroacetic acid. The gradient consisted of the following: 20% buffer B over the course of 10 min, advancing to 30% buffer B over the course of 20 min and then to 100% buffer B over the course of 45 min, and holding at 100% buffer B for an additional 10 min. The relevant peak fractions were dried in vacuo, stored at −20 °C, and prepared for N-terminal Edman sequencing and mass spectrometry.

### *Edman Degradation Amino Acid Sequence Analysis*

The purified radioactive peptide (0.6 pmol of radioactivity) was dissolved in 5 $\mu$L of 70% (v/v) HCOOH/20% (v/v) 2-propanol/10% (v/v) water and was applied to a Biobrene-coated filter. Sequencing was performed in an Applied Biosystems Procise 494 CLC Sequencer in pulsed-liquid mode according to the protocol from Applied Biosystems.

### *MALDI-TOF Mass Analysis*

The purified radioactive peptide (0.2 pmol of radioactivity) was dissolved in 1 $\mu$L of 80% (v/v) HCOOH, and 5 $\mu$L of a saturated solution of α-cyano-4-hydroxycinnamic acid in 0.05% (v/v) TFA/50% (v/v) acetonitrile/49.5% (v/v) water was added. The solution was applied to the target of the mass spectrometer and dried. Mass analysis was performed using a Bruker Reflex III instrument in reflectron mode. The acceleration voltage was 20 kV; the energy of the nitrogen laser ($\lambda$ = 337 nm) was attenuated by 79%. All signals with an *m*/*z* of <400 Da were faded out. One analysis was composed of 200 single experiments with a sampling rate of 2 GHz.

*Molecular Modeling of the $CCK_B$ Receptor−CCK-8s Complex*

The starting structure of the human cholecystokinin$_B$ ($CCK_B$) transmembrane domains has been built using the α-carbon template of the transmembrane helices of rhodopsin (*32*). This template is based on a three-dimensional electron density map of frog rhodopsin with an effective resolution of 7.5 Å in the membrane plane (*33*). It is known that the extracellular part of the human $CCK_B$ receptor plays an important role in receptor−ligand interactions (*20*, *34*, *35*). Therefore, we have completed our receptor model with the extracellular loops and a reduced N-terminus ($Cys^{22}$−$Leu^{54}$) using the loop search tool in SYBYL 6.4 (version 6.4, Tripos Inc., St. Louis, MO). *N*-Acetyl (ACE) and *N*-methyl (NME) groups were added to all N- and C-termini. The obtained $CCK_B$ model was minimized using the sander module in AMBER 4.1 (*36*). The $CCK_B$ agonist CCK-8s was docked manually into the putative binding site. According to the experimental results described in this report, the N-terminal sequence of CCK-8s was located near the upper part of TM 1 ($L^{52}$ELAIRITLY$^{61}$). On the basis of information from site-directed mutagenesis studies (*20*, *23*), the C-terminus of CCK-8s was arranged near $Ser^{219}$ and the hydrophobic residues of CCK-8s ($Trp^5$, $Met^6$, and $Phe^8$) were buried in the hydrophobic core built by TM 4−6. Bad contacts were removed using the docking tool in SYBYL 6.4. The atomic charges of the sulfated tyrosine were calculated using the RESP charge calculation formalism (*37*) with ab initio-derived [basis set 6G31*, program GAMESS_US (*38*)] molecular electrostatic potentials [MOLDEN 3.2 (*39*)] as input. To obtain a relaxed starting structure for molecular dynamics (MD), the receptor−CCK-8s complex was minimized again with the AMBER 4.1 force field. MD simulation of the $CCK_BR$−CCK-8s complex was carried out at 300 K for 250 ps in vacuo using the SANDER module of AMBER 4.1. The lack of intracellular loops and the C-terminal part requires the definition of structural restraints. To stabilize the helical conformation of transmembrane domains, all helical hydrogen bonds were restrained using distance constraints with a force constant of 10 kcal mol$^{-1}$ Å$^{-2}$. The arrangement of the seven-helix bundle was fixed by a set of distance restraints between the centers of three Cα atoms in each TM domain (TM 1, $Ser^{68}$, $Met^{67}$, and $Leu^{66}$; TM 2, $Asp^{100}$, $Ser^{99}$, and $Val^{98}$; TM 3, $Val^{138}$, $Ser^{137}$, and $Val^{136}$; TM 4, $Pro^{188}$, $Val^{187}$, and $Met^{186}$; TM 5, $Pro^{230}$, $Ile^{229}$, and $Phe^{228}$; TM 6, $Pro^{348}$, $Leu^{347}$, and $Trp^{346}$; and TM 7, $Pro^{387}$, $Asn^{386}$, and $Val^{385}$). This allows rotations of the TM domains around their own axis and enables helix kinks near prolines.

## RESULTS

*Biological Activity of the Photoreactive Probe [$^3$H]-p-Benzoylbenzoyl-Orn(propionyl)-CCK-8s*

The tritiated photoreactive probe [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s (Figure 1) bound to the membranes of stably transfected HEK 293 cells (*25*) bearing the *c-myc*-tagged human brain $CCK_BR$ saturably with high affinity ($K_D$ = 1.2 ± 0.1 nM), which was identical to the affinity of the radiolabeled physiological agonist [$^3$H]propionyl-CCK-8s for the wild-type $CCK_BR$ in human frontal cortex [$K_D$ = 1.1 nM (*2*)]. The receptor density in membranes derived from stably transfected HEK 293 cells was about 1000-fold higher than in natural tissues, and 8.4 ± 0.3 pmol/mg of membrane protein was found using [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s as a probe.

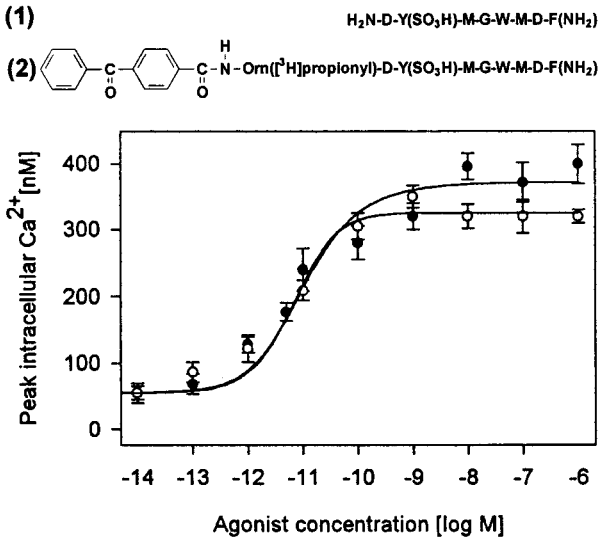


FIGURE 1: Structure and biological activity of the photoreactive CCK receptor probe [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s. The ability of the nonradioactive synthetic precursor of [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s, *p*-benzoylbenzoyl-Orn-CCK-8s (●), to stimulate an increase in the intracellular calcium concentration in HEK 293 cells stably transfected with the *c-myc*-tagged $CCK_BR$ in comparison with that of CCK-8s (○) is shown as means ± SE for three identical experiments performed in duplicate. Also shown is the structural formula of [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s (2) in comparison with that of the physiological agonist of the $CCK_BR$ in human brain, CCK-8s (1).

The nonradioactive synthetic precursor of the probe, *p*-benzoylbenzoyl-Orn-CCK-8s, represented a full agonist acting at the recombinant $CCK_BR$ by stimulating an increase in the intracellular concentration of $Ca^{2+}$ in a dose-dependent manner (Figure 1). The effect of the photoreactive probe was again nearly identical to that of the physiological agonist CCK-8s. $EC_{50}$ values for stimulation of the half-maximal concentration of intracellular $Ca^{2+}$ were (6.4 ± 0.5) × $10^{-12}$ M in the case of CCK-8s and (11.5 ± 0.9) × $10^{-12}$ M in the case of *p*-benzoylbenzoyl-Orn-CCK-8s. This indicates that residues of the receptor involved in the binding of [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s are part of the agonist binding site of the $CCK_BR$ and are implicated in signal transduction.

*Purification of the Photoaffinity-Labeled $CCK_B$ Receptor*

*Solubilization*. The $CCK_BR$ was solubilized using the nonionic detergent dodecyl β-D-maltopyranoside (Dβ M). The protein yield was 36 ± 2% of the total membrane protein, and the receptor yield was 86 ± 3% as determined by the radioligand binding test using [$^3$H]-*p*-benzoylbenzoyl-Orn-(propionyl)-CCK-8s as a probe. The probe binding affinity of the $CCK_BR$ was nearly not affected by this procedure ($K_D$ = 1.7 ± 0.4 nM), and the maximum binding capacity increased 2.4-fold, reaching 19.6 ± 1.2 pmol/mg of membrane protein.

*Lectin Affinity Chromatography*. Solubilized receptors were at first enriched using immobilized wheat germ

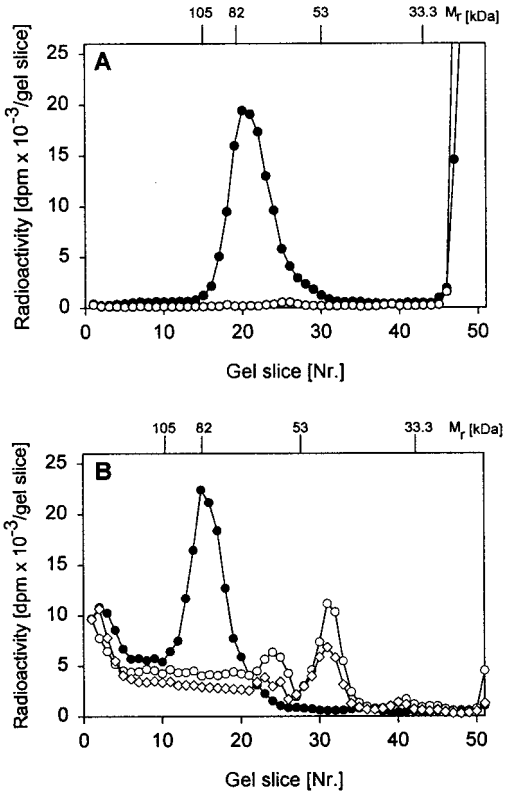


FIGURE 2: Photoaffinity labeling (A) and enzymatical deglycosylation (B) of the WGA−lectin-enriched $CCK_B$ receptor. (A) The pooled fractions after lectin chromatography were subjected to photoaffinity labeling with the photoreactive CCK analogue [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s in the absence (●) and presence (○) of the nonradioactive competitor CCK-8s at 1 $\mu$M. (B) Enzymatical deglycosylation was performed by adding 0 (●), 100 (○), and 250 (◇) units of *N*-glycosidase F as described in Materials and Methods. Shown are typical distributions of radioactivity in denaturing SDS−PAGE according to Laemmli (*28*) used to separate the products of photoaffinity labeling and deglycosylation of the $CCK_BR$ which are representative of three independent experiments.

agglutinin (WGA). In this manner, the receptor could be enriched 6-fold while retaining the high-affinity binding for [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s ($K_D$ = 8.1 ± 1.4 nM) with a maximum binding capacity of 119 ± 8 pmol/mg of membrane protein and a yield of about 30% of the total binding sites.

*Photoaffinity Labeling*. In the solubilized and WGA-enriched membrane fraction, a protein with an apparent molecular mass of about 80 kDa was labeled with an overall yield of 30 ± 5% of total binding sites (Figure 2A). Nonspecific labeling contributed to less than 1% of total binding. The broad peak of the affinity-labeled protein indicated a heterogeneously glycosylated protein. By enzymatical deglycosylation, the apparent molecular mass of the labeled protein decreased to about 50 kDa (Figure 2B) which was clearly consistent with the calculated mass of the $CCK_BR$ protein core with the additional *c-myc* tag estimated from the nucleotide sequence (50.294 kDa). This indicated a carbohydrate moiety of the recombinant receptor with a molecular mass of about 30 kDa.

*Gel Filtration Chromatography*. Gel filtration chromatography in the presence of D$\beta$M led to the formation of aggregates that eluted in the void volume of the column

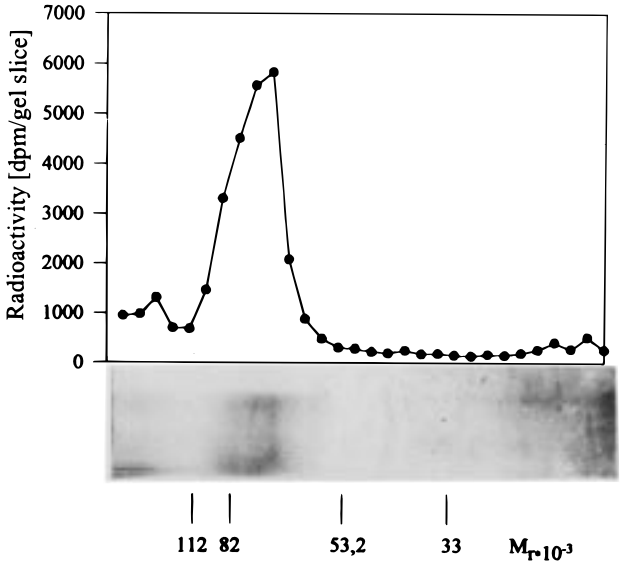


FIGURE 3: SDS−PAGE analysis of the purified photoaffinity-labeled human brain $CCK_B$ receptor. A portion of pooled eluate from nine HPEC runs was subjected to 9% denaturing SDS−PAGE analysis, and the gel was silver stained or sliced for $\beta$-scintillation counting as described in Materials and Methods.

containing both the receptor and free probe. A strongly denaturing detergent was necessary for obtaining suitable molecular mass resolution and recovery of proteins. Therefore, D$\beta$M was replaced by SDS. Two radioactive peaks were observed, representing free ligand and the photoaffinity-labeled receptor with the latter eluting in the molecular mass range between 44 and 158 kDa. The yield of the affinity-labeled receptor was 90% with a 3-fold enrichment of the protein. The eluate fractions containing high-molecular mass radioactivity were pooled and subjected to further purification.

*High-Performance Electrophoresis Chromatography (HPEC)*. The pooled samples after two runs of gel filtration chromatography were finally fractionated by high-performance electrophoresis chromatography (HPEC). By this, the photoaffinity-labeled $CCK_BR$ from the gel filtration eluate was enriched ~54-fold with a 26% yield. Fractions containing pure $CCK_BR$ were combined from nine HPEC runs, and the purified receptor was visualized by analytical SDS−PAGE and silver staining (Figure 3), showing the receptor to be purified to homogeneity. The receptor was present as a heterogeneously glycosylated protein with a molecular mass of about 80 kDa. Nearly all the radioactivity in the gel was due to the photoaffinity-labeled purified receptor; a smaller radioactive peak at a higher molecular mass indicated the minor formation of some aggregates. The amount of total receptor-bound radioactivity in the combined eluate fractions from nine HPEC runs corresponded to 120 pmol of the covalent radioactive $CCK_B$ receptor complex, which is 6 $\mu$g of protein. Since it was not possible to distinguish between the excess of receptor, which did not react with the photoreactive probe, and the covalently labeled, radioactive receptor (30 ± 5% of the total), the fact that the protein amount in the purified sample was considerably higher than 6 $\mu$g must be taken into account. As summarized in Table 1, the photoaffinity-labeled $CCK_BR$ was obtained with an overall yield of 1.8%.

Table 1: Purification of the Recombinant Human Brain $CCK_BR$[a]

| sample | protein (mg) | receptor pmol | % yield | pmol/mg | purification (*x*-fold) |
|---|---|---|---|---|---|
| membrane suspension | 800 | 6720 | 100 | 8.4 | 1 |
| detergent extract | 298 | 5833 | 86.8 | 19.6 | 2.3 |
| WGA eluate | 14.7 | 1747 | 26 | 119 | 14.2 |
| photoaffinity-labeled WGA eluate | 14.7 | 524[b] | 7.8[b] | nd | nd |
| Superdex eluate | 4.8 | 472[b] | 7[b] | nd | nd |
| HPEC eluate | nd[c] | 120[b] | ~1.8[b] | nd | nd |

[a] Experimental procedures are described in Materials and Methods. [b] Corresponding to the photoaffinity-labeled receptor. [c] nd, not determined.

### *Identification of the Agonist Binding Site*

*Cyanogen Bromide Cleavage.* Chemical cleavage of the photoaffinity-labeled $CCK_BR$ with cyanogen bromide yielded a fragment migrating at an apparent molecular mass of about 35−40 kDa by comparison with standard proteins and peptides (Figure 4A), which was reduced to about 8 kDa by enzymatical deglycosylation. The carbohydrate moiety of the glycosylated fragment contributed about 30 kDa in mass which was consistent with the molecular mass shift occurring during deglycosylation of the intact receptor protein. There is only one possible cyanogen bromide fragment of the $CCK_BR$ that could account for these properties. The fragment extending from residue $Glu^2$ to $Met^{67}$ of the wild-type $CCK_BR$ contains all three putative sites for N-linked glycosylation at N7, N30, and N36 (*10*). The calculated molecular mass of the protein core of this fragment amounts to 6933 Da which was consistent with the observed value of the deglycosylated fragment considering the mass of the covalently attached photoreactive probe (759.8 Da). This indicated that the photoreactive probe was incorporated into the receptor somewhere in the sequence between residues $Glu^2$ and $Met^{67}$ which represent the N-terminal domain of the $CCK_BR$, including the extracellular half of the predicted first transmembrane domain (Figure 4B).

*Isolation of the Agonist Binding Domain.* Five micrograms of the purified photoaffinity-labeled human brain $CCK_BR$ (100 pmol) was subjected to cyanogen bromide cleavage, and the 65-amino acid radiolabeled cyanogen bromide fragment was purified by analytical high-performance liquid chromatography (HPLC) (Figure 5A). The fragment containing the agonist binding site eluted in fractions 34−38 with an overall yield of about 10%, corresponding to 9.8 pmol of the radioactive peptide. The identity of the eluted radioactivity was examined by subsequent deglycosylation with *N*-glycosidase F and SDS−PAGE analysis, indicating that the properties of the radiolabeled cyanogen bromide fragment were not affected by HPLC purification (data not shown). After tryptic digestion of the purified cyanogen bromide fragment, a single radioactive peptide representing the agonist binding domain of the human brain $CCK_BR$ was isolated with a yield of 1.8 pmol, corresponding to an overall yield of 0.03% compared to cell membranes (Figure 5B).

*Edman Degradation Amino Acid Sequence Analysis.* The isolated peptide was identified by automated Edman degradation amino acid sequence analysis. The sequence that was identified was XXGTRELELAIRITLYXXIFL, with the X indicating residues which could not clearly be assigned either by contaminations in the two first cycles or by signal intensity in the latter. The sequence identified by comparison with the cDNA sequence of the $CCK_BR$ was GAGTRELELAIRITLYAVIFLM, corresponding to residues $Gly^{46}$−$Met^{67}$ representing the transition between the N-terminal domain and the predicted first transmembrane domain (Figure 4B). Tryptic cleavage of the precursor 65-amino acid cyanogen bromide fragment was incomplete since the isolated peptide still contained two arginine residues. The position of the incorporated radioactive probe could not be determined by amino acid sequencing. All radioactivity stayed bound to the Biobrene-coated filter which was caused by the insolubility of the probe-bound thiocarbamoyl derivative of the radio-

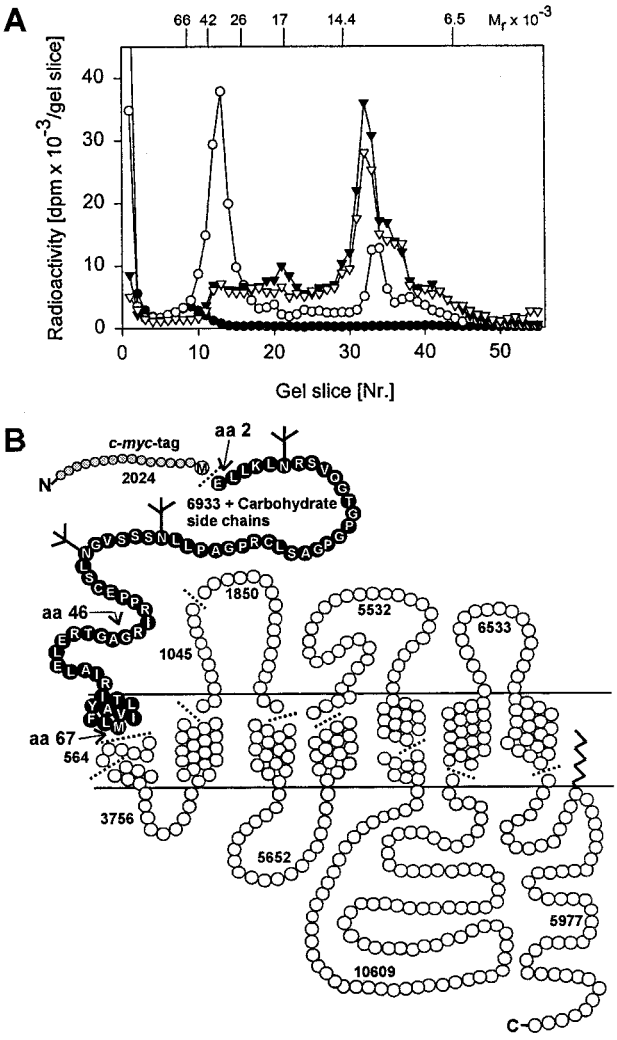


FIGURE 4: Cyanogen bromide cleavage of the photoaffinity-labeled human brain $CCK_B$ receptor. Shown is a typical distribution of radioactivity in a 16.5% T/6% C Tris/Tricine/urea SDS−PAGE analysis used to separate the products of CNBr digestion of the $CCK_BR$ with or without further enzymatical deglycosylation using 0 (○), 250 (▼), or 1000 units (▽) of *N*-glycosidase F (PNGAse F, NEB), which is representative of three independent experiments (A). As a control, the photoaffinity-labeled $CCK_BR$ without addition of CNBr was used (●). Also shown are the theoretical cleavage sites of the human $CCK_BR$, including the additional *c-myc* tag and the masses of the expected fragments independent of carbohydrate moieties and the mass of the eventually incorporated photoreactive probe (B). Only the fragment corresponding to residues $Glu^2$−$Met^{67}$ starting at the first residue of the receptor without numbering the additional *c-myc* tag is consistent with the migrations of glycosylated and deglycosylated fragments observed on the gel.

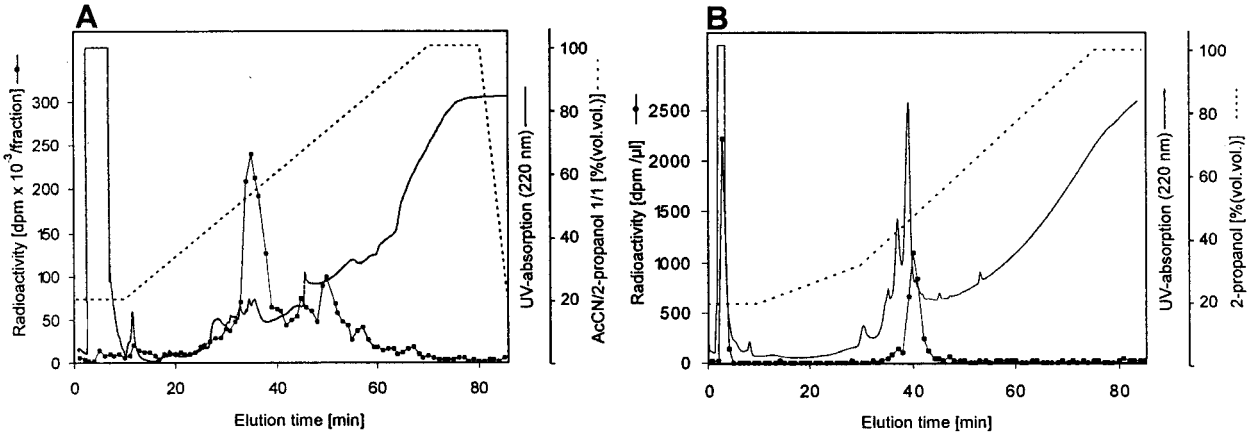


FIGURE 5: Reversed-phase HPLC purification of the radioactive photoaffinity-labeled agonist binding domain of the $CCK_B$ receptor. (A) One hundred picomoles of the purified photoaffinity-labeled $CCK_BR$ was subjected to cyanogen bromide cleavage, and the fragments were separated by chromatography on a $C_4$ reversed-phase analytical column (Vydac 215TP5415, 4.6 mm × 250 mm). (B) The main radioactive peak eluting in fractions 34−38 was pooled and subjected to tryptic cleavage. The tryptic digest was applied on a $C_4$ reversed-phase microbore column (Vydac 214TP215, 2.1 mm × 150 mm). A single radioactive peak eluting in fractions 40 and 41 was isolated and pooled for further characterization.

labeled amino acid(s) in *n*-butyl chloride, and the sequencers transfer solvent. The total peptide amount was determined with ~2 pmol by calculation of the initial yield, which indicated that unlabeled and labeled agonist binding domains were purified together.

*MALDI-TOF Mass Analysis.* The agonist binding site was further analyzed by MALDI-TOF mass analysis. Two series of peaks could be assigned to the photoaffinity-labeled and nonlabeled peptide GAGTRELELAIRITLYAVIFLM*, with M* denoting the carboxy-terminal homoserine resulting from cyanogen bromide cleavage of the photoaffinity-labeled $CCK_BR$ (Figure 6A). Mass analysis of the agonist binding domain was impeded by splitting of the masses of each sequence into formylated or lactone species, which considerably decreased the signal intensity. Since peaks corresponding to the affinity-labeled agonist binding domain appeared with an intensity near the threshold, it was difficult to assign monoisotopic masses, and therefore, average masses were determined.

The signal at *m*/*z* 2403.2 (monoisotopic) shown in the enlarged section of Figure 6B corresponded to the peptide GAGTRELELAIRITLYAVIFLM*-lactone (calculated monoisotopic mass of 2402.4 Da), with the carboxy-terminal homoserine as lactone. A series of peaks derived from this mass indicated formylated and homoserine species of this peptide (Figure 6B). By hydrolysis of the carboxy-terminal homoserine lactone during purification in aqueous acidic medium, the homoserine species of this peptide occurred with a mass of 2421.2 Da (+18 for added water). Both peptide species, homoserine lactone and homoserine, were present as their formylated derivatives (+28 Da for added CO) with *m*/*z* values of 2432.2 and 2448.3 Da, respectively. To a lesser extent, a twice formylated homoserine peptide with an *m*/*z* value of 2476.2 Da appeared.

The incorporation of the photoreactive probe [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s within the sequence GAGTRELELAIRITLYAVIFLM* was shown by a second series of signals with *m*/*z* values corresponding to the sum of the nonlabeled peptide species and the covalently attached probe (Figure 6C). The molecular mass of *p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s [FAB-MS 1519 Da $(M - H)^-$ (*24*)] was reduced by cyanogen bromide cleavage of the photoaffinity-labeled receptor. After covalent attachment to the $CCK_BR$, the probe too will be cleaved at its two methionine residues (Figure 1) and the tyrosine sulfate will be completely hydrolyzed during cyanogen bromide cleavage in 70% formic acid. Additionally, the amide bond between the side chain of ornithine and tritiated propionic acid may be hydrolyzed due to the harsh acidic conditions. Synthesis of the radioactive probe will yield both tritiated and nontritiated compounds since radioactivity was introduced by a commercially available activated ester with a known, but lot-dependent, specific radioactivity. Species in which tritium is replaced by hydrogen will have a mass difference of 2 Da. The resulting propionyl derivative of the radioactive probe fragment covalently attached to the $CCK_BR$ therefore should have a molecular mass of 759.8 or 757.8 Da with or without tritium, respectively. The resulting despropionyl probe fragment will have a mass reduced by 58 Da which would amount to 699 Da.

The photoaffinity-labeled sequence GAGTRELELAIRITLYAVIFLM* should therefore have an additional mass of either 757.8 and 759.8, or 699, Da compared with the nonlabeled species. In Figure 6C, the signal with an *m*/*z* value of 3175.6 corresponds to the nontritiated photoaffinity-labeled homoserine peptide GAGTRELELAIRITLYAVIFLM*. The despropionyl species of the photoaffinity-labeled peptide was detected with an *m*/*z* value of 3118.7 Da, and its formylated species appeared with an *m*/*z* value of 3147.7 Da. Since the yield of the covalent attachment of the probe and receptor amounted to 30% of the total binding sites and parts of the covalent bond between the probe and receptor may be cleaved during purification of the agonist binding domain, the labeled peptides occurred in a clearly weaker extent than the signals corresponding to the nonlabeled peptide. Since available substance amounts were too low for recording post source decay spectra, no information about the position of the covalently attached probe within the identified peptide could be obtained by MALDI-TOF mass analysis.

*Electrospray Tandem Mass Analysis.* To refine the position of the probe, the purified agonist binding domain was

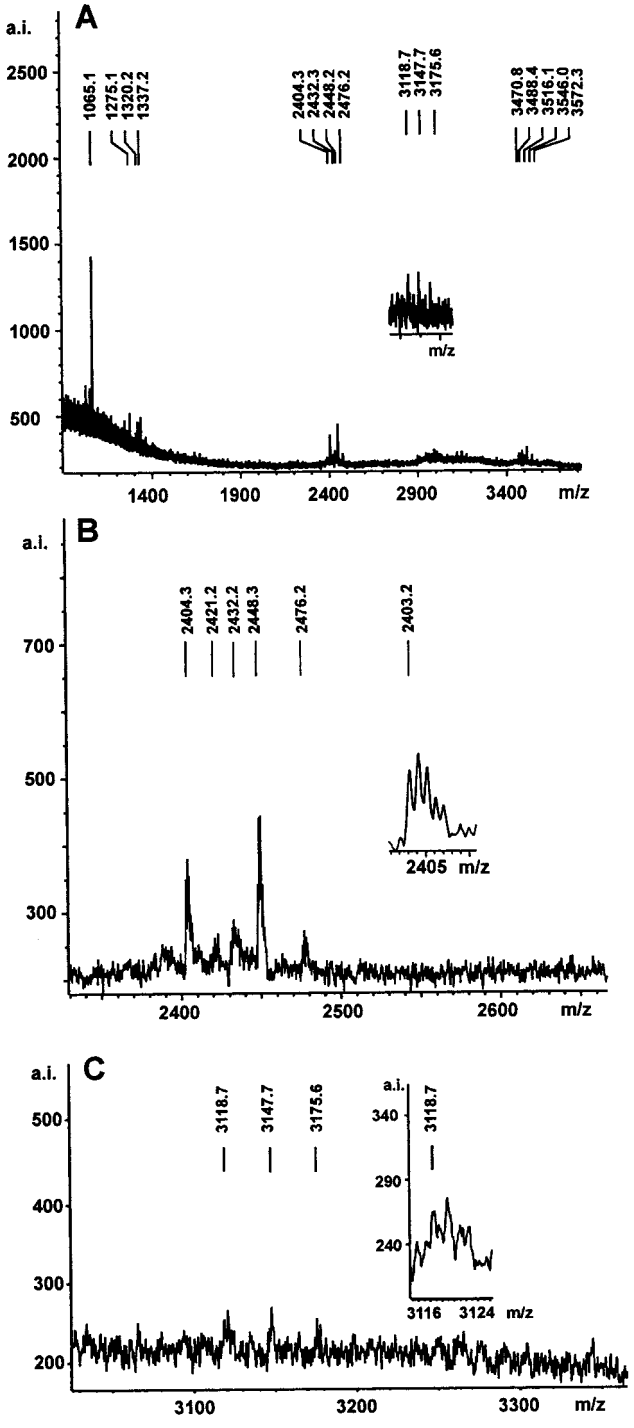


FIGURE 6: MALDI-TOF mass analysis. Shown is the MALDI-TOF mass spectrum obtained from 0.2 pmol of isolated agonist binding domain using α-cyano-4-hydroxycinnamic acid as a matrix (A). Two series of peaks could be assigned to the photoaffinity-labeled (C) and nonlabeled (B) peptide GAGTRELELAIRIT-LYAVIFLM*, with M* denoting the C-terminal homoserine resulting from cyanogen bromide cleavage of the photoaffinity-labeled CCK$_B$R, which are shown in the insets. The assignment of the peaks is described in the text.

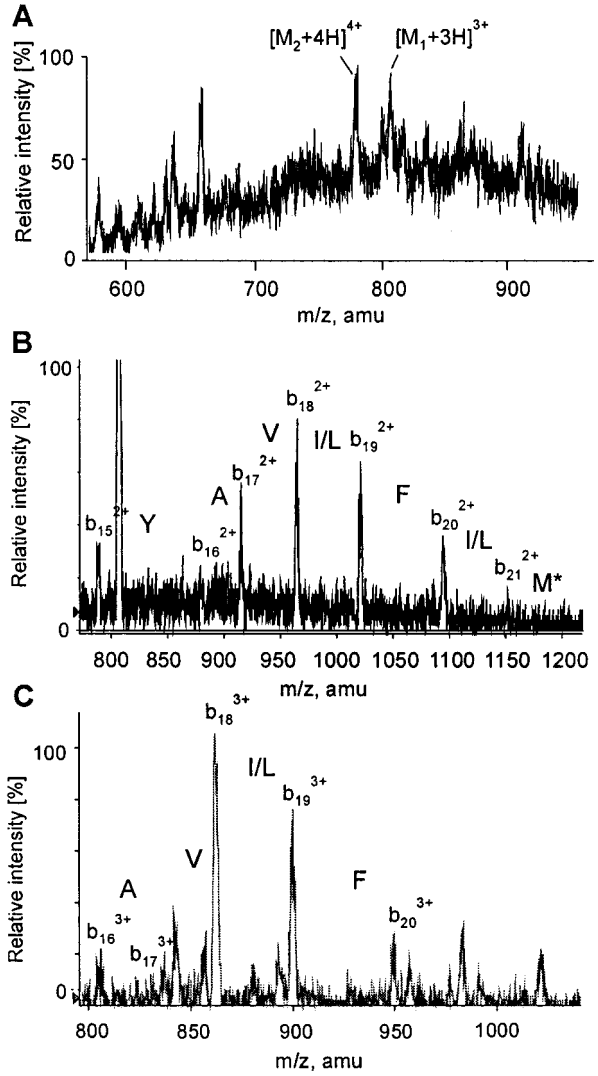


FIGURE 7: ESI tandem mass analysis. Shown is the peptide ion spectrum obtained with 0.9 pmol of the agonist binding domain (A). Two multiply charged precursor ions were selected for fragmentation which were labeled $[M_1 + 3H]^{3+}$ and $[M_2 + 4H]^{4+}$. The first $[M_1 + 3H]^{3+}$ (*m/z* of 806.7, calculated monoisotopic mass of 2419.4 Da, and detected average mass of 2416.9 Da) corresponded to the nonlabeled sequence GAGTRELELAIRITLYAV-IFLM*. A continuous series of doubly charged $b_i$ ions could be assigned to the dominant peaks in the fragment spectrum yielding the sequence YAVI/LFI/LM* (B). The second selected peak $[M_2 + 4H]^{4+}$ (*m/z* of 780.2, calculated monoisotopic mass of 3119.2 Da, and detected average mass of 3116.8 Da) could be assigned to the despropionyl photoaffinity-labeled peptide GAGTRELELAIR-ITLYAVIFLM* appearing in the MALDI-TOF mass spectra with an average *m/z* value of 3118.7 Da $[(M + H)^+]$ (Figure 8C). A continuous series of triply charged $b_i$ ions with an assumed N-terminal mass of 656 Da could be assigned to the dominant peaks in the spectrum yielding the sequence tag AVI/LF located near the C-terminus of the agonist binding site (C).

subjected to electrospray tandem mass spectrometry (*40*−*42*). Two multiply charged precursor ions were selected for fragmentation, which could be assigned to the photoaffinity-labeled ($[M_2 + 4H]^{4+}$) and the nonlabeled ($[M_1 + 3H]^{3+}$) GAGTRELELAIRITLYAVIFLM* peptide (Figure 7A). The masses of multiply charged precursor ions were determined by a precursor ion scan detecting I/L-containing peptides. Fragmentation of these peaks produced ion series containing parts of the amino acid sequences located at the C-terminus of the peptides (Figure 7B,C).

The first selected peak $[M_1 + 3H]^{3+}$ (*m/z* of 806.7, calculated monoisotopic mass of 2419.4 Da, and detected average mass of 2416.9 Da) corresponded to the nonlabeled sequence GAGTRELELAIRITLYAVIFLM* appearing in the MALDI-TOF mass spectra with an average *m/z* value of 2421.2 Da $[(M + H)^+]$ (Figure 6B). A continuous series

of doubly charged $b_i$ ions could be assigned to the dominant peaks in the spectrum yielding the sequence YAVI/LFI/LM*, which corresponds to the seven C-terminal amino acids of the detected agonist binding site of the $CCK_BR$.

The second selected peak $[M_2 + 4H]^{4+}$ (*m*/*z* of 780.2, calculated monoisotopic mass of 3119.2 Da, and detected average mass of 3116.8 Da) could be assigned to the despropionyl photoaffinity-labeled peptide GAGTRELELAIRITLYAVIFLM* appearing in the MALDI-TOF mass spectra with an average *m*/*z* value of 3118.7 Da $[(M + H)^+]$ (Figure 7C). Due to the despropionyl species of the ornithine residue in the fragment of the photoreactive CCK receptor probe (see Figure 1A), this peptide could be protonated 4-fold. A continuous series of triply charged $b_i$ ions with an assumed N-terminal mass of 656 Da could be assigned to the dominant peaks in the spectrum yielding the sequence AVI/LF, which corresponds to the sequence AVIF near the C-terminus of the detected agonist binding site of the $CCK_BR$.

Electrospray tandem mass analysis of the isolated agonist binding site clearly verifies the interpretation of the MALDI-TOF mass spectra shown in Figure 6. Additionally, the position of the incorporated probe [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s within the sequence GAGTRELELAIRITLYAVIFLM was defined. Since C-terminal amino acids of the photoaffinity-labeled peptide could be assigned due to mass differences of fragment ions, they cannot carry the additional mass of the probe fragment. Therefore, it could be shown that the probe was incorporated within the sequence GAGTRELELAIRITLY excluding the inner transmembrane part of the peptide identified by Edman degradation amino acid sequence analysis (Figure 8B).

*V8 Protease Peptide Mapping*. Further definition of the position of the probe [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s within the identified peptide was achieved by V8 protease peptide mapping of the radiolabeled cyanogen bromide fragment extending from residue $Glu^2$ to $Met^{67}$ purified from the gel filtration eluate (Figure 8A). A single fragment with a molecular mass of about 2 kDa by comparison with standard peptides existed. Only the non-glycosylated fragments corresponding to residues $Leu^{52}$ and $Glu^{53}$, and $Leu^{54}-Met^{67}$, were consistent with the low molecular mass of the radiolabeled peptide (Figure 8A).

When the results obtained by both electrospray tandem mass analysis and V8 protease peptide mapping were considered, it became clear that the photoreactive probe [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s was incorporated within the sequence LELAIRITLY ($Leu^{52}-Tyr^{61}$ of the $CCK_BR$) which represents the transition region between the N-terminal domain and the predicted first transmembrane domain (Figure 8B). Since the probe [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s carried the photoreactive *p*-benzoylbenzoyl moiety at the peptide amino terminus (Figure 1A), the identified sequence LELAIRITLY ($Leu^{52}-Tyr^{61}$ of the wild-type $CCK_BR$) corresponded to the contact site between the amino terminus of the probe and the $CCK_BR$.

### *Molecular Modeling of the $CCK_BR-CCK$-8s Complex*

On the basis of the experimental results described above, we studied the CCK-8s binding to the $CCK_B$ receptor using a 250 ps molecular dynamics simulation with respect to an orientation of the ligand N-terminus pointing toward the identified sequence $L^{52}ELAIRITLY^{61}$. The putative structure of the $CCK_BR-CCK$-8s complex with the lowest potential energy is shown in Figure 9.

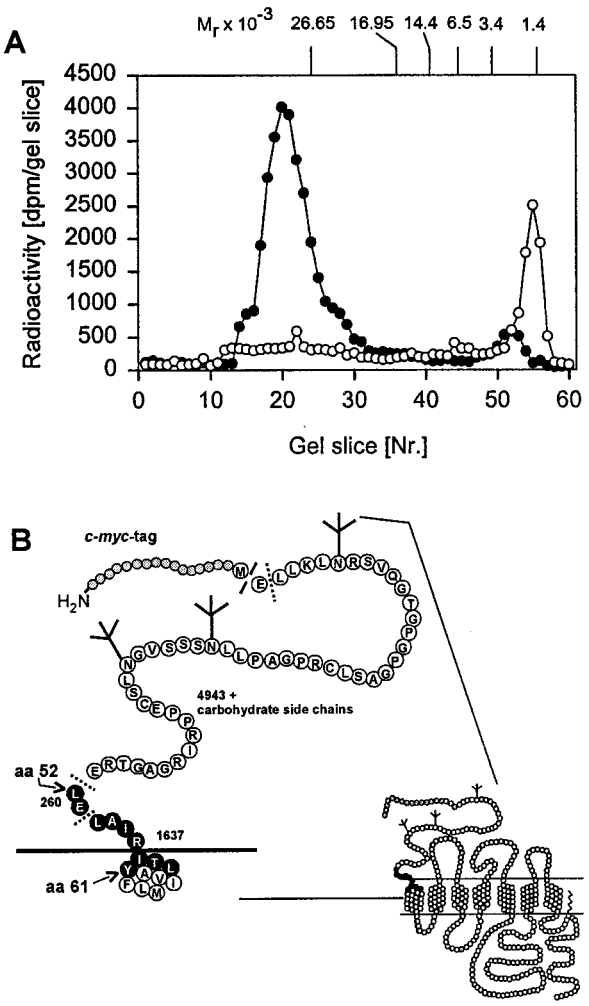


FIGURE 8: Peptide mapping of the agonist binding site in the $CCK_B$ receptor. The radioactive photoaffinity-labeled cyanogen bromide fragment $Glu^2-Met^{67}$ was purified from the gel filtration eluate and subjected to enzymatical cleavage with V8 protease. Shown is a typical distribution of radioactivity in a 16.5% T/6% C Tris/Tricine/urea SDS−PAGE analysis used to separate the products of V8 protease digestion of the $CCK_BR$ fragment of $Glu^2-Met^{67}$ with (○) or without (●) enzyme which is representative of three independent experiments (A). Also shown are the theoretical cleavage sites of the fragment, including the additional *c-myc* tag and the masses of the expected fragments wihch are independent of carbohydrate moieties and the mass of the eventually incorporated photoreactive probe (B). Only the fragments corresponding to residues $Leu^{52}$ and $Glu^{53}$, and $Leu^{54}-Met^{67}$, starting at the first residue of the $CCK_BR$ without numbering the additional *c-myc* tag are consistent with the migrations of fragments observed on the gel. Residues which are shown to be part of the agonist binding site are depicted in black circles.

The model implies that the $L^{52}ELAIRITLY^{61}$ motif builds a helix extending the first transmembrane domain of the $CCK_B$ receptor to the extracellular site. The side chains of only three residues ($Leu^{54}$, $Arg^{57}$, and $Tyr^{61}$) of the LELAIRITLY sequence are pointing toward the putative binding pocket for CCK-8s. Two of them, $Arg^{57}$ and $Tyr^{61}$, appear to interact directly with $Asp^1$ of CCK-8s, forming hydrogen bonds. This is in good agreement with a result obtained by alanine scanning mutagenesis, which reported $Arg^{57}$ to be

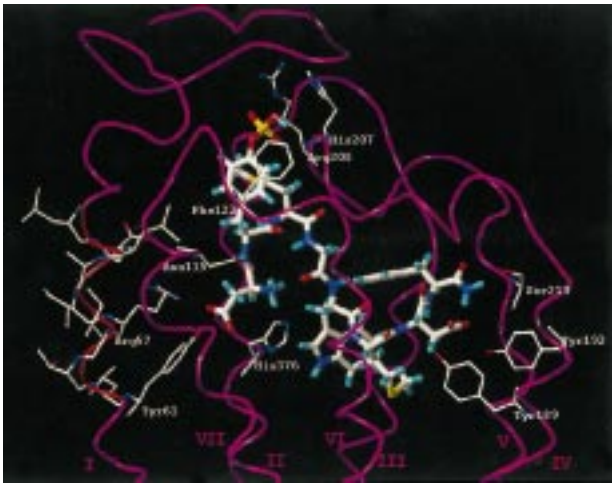

FIGURE 9: Molecular modeling of the $CCK_BR$−CCK-8s complex. Representation of the minimized lowest-potential energy conformation of the CCK-8s−$CCK_BR$ complex (248 ps) taken from the MD trajectory. The backbone of the experimentally identified binding region ($L^{52}$ELAIRITLY$^{61}$) is shown as a red-colored tube.

important for high-affinity CCK-8s binding to the $CCK_BR$ (*19*).

Four other residues of the $CCK_BR$ identified in this study to be critical for high-affinity CCK-8s binding ($Asn^{115}$, $Leu^{116}$, $Phe^{120}$, and $Phe^{122}$) are clustered in the first extracellular loop. When the peptide ligand is docked in a $\beta$-hairpin conformation into the putative binding site of $CCK_BR$, the sulfated tyrosine of the ligand is well positioned between the first and second extracellular loops for forming a good interaction with $Phe^{122}$, making an aromatic stack. Such aromatic−aromatic interactions have proven to be very important for the stabilization of proteins or in ligand−protein interactions (*43*). Furthermore, we found that $Asn^{115}$ forms a hydrogen bond to the backbone NH group of sulfated tyrosine and $Phe^{120}$ is located near this residue, shielding the binding pocket from the extracellular space. In the second extracellular loop of the $CCK_BR$, $His^{207}$ and $Arg^{208}$ were found in the model to interact with the sulfonyl group of CCK-8s. This finding agrees very well with the remarkable effect of the $His^{207}$Phe mutation on CCK-8s binding (*19*).

On the basis of information from site-directed mutagenesis studies (*20*, *25*), we have located the C-terminal part of CCK-8s near $Ser^{219}$ (TM VI) using a distance restraint. In our molecular dynamics simulation, this interaction is responsible for placing the C-terminal hydrophobic residues of CCK-8s deep in a hydrophobic core of the receptor.

In summary, the model shows that the N-terminus of CCK-8s mainly interacts with $Arg^{57}$ and $Tyr^{61}$ which are part of the receptor region identified in this work using photoaffinity labeling of the $CCK_BR$ with a CCK probe with its photoreactive moiety at the peptide N-terminus.

## DISCUSSION

Understanding the molecular interactions involved in agonist binding provides insights that could be helpful for targeted drug design and for understanding the basic mechanism of G-protein-coupled receptor activation. The cholecystokinin$_B$ receptor is a particularly important drug target due to its widespread expression throughout the central nervous system and in certain peripheral tissues, including stomach, pancreas, and smooth muscle, and since it mediates a broad variety of physiological actions.

In this paper, we describe the purification and photoaffinity labeling of the human brain cholecystokinin$_B$ receptor from stably transfected HEK 293 cells and the identification of its agonist binding site. The photoreactive probe [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s bound to the recombinant $CCK_BR$ with an affinity of $1.2 \pm 0.1$ nM and triggered calcium responses of stably transfected HEK 293 cells with the same potency as the parent peptide CCK-8s (Figure 1), thus representing a high-affinity agonist of the $CCK_BR$. By mass spectrometry of peptide fragments of the purified photoaffinity-labeled receptor, we were able to prove directly that the receptor sequence $L^{52}$ELAIRITLY$^{61}$ representing the transition region between the N-terminal domain and transmembrane domain 1 of the $CCK_BR$ is the contact site between the receptor and the N-terminus of the photoreactive CCK receptor probe. These data represent the first direct structural information about the $CCK_BR$−CCK complex reported so far and provide a structural basis for understanding the results of previously published mutagenesis studies.

A recent study in which alanine scanning mutagenesis of residues that are nonconserved between the two receptor subtypes was used to identify residues involved in CCK-8s binding reported five residues being critical for high-affinity CCK-8s binding (*19*). One of these residues, $Arg^{57}$, was shown by our protein chemical approach to be part of the $CCK_B$ receptor sequence covalently linked to the CCK receptor probe [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s and therefore is part of the contact site between the peptides N-termini and the $CCK_BR$. Additionally, it was found in our modeling study that $Arg^{57}$ forms a hydrogen bond to $Asp^1$ of CCK-8s, implying that this interaction is responsible for the orientation of CCK in the putative agonist binding pocket. The importance of the N-terminal region of the $CCK_B$ receptor for agonist binding is further supported by the discovery of a splice variant from which the N-terminal extracellular domain and the upper part of the first transmembrane domain had been removed and which had an altered pharmacology (*44*). The four remaining critical residues identified by alanine scanning mutagenesis ($Asn^{115}$, $Leu^{116}$, $Phe^{120}$, and $Phe^{122}$) are clustered in the first extracellular loop of the receptor. In addition, a sequence of five residues ($Q^{204}$CVHRW$^{209}$) in the second extracellular loop of the $CCK_B$ receptor was shown to be essential for its high affinity for the natural peptide agonist gastrin (*19*), with $His^{207}$ additionally being crucial for high-affinity CCK-8s binding (*18*). In our modeling study, residues $Asn^{115}$, $Phe^{120}$, and $Phe^{122}$ of the first extracellular loop and $His^{207}$ together with $Asp^{208}$ of the second extracellular loop appear to interact with sulfated tyrosine of CCK-8s, with the model suggesting a putative binding pocket for this residue. These findings indicate that determinants of the ligand selectivity other than the decapeptide identified in this study are located within the extracellular domains of the $CCK_BR$. In the case of the $CCK_A$ receptor, $Met^{195}$ in the second extracellular loop of the $CCK_AR$ was reported to be crucial for high-affinity CCK-8s binding and was found by molecular modeling to interact with the sulfated tyrosine of the $CCK_AR$ (*45*). A recently published photoaffinity labeling study in which a photolabile CCK probe with its photoreactive moiety located in the

middle of the molecule at position 4 was used reported that the probe is incorporated into residues $His^{347}$ and $Leu^{348}$ in the third extracellular loop of the $CCK_AR$ (*46*). Both findings suggest a similar important role of extracellular loops in the $CCK_AR$ for high-affinity agonist binding as was reported in case of the $CCK_BR$. In contrast, mutations of transmembrane domain residues interfere only slightly with agonist binding (*20*, *23*).

The comparison of our protein chemical result with data previously presented for the $CCK_A$ receptor is of particular interest, since both receptor subtypes share approximately 50% amino acid identity and are involved in the same signal transduction mechanism but differ clearly in their agonist binding profile. In the case of the $CCK_AR$, the agonist binding site has been characterized by both mutagenesis and photoaffinity labeling studies together with molecular modeling providing detailed, but in some respects controversial, insights. With a mutagenesis and molecular modeling approach, Kennedy et al. (*47*) reported that residues $Trp^{39}$ and $Gln^{40}$ located adjacent to the top of transmembrane domain 1 interact with the N-terminal amino group of the CCK nonapeptide [Thr,Nle]CCK-9. In the alignment of both human $CCK_A$ and $CCK_B$ receptors, these residues correspond to the positions $Leu^{52}$ and $Glu^{53}$ of the wild-type $CCK_B$ receptor determined by our protein chemical approach to be part of the identified decapeptide $L^{52}ELAIRITLY^{61}$ covalently linked to the [$^3$H]-*p*-benzoylbenzoyl-Orn(propionyl)-CCK-8s amino terminus. A more recent study in which photoaffinity labeling of the $CCK_AR$ with a radiolabeled probe that was chemically modified with the photoreactive group at the peptide C-terminus was used reported that this probe was incorporated into residue $Trp^{39}$, indicating that this residue is the critical point of contact with the CCK-8s C-terminus (*48*). The authors investigated the position of the probe within the receptor by radioactive sequencing of a radiolabeled peptide on a glass support; however, this report lacks a direct Edman amino acid sequence analysis or mass spectrometry. The identity of the radiolabeled peptide was only based on the estimation of a peptide molecular mass by SDS−PAGE analysis.

Taken together, the available data indicate that the CCK-binding sites of both the $CCK_AR$ and $CCK_BR$ might be quite similar. In both the $CCK_AR$ and $CCK_BR$, CCK agonists were shown to be in contact with receptor residues located in the N-terminal part adjacent to or just on top of transmembrane domain 1 with respect to the same orientation of CCK ligands within the putative binding pockets of both receptors. Within these pockets, different nonconserved residues seem to be directly involved in agonist binding. In our modeling study, the C-terminal amino acids of CCK-8s were proposed to be buried in a hydrophobic core of $CCK_BR$ with respect to results obtained by site-directed mutagenesis (*20*, *23*). A similar finding was presented in the case of the complex between the CCK nonapeptide [Thr,Nle]CCK-9 and the $CCK_AR$ (*47*). Since residues of transmembrane domains are highly conserved among $CCK_B$ and $CCK_A$ receptors of different species and since the hydrophobic C-termini of both gastric and CCK peptides are identical, this suggests a very similar mode of agonist C-terminus binding to both CCK receptor subtypes.

However, despite their similar agonist binding modes, the $CCK_AR$ and $CCK_BR$ differ clearly in their agonist binding selectivity, raising the question of the structural basis of this difference. With respect to the available data, one may suggest that the different agonist binding selectivities of both receptor subtypes may be due to specific contact residues within a quite similar structure motif. Further refinement of the structural information for the interaction between CCK and the CCK receptors will be required to elucidate whether the corresponding binding sites for CCK and gastrin are similar in the $CCK_AR$ and $CCK_BR$ and which structural motifs of both receptors are responsible for the agonist selectivity. The results of our protein chemical analysis will provide the basis for further investigations of the molecular basis of two receptor subtypes that induce the same signal transduction mechanisms upon agonist binding but clearly differ in their agonist selectivities.

## ACKNOWLEDGMENT

We thank M. Wilm (EMBL, Heidelberg, Germany) for the recording of the ESI tandem mass spectra and A. Kopin (New England Medical Center Hospitals, Boston, MA) for generously providing the human brain $CCK_B$ receptor cDNA. We further want to thank F. Bender (Johannes Gutenberg Universität Mainz) for technical advice and help concerning preparation of the manuscript and figures.

## REFERENCES

1. Wank, S. A. (1995) *Am. J. Physiol. 269*, G628−G646.
2. Kinze, S., Schöneberg, T., Meyer, R., Martin, H., and Kaufmann, R. (1996) *Neurosci. Lett. 217*, 45−49.
3. Faris, P. L., Komisaruk, B. R., Watkins, L. R., and Mayer, D. J. (1983) *Science 219*, 310−312.
4. Dourish, C. T., O'Neill, M. F., Coughlan, J., Kitchener, S. J., Hawley, D., and Iversen, S. D (1990) *Eur. J. Pharmacol. 176*, 35−44.
5. Wiertelak, E. P., Maier, S. F., and Watkins, L. R. (1992) *Science 256*, 830−833.
6. Harro, J., Vasar, E., and Bradwejn, J. (1993) *Trends Pharmacol. Sci. 14*, 244−249.
7. Bradwejn, J., Koszycki, D., Couetoux de Tertre, A., van Megen, H., den Boer, J., and Westenberg, H. (1994) *Arch. Gen. Psychiatry 51*, 486−493.
8. Kopin, A. S., Lee, Y.-M., McBride, E. W., Miller, L. J., Lu, M., Lin, H. Y., Kolakowski, L. F., and Beinborn, M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3605−3609.
9. Wank, S. A., Pisegna, J. R., and De Weerth, A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8691−8695.
10. Lee, Y. M., Beinborn, M., McBridge, E. W., Lu, M., Kolakowski, L. F., and Kopin, A. S. (1993) *J. Biol. Chem. 268*, 8164−8169.
11. Silvente-Poirot, S., Dufresne, M., Vaysse, N., and Fourmy, D. (1993) *Eur. J. Biochem. 215*, 513−529.
12. Strader, C. D., Sigal, I. S., and Dixon, R. A. (1989) *Trends Pharmacol. Sci. 10* (Dec. Suppl.), 26−30.
13. Caron, M. G., and Lefkowitz, R. J. (1993) *Recent Prog. Horm. Res. 48*, 277−290.
14. Kojro, E., Eich, P., Gimpl, G., and Fahrenholz, F. (1993) *Biochemistry 32*, 13537−13544.
15. Schwartz, T. W. (1994) *Curr. Opin. Biotechnol. 5*, 434−444.
16. Strader, C. D., Fong, T. M., Tota, M. R., Underwood, D., and Dixon, R. A. (1994) *Annu. Rev. Biochem. 63*, 101−132.
17. Postina, R., Kojro, E., and Fahrenholz, F. (1996) *J. Biol. Chem. 271*, 31593−31601.
18. Silvente-Poirot, S., and Wank, S. A. (1996) *J. Biol. Chem. 271*, 14698−14706.
19. Silvente-Poirot, S., Escrieut, C., and Wank, S. A. (1998) *Mol. Pharmacol. 54*, 364−371.

20. Kopin, A. S., McBride, E. W., Quinn, S. M., Kolakowski, L. F., Jr., and Beinborn, M. (1995) *J. Biol. Chem. 270*, 5019−5023.
21. Jagerschmidt, A., Guillaume, N., Goudreau, N., Maiget, B., and Roques, B.-P. (1995) *Mol. Pharmacol. 48*, 783−789.
22. Jagerschmidt, A., Guillaume-Rousselet, N., Vikland, M.-L., Goudreau, N., Maigret, B., and Roques, B.-P. (1996) *Eur. J. Pharmacol. 296*, 97−106.
23. Jagerschmidt, A., Guillaume, N., Roques, B. P., and Noble, F. (1998) *Mol. Pharmacol. 53*, 878−885.
24. Thiele, C., and Fahrenholz, F. (1993) *Biochemistry 32*, 2741−2746.
25. Gimpl, G., Burger, K., and Fahrenholz, F. (1997) *Biochemistry 36*, 10959−10974.
26. Munson, P. J., and Rodbard, D. (1980) *Anal. Biochem. 107*, 220−239.
27. Grynkiewicz, G., Poenie, M., and Tsien, R. Y. (1984) *J. Biol. Chem. 260*, 3440−3450.
28. Laemmli, U. K. (1970) *Nature 227*, 680−685.
29. Wessel, P. J., and Flügge, U. I. (1984) *Anal. Biochem. 138*, 141−143.
30. Schägger, H., and von Jagow, G. (1987) *Anal. Biochem. 166*, 368−379.
31. Gimpl, G., Anders, J., Thiele, C., and Fahrenholz, F. (1996) *Eur. J. Biochem. 237*, 768−777.
32. Baldwin, J. M., Schertler, G. F. X., and Unger, V. M. (1997) *J. Mol. Biol. 272*, 144−154.
33. Unger, V. M., Hargrave, P. A., Baldwin, J. M., and Schertler, G. F. X. (1997) *Nature 389*, 203−206.
34. Lutz, J., Romano-Götsch, R., Escrieut, D., Fourmy, D., Mathä, B., Müller, G., Kessler, H., and Moroder, L. (1997) *J. Pept. Sci. 3*, 1−14.
35. Moroder, L. (1997) *Biopolymers 41*, 799−817.
36. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., III, Ferguson, D. M., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1995) *AMBER 4.1*, University of California, San Francisco.
37. Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) *J. Phys. Chem. 97*, 10269.
38. Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunage, N., Nguyen, K. A., Su, S. J., Windus, T. L., Dupuis, M., and Montgomery, J. A. (1993) *J. Comput. Chem. 14*, 1347−1363.
39. Schaftenaar, G. (1992) *QCPE Bull. 12*, 3 (QCPE619).
40. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) *Nature 379*, 466−469.
41. Mann, M., and Wilm, M. (1995) *Trends Biochem. Sci. 20*, 219−224.
42. Mørtz, E., O'Connor, P. B., Roepstorff, P., Kelleher, N. L., Wood, T. D., McLafferty, F. W., and Mann, M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 8264−8267.
43. Hunter, C. A., Singh, J., and Thornton, J. M. (1991) *J. Mol. Biol. 218*, 837−846.
44. Miyake, A. (1995) *Biochem. Biophys. Res. Commun. 208*, 230−237.
45. Gigoux, V., Escrieut, C., Silvente-Poirot, S., Maigret, B., Gouilleux, L., Fehrentz, J.-A., Gully, D., Moroder, L., Vaysse, N., and Fourmy, D. (1998) *J. Biol. Chem. 273*, 14380−14386.
46. Hadac, E. M., Pinon, D. I., Ji, Z., Holicky, E. L., Henne, R. M., Lybrand, T. P., and Miller, L. J. (1998) *J. Biol. Chem. 273*, 12988−12993.
47. Kennedy, K., Gigoux, V., Escrieut, C., Maigret, B., Martinez, J., Moroder, L., Frehel, D., Gully, D., Vaysse, N., and Fourmy, D. (1997) *J. Biol. Chem. 272*, 2920−2926.
48. Ji, Z., Hadac, E. M., Henne, R. M., Patel, S. A., Lybrand, T. P., and Miller, L. J. (1997) *J. Biol. Chem. 272*, 24393−24401.

BI990269Z