Molecular Pharmacology Fast Forward. Published on May 2, 2007 as doi:10.1124/mol.107.035402

Molecular Approximations between Residues 21 and 23 of Secretin and its Receptor. Development of a Model for Peptide Docking with the Amino Terminus of the Secretin Receptor

Maoqing Dong, Polo C.-H. Lam, Fan Gao, Keiko Hosohata, Delia I. Pinon, Patrick M. Sexton, Ruben Abagyan, Laurence J. Miller

Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic Scottsdale, Scottsdale, AZ 85259 (M.D., F.G., K.H., D.I.P., L.J.M.), Department of Molecular Biology, Scripps Research Institute and Molsoft LLC, La Jolla, CA 92037 (P.C.-H.L., R.A.), and Department of Pharmacology, Monash University, Clayton, Victoria 3800, Australia (P.M.S.).

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Running title: Secretin receptor ligand binding

Address correspondence to: Laurence J. Miller, M.D., Mayo Clinic, 13400 East Shea Blvd., Scottsdale, AZ 85259, Telephone: (480) 301-6650; Fax: (480) 301-6969, E-mail: miller@mayo.edu

Manuscript information:

Number of text pages: 40

Number of figures: 10

Number of tables: 2

Number of References: 40

Number of words in Abstract: 223

Number of words in Introduction: 890

Number of words in Discussion: 1756

Abbreviations used:

Bpa, *p*-benzoyl-*L*-phenylalanine; CHO, Chinese hamster ovary; CNBr, cyanogen bromide; CRF, corticotrophin-releasing factor; GPCR, G protein-coupled receptor; HA, hemagglutinin; ICM, internal coordinate mechanics; KRH, Krebs-Ringers/HEPES; rmsd, root mean square deviation; SecR, secretin receptor; VPAC1, vasoactive intestinal polypeptide receptor.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract

The structurally-unique amino-terminal domain of Class II G protein-coupled receptors is critically important for ligand binding and receptor activation. Understanding the precise role it plays requires detailed insights into the molecular basis of its ligand interactions and the conformation of the ligand-receptor complex. In this work, we have used two high-affinity, full-agonist, secretin-like photolabile probes having sites for covalent attachment in positions 21 and 23, and have utilized sequential proteolysis and sequencing of the labeled region of the receptor to identify two new spatial approximation constraints. The position 21 probe labeled receptor residue Arg^{15}, while the position 23 probe labeled receptor residue Arg^{21}. A homology model of the amino-terminal domain of the secretin receptor was developed using the NMR structure of the analogous domain of the corticotrophin releasing factor receptor. This was attached to a homology model of the secretin receptor transmembrane bundle, with the two domains oriented relative to each other based on continuity of the peptide backbone and by imposing a distance restraint recently identified between the amino-terminal WDN sequence and the region of the helical bundle above transmembrane segment six. Secretin was docked to this model using seven sets of spatial approximation constraints identified in previous photoaffinity labeling studies. This model was found to fully accommodate all existing constraints, as well as the two new approximations identified in this work.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

The secretin receptor is a prototypic member of the Class II family of G protein-coupled receptors (GPCRs) that contains several potentially important drug targets, such as receptors for calcitonin, vasoactive intestinal polypeptide, glucagon, glucagon-like peptide, corticotrophin-releasing factor (CRF) and parathyroid hormone (Ulrich et al., 1998). Despite recent advances in the study of this receptor family, our understanding of receptor structure-function remains limited. Refinement of our understanding of the molecular basis of agonist ligand binding and activation of the secretin receptor should contribute insights relevant to the entire family and facilitate the design and refinement of potential receptor-active drugs.

Insights into mechanisms of agonist ligand binding and activation of the secretin receptor have come from natural ligand structure-activity relationships and from receptor mutagenesis and photoaffinity labeling studies (Dong et al., 2003; Gourlet et al., 1996). These have identified the critical importance of the long and structurally complex amino terminus of this receptor for agonist ligand binding. This theme has been consistent for multiple other Class II GPCR family members (Cao et al., 1995; Gourlet et al., 1996; Juppner et al., 1994). Alignment of the receptor amino terminal sequences reveals conserved features that include six cysteine residues with three intradomain disulfide bonds (Grauschopf et al., 2000; Lisenbee et al., 2005). Insights into the structure of this region was substantially advanced with the solution of an NMR structure of the isolated amino terminus of the corticotropin-releasing factor (CRF2ß) receptor (Grace et al., 2004). However, it is not yet clear whether the amino terminus of the intact CRF receptor will have the same or related conformation, how similar the amino terminal region of other receptors in this family will be to that structure, or how ligands dock to these receptors.

Natural ligands for Class II GPCRs are moderately long peptides (greater than 25 residues) having diffuse pharmacophoric domains (Ulrich et al., 1998). Like natural ligands for other members of this family, the amino-terminal region of secretin contains key determinants for

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

receptor selectivity and activation, while its carboxyl-terminal region contains determinants for high affinity binding. Considerable experimental evidence indicates that the carboxyl-terminal regions of ligands for Class II receptors interact with the amino-terminal regions of their receptors (Gourlet et al., 1996; Juppner et al., 1994; Pham et al., 2004; Tan et al., 2006), while the amino terminus of the peptides is critical for agonist activity and interacts with parts of the receptor core and loop regions (Bisello et al., 1998; Dong et al., 2004). Critical to understanding receptor activation is knowledge of how these ligands dock to their receptors. The most useful insights have come from intrinsic photoaffinity labeling studies that provide physical constraints on the relative position of peptide and receptor amino acids. The Class II GPCR that has been most extensively studied this way is the secretin receptor, where spatial approximations have been established between residues scattered throughout the pharmacophore of secretin, in positions 6, 12, 13, 14, 18, 22, and 26, and residues within the receptor amino terminus (Dong et al., 2000; Dong et al., 2003; Dong et al., 1999a; Dong et al., 1999b; Dong et al., 2002; Zang et al., 2003).

To date, the only photoaffinity labeling probes shown to covalently interact with receptor residues in regions other than the amino terminus are those having photolabile residues substituted into the amino-terminal end of the peptide ligand (Dong et al., 2004). This type of observation, also reported for the parathyroid hormone receptor, has resulted in the hypothesis that natural peptide ligands of Class II GPCRs can act as tethers interacting with both amino terminus and body of their receptors, exerting tension to induce the conformational change associated with activation and G protein association (Bisello et al., 1998). We have proposed an alternative hypothesis (Dong et al., 2006) that describes a ligand-induced conformational change in the receptor amino terminus that exposes a previously “constrained” epitope (WDN in the secretin receptor) that can act as an endogenous agonist that interacts with and activates the body of the receptor.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Evaluation of these two apparently divergent activation mechanisms requires a more detailed understanding of the mechanism of binding and of the conformation of the complex between secretin and the amino terminus of the secretin receptor. We have, therefore, performed additional photoaffinity labeling experiments with new probes that incorporate photolabile residues into the important carboxyl-terminal region, in positions 21 or 23 of secretin. Both probes bound specifically and saturably to the secretin receptor and were full agonists. They were also able to covalently label single and distinct residues within the amino terminus of the receptor, at residues Arg^{15} (for position 21 probe) and Arg^{21} (for position 23 probe). We also have now utilized the NMR data for the amino-terminal domain of the structurally-related CRF2ß receptor (Grace et al., 2004) to refine that structure and to build a homology model of the analogous domain of the secretin receptor. We have oriented this relative to a homology model of the bovine rhodopsin helical bundle (Palczewski et al., 2000) to yield a preliminary full secretin receptor model. The previously-identified spatial approximation constraints from photoaffinity labeling studies were utilized to provide initial ligand docking. This model was found to also nicely accommodate the two new spatial approximation constraints identified in the current work, with all of the currently available constraints contributing toward the most meaningful model of ligand docking yet proposed.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Materials and Methods

Materials. The solid-phase oxidant, *N*-chlorobenzenesulfonamide (Iodobeads), cyanogen bromide (CNBr), and *m*-maleimidobenzoyl-*N*-hydroxysulfosuccinimide ester were purchased from Pierce Chemical Co. Phenylmethylsulfonyl fluoride, 3-isobutyl-1-methylxanthine, and *N*-(2-aminoethyl-1)-3-aminopropyl glass beads were from Sigma. Endoproteinase Lys-C and the 12CA5 monoclonal antibody that recognizes the hemagglutinin (HA) epitope were from Roche Applied Science. Secretin and endoglycosidase F were prepared in our laboratory, as described previously (Pearson et al., 1987). All other reagents were of analytical grade.

Synthetic peptides. Photolabile secretin probes used in this study were [Tyr^{10},Bpa^{21}]rat secretin (Bpa^{21} probe) and [Tyr^{10},Bpa^{23}]rat secretin (Bpa^{23} probe). They incorporated a photolabile *p*-benzoyl-*L*-phenylalanine (Bpa) to replace Arg^{21} and Leu^{23} located at the carboxyl terminal half of the ligand, respectively. Like the well characterized radioligand, [Tyr^{10}]rat secretin, they incorporated a Tyr to replace Leu^{10} that has been shown to be well tolerated. Peptides were synthesized as described previously (Powers et al., 1988), radioiodinated oxidatively using the solid-phase oxidant, Iodobeads, and purified to homogeneity using reversed-phase HPLC to yield specific radioactivities of 2,000 Ci/mmol.

Receptor-expressing cell lines. Several cell lines that had been established and well characterized previously were utilized as sources of receptor for this study. These included Chinese hamster ovary (CHO) cell lines stably expressing the wild type secretin receptor (SecR) (Ulrich et al., 1993), the HA-tagged wild type secretin receptor (SecR-HA37) (Dong et al., 1999b), and HA-tagged mutant secretin receptors (SecR-V16M-HA37 and SecR-V13M-HA37) (Dong et al., 2000). Numbering of the residues in the secretin receptor reflects the cleavage of the 22-residue signal peptide sequence (Dong et al., 1999a). These cell lines were cultured at 37°C in

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

a 5% CO_2 environment on Falcon tissue culture plasticware in Ham's F-12 medium supplemented with 5% Fetal Clone-2 (HyClone Laboratories, Logan, UT). Cells were passaged twice a week and lifted mechanically before use. Plasma membranes were prepared using discontinuous sucrose gradient centrifugation (Hadac et al., 1996). Membranes were then resuspended in Krebs-Ringers/HEPES (KRH) buffer containing 25 mM HEPES, pH 7.4, 104 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM KH_2PO_4, 1.2 mM $MgSO_4$, 0.01% soybean trypsin inhibitor, and 1 mM phenylmethylsulfonyl fluoride for storage at -80°C until they were to be used.

Binding studies. Binding activities of the Bpa^{21} and Bpa^{23} probes were performed with enriched plasma membranes from the CHO-SecR cells in standard competition-binding assays, using conditions previously established (Hadac et al., 1996). Membranes (containing ~10 µg protein) were incubated with a constant amount of radioligand, ^{125}I-[Tyr^{10}]rat secretin (5-10 pM), and increasing concentrations of non-radioactive ligand (0-1 µM Bpa^{21} or Bpa^{23} probe) in KRH buffer containing 0.2% bovine serum albumin for 1 h at room temperature (final volume of 500 µl). Separation of bound from free radioligand was performed with a Skatron cell harvester (Molecular Devices, Sunnyvale, CA), using receptor-binding filtermats that had been pre-soaked in 0.3% polybrene. Bound radioactivity was quantified using a γ-spectrometer. Non-specific binding was determined in the presence of 1 µM unlabeled secretin and represented less than 15% of total binding. Data were graphed using Prism software (GraphPad software, San Diego, CA) and were analyzed using the nonlinear least-squares curve-fitting routine in LIGAND (Munson and Rodbard, 1980).

Biological activity assay. Biological activities of the Bpa^{21} and Bpa^{23} probes were studied using assays for stimulation of intracellular cAMP in the receptor-bearing CHO-SecR cell line, as previously described (Ganguli et al., 1998). CHO-SecR cells were incubated with

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

increasing concentrations of the probe (0-1 μM) for 30 min at 37°C in KRH buffer containing 0.2% bovine serum albumin, 0.01% soybean trypsin inhibitor, 0.1% bacitracin, and 1 mM 3-isobutyl-1-methylxanthine. Reactions were stopped by adding 6% (w/v) perchloric acid and the pH was adjusted to 6.0 with 30% $KHCO_3$. The cAMP concentrations in cell lysates were measured using a competition-binding assay (Diagnostic Products Corporation; Los Angeles, CA) (Ganguli et al., 1998). Radioactivity was quantified by scintillation counting in a Beckman LS6000.

Photoaffinity labeling. The detailed procedure has been described (Dong et al., 1999b). Plasma membranes (~50 μg) were incubated with the ^{125}I-labeled Bpa^{21} or Bpa^{23} probes (~0.1 nM) in KRH buffer for 1 h in the dark at room temperature in the presence of increasing concentrations of unlabeled secretin. The reaction was then exposed to photolysis for 30 min at 4°C using a Rayonet Photochemical Reactor (Southern New England Ultraviolet Co., Bradford, CT) equipped with 3500 Å lamps. Membranes were then washed with KRH buffer, solubilized in SDS sample buffer and separated by gel electrophoresis on 10% SDS-PAGE gels (Laemmli, 1970). The apparent molecular masses of radiolabeled receptors, visualized by autoradiography, were determined by interpolation on a plot of the mobility of ProSieve protein standards (Cambrex BioScience, Rockland) versus the log values of their apparent masses. .

Peptide mapping. Photoaffinity labeled secretin receptor was prepared in larger scale using ~200 μg aliquots of plasma membranes and 0.5 nM ^{125}I-labeled Bpa^{21} or Bpa^{23} probes. Labeled bands were cut from gels, eluted, lyophilized and precipitated with ethanol prior to subsequent cleavage. Purified labeled receptor was deglycosylated with endoglycosidase F and digested with CNBr and endoproteinase Lys-C using conditions previously described (Dong et al., 1999b). Cleavage products were separated on 10% NuPAGE gels (Invitrogen, Carlsbad, CA) using MES running buffer. The apparent molecular masses of radiolabeled receptor fragments

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

were determined by interpolation on a plot of the mobility of Multimark protein standards (Invitrogen) versus the log values of their apparent masses.

The identity of the affinity-labeled secretin receptor fragment resulting from CNBr cleavage was further established by immunoprecipitation with the 12CA5 anti-HA monoclonal antibody using the affinity-labeled HA-tagged receptor construct (SecR-HA37), previously shown to bind and signal identically to wild type receptor (Dong et al., 1999b). The immunoprecipitated material was eluted into sample buffer and resolved by NuPAGE gel electrophoresis.

Radiochemical Edman degradation sequencing (Dong et al., 1999b) was used to determine the specific sites of attachment after achieving definitive identification of the receptor fragments labeled with each probe. This involved the cross-linking of the receptor fragment of interest through cysteine residues to *N*-(2-aminoethyl-1)-3-aminopropyl glass beads. We used V13M-HA37 and V16M-HA37 mutant receptor constructs expressed in CHO cell lines for the Bpa^{21} and Bpa^{23} probes, respectively. Labeled receptor was purified and cleaved by CNBr and the resultant fragments were gel-purified to radioactive homogeneity before being covalently coupled through cysteine residues to maleimidobenzoyl succinimide-activated *N*-(2-aminoethyl-1)-3-aminopropyl glass beads. Following this, repetitive cycles of manual Edman degradation were performed with quantitation of radioactivity released in each cycle. This procedure was repeated three times in independent experiments.

Molecular modeling. All molecular modeling activities were conducted using a stochastic global energy optimization procedure implemented in Internal Coordinate Mechanics (ICM) (Abagyan, 1994). This procedure consisted of the following iterative steps: (a) random conformational change of a dihedral angle according to the biased-probability Monte Carlo

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

(BPMC) method (Abagyan and Totrov, 1994); (b) local minimization of all free dihedral angles; and (c) acceptance or rejection of the new conformation based on the Metropolis criterion at the simulation temperature, usually at 600 K (Metropolis et al., 1953). This procedure can generate and search through diverse sets of conformations by actively sampling a selected set of dihedral angles. All calculations were carried out on 3.4 GHZ Intel XEON-EMT processors.

Refinement of the NMR models of the mouse CRF2ß receptor amino-terminal domain. The 20 NMR models of the mouse CRF2ß receptor amino-terminal domain and the associated NMR distance restraints that were deposited into Protein Data Bank (PDB: 1U34) were obtained. The highly flexible regions at both ends of this domain were removed, retaining only residues 37-122. The models were then completed by attaching missing hydrogens, assigning ECEPP/3 atom types and charges (Nemethy et al., 1992), followed by regularization to impose ideal covalent geometries on all bond angles and bond lengths. 803 NMR restraints were converted to ICM-readable format and were imposed on the models using a quadratic restraint energy with a strength of 10 kcal/(mol$Å^2$) (Bordner and Abagyan, 2006). Each of the 20 models was then subjected to local minimization, followed by ten independent cycles of Monte Carlo sampling of both the backbone and side chain dihedral angles at 600 K in ICM, each simulation lasting for more than 144 h (Abagyan and Totrov, 1999). The lowest energy conformation for each of the 20 models was retained.

Homology modeling of the amino-terminal domain of the rat secretin receptor. The multiple sequence alignment of Class II GPCRs from the Swissprot database was used as a guide to align the sequence of the rat secretin receptor to the mouse CRF2ß receptor amino-terminal domain (Boeckmann et al., 2003). The 20 refined models of this domain of the CRF2ß receptor were then used to generate 20 homology models of the amino terminus of the secretin receptor (Cardozo et al., 1995). A subset of 115 NMR restraints of the CRF2ß receptor amino terminus

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

was imposed on the residues that are conserved between the secretin receptor and the CRF2ß receptor. Each of the 20 models was then subjected to local minimization, followed by ten independent cycles of Monte Carlo sampling of backbone and side chain conformations at 600 K, each simulation lasting for more than 216 h. The lowest energy conformation for each of the 20 models was retained and further refined by Monte Carlo simulation at 100 K for 36 h.

Homology modeling of the transmembrane domain of the rat secretin receptor. The sequence of the transmembrane helices of the rat secretin receptor were aligned to those of bovine rhodopsin (PDB code: 1U18) using the "cold spot" method (Frimurer and Bywater, 1999). A homology model of the transmembrane helical bundle of the rat secretin receptor was then built in ICM. The packing of the model was further optimized by side chain sampling and backbone minimization.

Alignment of the amino-terminal domain of the secretin receptor with its transmembrane helical bundle. A distance restraint was imposed between the WDN epitope and the top of the sixth transmembrane segment. A second distance restraint was imposed between the carboxyl end of the amino-terminal domain and the amino-terminal end of helix 1. The amino-terminal domain was then docked onto the helical bundle by Monte Carlo sampling of the six positional variables of the amino-terminal domain. The contact between the two domains was further optimized by side chain sampling to generate a preliminary model of the secretin receptor incorporating both the amino terminus and receptor core (Fernandez-Recio et al., 2003).

Docking of secretin to the full model of the secretin receptor. The initial conformation of secretin to be utilized in docking studies was taken from a previous solution-phase NMR determination of the porcine secretin structure (differs from the rat sequence only with a glutamine in position 14 in place of an arginine) (Clore et al., 1988). A maximum

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

distance of 3.5 Å was imposed between the residues Phe^{6}, Arg^{12}, Leu^{13}, Gln^{14}, Arg^{18}, Leu^{22}, Leu^{26} of secretin and the residues Val^{4}, Val^{6}, Val^{103}, Pro^{38}, Arg^{14}, Leu^{17}, Leu^{36} of the secretin receptor, respectively. Secretin was first docked to the grid potentials derived from the model of the intact secretin receptor in ten independent runs. During the grid docking, the six positional variables of secretin were actively sampled, with rigid backbone and flexible side chain during local minimization cycles. All of the ten independent runs converged to a single solution after 8 h of docking. Subsequently, the secretin/secretin receptor complex was refined. During refinement, the side chains of both secretin and the secretin receptor were actively sampled, while the backbone of secretin and residues 1-43 and 102-110 of the secretin receptor were allowed to be more flexible during cycles of local minimization. The refinement process typically lasted for 70 h, with the lowest energy conformation among the ten independent runs retained. Finally, the two new residue-residue spatial approximation constraints determined by photoaffinity labeling in the current work were assessed in the model. The health of the models was established by PROCHECK and WHAT_CHECK evaluations (Hooft et al., 1996; Laskowski et al., 1993).

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Results

Characterization of photolabile probes. Both the Bpa^{21} and Bpa^{23} probes were synthesized by manual solid phase techniques and purified by reversed-phase HPLC, and their identities were verified by mass spectrometry. They were functionally characterized in a competition radioligand-binding assay. As shown in Figure 1, both probes bound to the secretin receptor specifically and saturably, although with affinities lower than that of natural secretin (K_i values in nM: secretin, 2.7 ± 0.1; Bpa^{21} probe, 17.9 ± 1.2; Bpa^{23} probe, 47.1 ± 14.2).

The probes were then tested for their ability to stimulate intracellular cAMP accumulation in CHO-SecR cells. As shown in Figure 1, both probes were fully efficacious agonists, stimulating full intracellular cAMP responses that were not different from those achieved in response to natural secretin. They were, however, less potent than natural secretin (EC_{50} values in nM: secretin, 0.06 ± 0.01; Bpa^{21} probe, 0.37 ± 0.02; Bpa^{23} probe, 5.22 ± 0.97).

Photoaffinity labeling of the secretin receptor. We explored the ability of the Bpa^{21} and Bpa^{23} probes to covalently label the secretin receptor. As shown in Figure 2, both probes specifically and saturably labeled the secretin receptor, with the labeled protein bands migrating at approximate M_r = 70,000 and shifting to approximate M_r = 42,000 after deglycosylation with endoglycosidase F. This labeling was competed by secretin in a concentration-dependent manner (Bpa^{21} probe, IC_{50} = 18.1 ± 2.8 nM; Bpa^{23} probe, IC_{50} = 3.8 ± 0.5 nM). No radioactive band was observed in the affinity-labeled non-receptor bearing CHO cell membranes.

Identification of sites of labeling by the Bpa^{21} and Bpa^{23} probes. To gain insights into the domains of labeling by each of the probes, CNBr cleavage was first used. The secretin receptor contains nine Met residues and theoretically, CNBr cleavage would yield ten fragments

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

ranging in molecular weight from 1 to 11 kDa, three of which contain sites of glycosyaltion (Fig. 3). As shown in Figure 3, CNBr cleavage of the secretin receptor labeled by each probe resulted in a fragment migrating at approximate M_r = 19,000 and shifting to M_r = 10,000 after deglycosylation. This pattern was the same for both probes, indicating they might label the same region of the receptor. Based on the molecular masses of the attached probes (Bpa^{21} probe, 3,173 Da; Bpa^{23} probe, 3,216 Da) and on the glycosylated nature of the labeled bands, two fragments were felt to be candidates (segments including amino acid residues 1-51 and 74-123). Both of these fragments are within the amino-terminal tail of the secretin receptor, with one at the amino-terminal end of the receptor (fragment 1) and the other (fragment 3) adjacent to the first transmembrane segment (Fig. 3). Previous experience with the electrophoretic migration of each of these CNBr fragments makes the first fragment most likely (Dong et al., 1999b, Zang et al., 2003).

To definitively establish receptor residues 1-51 as the region of labeling, a well-characterized receptor mutant that incorporates an HA epitope within this fragment (SecR-HA37) (Dong et al., 1999b) was used in immunoprecipitation experiments. Figure 4 shows that the M_r = 19,000 CNBr fragment radiolabeled by Bpa^{21} and Bpa^{23} probes was immunoprecipitated by the anti-HA monoclonal antibody and precipitation was prevented in the presence of excess HA peptide.

Endoproteinase Lys-C that cleaves at the carboxyl-terminal side of lysine residues was next used to further localize the region of covalent labeling by each probe. Figure 5 shows that endoproteinase Lys-C cleavage of labeled CNBr fragment 1 yielded a labeled fragment migrating at approximate M_r = 6,000 that did not shift further after deglycosylation. This identified the non-glycosylated 30 residue fragment at the distal amino terminus of the secretin receptor as the domain of labeling by each of the probes.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

To further localize the region of labeling by each of the probes, two previously established secretin receptor mutant constructs, V13M-HA37 and V16M-HA37 (Dong et al., 2000), were used. Both receptor mutants were specifically and saturably labeled with each probe (Fig. 6). CNBr cleavage of the V13M-HA37 mutant labeled with the Bpa^{21} probe yielded a fragment migrating on a 10% NuPAGE gel at approximate M_r = 19,000 that shifted to M_r = 9,000 after deglycosylation, representing the fragment between Arg^{14} and Met^{51}. CNBr cleavage of the V16M-HA37 mutant labeled with the Bpa^{21} probe yielded a non-glycosylated fragment migrating at approximate M_r = 5,000, representing the fragment between Ala^{1} and Met^{16} (Fig. 6). Taken together, these data indicate that the site of labeling with the Bpa^{21} probe was within the small segment between Arg^{14} and Val^{16}.

CNBr cleavage of the V13M-HA37 and the V16M-HA37 receptor mutants labeled with the Bpa^{23} probe yielded fragments migrating similarly on a 10% NuPAGE gel at approximate M_r = 19,000 that shifted to M_r = 9,000 after deglycosylation, representing the fragment between Arg^{14} and Met^{51}. Taking into account the endoproteinase Lys-C cleavage data above, the site of labeling with the Bpa^{23} probe was within the region between Leu^{17} and Lys^{30}.

Radiochemical Edman degradation sequencing of the purified labeled fragments resulting from the CNBr cleavage of the labeled V13M-HA37 (for the Bpa^{21} probe) and V16M-HA37 (for the Bpa^{23} probe) receptor mutants was performed. Shown in Figure 7 are the profiles of eluted radioactivity in which a peak was found in cycle 2 for the Bpa^{21} probe and in cycle 5 for the Bpa^{23} probe. This identified Arg^{15} and Arg^{21} of the secretin receptor as the sites of labeling with the Bpa^{21} probe and for the Bpa^{23} probe, respectively.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Refinement of the NMR models of the mouse CRF2ß receptor amino terminus. PROCHECK and WHAT_CHECK were utilized to compare the quality of the ensemble of 20 models of receptor residues 37-122 before and after Monte Carlo simulation (Table 1). The results from PROCHECK show that after simulation, 25% more residues occupied the most favored regions, while the percentages of residues in the generously allowed and disallowed regions dropped from 11.6% and 3.3% to 1.9% and 1.0%, respectively. Similarly, results from WHAT_CHECK indicated significant improvement in quality of second generation packing, Ramachandran plot appearance, $\chi 1/\chi 2$ rotamer normality, and backbone conformation after simulation.

Recently, Nederveen et al. (Nederveen et al., 2005) published the RECOORD database, in which more than 500 NMR models from the Protein Data Bank were recalculated. The largest improvement in the model quality was achieved by restrained molecular dynamics in a hydrated environment. Comparing their study with ours shows that refinement in ICM improved the model quality similar to that of Nederveen's study in all categories in PROCHECK and WHAT_CHECK, bringing the overall quality of the amino terminus of the CRF2ß receptor to that of an average NMR model. Note that although the Z-score of the backbone conformation after refinement was still -8.4, it represented an improvement over the Z-score of the original models of -11.4. Consistent with this, Nederveen's refinement also did not improve the backbone conformations significantly.

To see how the refinement affected the overall structure of the models, we calculated the pairwise C_α root mean square deviation (rmsd) by superimposing the C_α of the stable central core residues 59-64, 70-81, and 100-113. For each of the 20 models, we superimposed the model before and after refinement, the resultant C_α rmsd for the central core and the remaining

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

disordered region were 1.22 ± 0.16 Å and 7.8 ± 1.6 Å, respectively. These values reflected the degree of movement in the central core and the disordered region during refinement. As a reference, the pairwise C_α rmsd for the central core and the disordered region within the ensemble of the original 20 NMR models were 0.84 ± 0.17 Å and 7.7 ± 1.9 Å, respectively, which indicate the intrinsic flexibility/uncertainty in the two regions in the original NMR models. These results demonstrate that our simulation enables large movement in the disordered region, while limiting the movement of the stable central core within experimental uncertainty, as shown in Figure 8.

Homology modeling of the amino terminus of the secretin receptor. The aligned amino terminal regions of secretin and CRF2ß receptors share 19% sequence identity and 36% sequence homology, with the three disulfide bonds that are conserved throughout this family (Grace et al., 2004; Lisenbee et al., 2005). An initial homology model of the amino terminus of the secretin receptor was produced using ICM, and was refined based on a subset of the original NMR distance restraints of the CRF2ß receptor amino terminus that included 115 distance restraints between residues that are conserved among secretin and CRF2ß receptors. The backbone and side-chain dihedral angles of the secretin receptor models were then sampled by BPMC in the presence of these distance restraints. This refinement was necessary because of the low sequence identity and the highly undefined nature of the amino-terminal domain outside of its central core, as shown in the CRF2ß receptor in Figure 8.

The ensemble of 20 rat secretin receptor amino terminus models after homology modeling and refinement is shown in Figure 9. We calculated the pairwise C_α rmsd within the ensemble by superimposing the C_α of the stable central core residues 43-48, 53-64, and 82-93. The C_α rmsd for the central core and the remaining disordered region were 2.11 ± 0.60 Å and 11.3 ± 2.0 Å, respectively. Comparing these numbers with those of the original NMR models of

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

the amino terminus of the CRF2ß receptor (0.84 ± 0.17 Å and 7.7 ± 1.9 Å) show a wider spread of the supposed stable central core, probably reflecting the smaller number of distance restraints used in the modeling of this region of the secretin receptor.

Only seven of the 20 models were found to contain an intact salt bridge between Asp^{49} and Arg^{83}, corresponding to the salt bridge between Asp^{65} and Arg^{101} proposed to be present in the amino terminus of the mouse CRF2ß receptor. We inspected the distance restraints deposited in Protein Data Bank and found that there was no direct distance restraint connecting Asp^{65} and Arg^{101} in the original NMR models, even though this salt bridge was reported as a prominent feature of those conformations.

Molecular model of the intact secretin receptor. To study how secretin interacts with its receptor, we first constructed an initial model of the receptor including the receptor amino terminus and the seven-transmembrane segment bundle. A preliminary model of the transmembrane helical bundle domain was generated using a previously-reported "cold-spot" method (Frimurer and Bywater, 1999). We attached the seven secretin receptor amino-terminal domain models having an intact Asp^{49}-Arg^{83} salt bridge (that were most similar to the reported conformation of the CRF2ß receptor) to this receptor transmembrane domain. The relative orientation of the two domains was restrained by the connection of the two domains at the amino-terminal end of helix 1 and by the proposed interaction between the WDN sequence and the transmembrane domain (Dong et al., 2006). Seven models of the receptor that were consistent with these restraints were generated.

Secretin was first docked to the grid representation of the seven full models to maximize sampling efficiency. Due to the rigid nature of the receptor in grid docking, we subsequently carried out flexible backbone refinement. The C_α distances between cross-linked residues in

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

secretin and the secretin receptor from photoaffinity labeling studies are shown for three of the best models in Table 2. Note that the two photoaffinity labeling groups, *p*-benzoyl-*L*-phenylalanine and *p*-benzoylbenzoyl-*L*-lysine, can span distances of 10 Å and 15 Å, respectively. Therefore, the seven distance restraints between secretin and the amino-terminal domain of the secretin receptor coming from previous photoaffinity labeling studies were all satisfied.

Most importantly, the last two entries of Table 2 show that the two pairs of cross-linked residues being demonstrated for the first time in the current study; Arg^{21} of secretin to receptor residue Arg^{15} and Leu^{23} of secretin to receptor residue Arg^{21}, were found to be within cross-linking distances in our models, even though these two sets of distance restraints had not been utilized during the peptide docking and refinement. Figure 10 shows the best working model for secretin docking to the receptor amino terminus and highlights the potential accessibility of the WDN sequence within the amino terminus of the receptor previously postulated as playing an endogenous agonist role by interacting with the body of the secretin receptor (Dong et al., 2006).

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Discussion

The molecular basis of natural peptide ligand binding to the amino-terminal domain of Class II G protein-coupled receptors and the conformation of this domain are of substantial current interest. This is based on a series of observations showing that this region is key for agonist binding and receptor activation for multiple members of this family (Cao et al., 1995; Gourlet et al., 1996; Juppner et al., 1994). The current work has focused on these aspects of the prototypic secretin receptor. This receptor was the first member of the Class II GPCR family to be cloned and displays conservation of important domains in the natural peptide agonist ligand, as well as receptor structural signatures and functional mechanisms examined. This model system was chosen, since extensive experimental data are already available to facilitate the meaningful docking of secretin to this receptor.

Useful structural insights have come from the solution of the NMR structure of the amino-terminal domain of another Class II GPCR, the CRF2ß receptor (Grace et al., 2004). In the current work, we have carefully examined those reported structures and the distance constraints that had contributed to the reported models. In the current report, the CRF2ß receptor model was refined and the insights coming from it were applied to the analogous region of the structurally-related secretin receptor. This likely represents a more meaningful template for a homology model for this region of the secretin receptor than the amino-terminal region of ribonuclease Mc1, which was previously utilized as template in the absence of the CRF2ß receptor structure (Dong et al., 2003). Indeed, the core regions of the amino-terminal structures of the secretin receptor and of the CRF2ß receptor have now been found to be very similar.

Most of the experimentally-defined sites of residue-residue approximation that have been determined for the secretin receptor are within the first 38 residues of the distal amino terminus in

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

poorly-constrained regions of this structure. Of particular interest, the regions of the secretin receptor that are homologous to the relatively stable core of the amino terminus of the CRF2ß receptor have not been covalently labeled by any of positions of photolabile residues spread throughout the pharmacophoric domain of secretin receptor probes that have been used to date. This raises the possibility that this stable core provides a lattice deep in the structure with the portions of the receptor that interact with the natural peptide ligands situated above it.

Although placement of the peptide ligands within the amino-terminal domain is crucial to understanding high affinity binding for Class II receptors, an additional key factor to resolving the mechanism of receptor activation by agonists is the positioning of both the peptide ligand and receptor amino-terminal domain relative to the heptahelical transmembrane core.

There is currently little empirical data to direct the organization of the receptor amino-terminal domain and the transmembrane receptor core, and this has led to disparate models of the relative positioning of these two domains. For the CRF2β receptor, the orientation used was based on a proposed interaction between a negatively charged surface of the receptor amino-terminal domain and positively charged residues present in the extracellular loops of the receptor core (Grace et al., 2004), while the VPAC1 receptor amino-terminal domain was arbitrarily placed above the receptor core in the recently proposed model of the intact VPAC1 receptor (Tan et al., 2006). In both cases, the docking of peptides to their receptors relied significantly on general consideration of the functional roles of the amino-terminal and carboxyl-terminal ends of the peptide ligands across Class II receptors that are dependent upon the orientation of the receptor amino-terminal domain and receptor core. These models are further complicated by lack of direct knowledge of regions of the amino terminus that are disordered in the NMR structure that form the basis for receptor amino-terminal domain modeling.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

In the current report, we have utilized two restraints in determining how to best position the amino-terminal domain of the secretin receptor with its helical bundle domain. The first restraint is the integrity of the peptide backbone involved in connecting the two domains, bringing the amino-terminal end of the first transmembrane helix to the carboxyl-terminal end of the receptor amino terminus. This is a common feature of each of the receptor models, however, the innate flexibility of the distal end of the receptor amino-terminal domain allows multiple surfaces to potentially abut the receptor core. The second restraint is specific to the secretin receptor, and assumes that the WDN sequence within the amino terminus of the receptor interacts with the helical bundle above transmembrane segment six (Dong et al., 2006). This was based on a series of studies demonstrating the ability of this receptor sequence to act as an endogenous agonist ligand, with its site of interaction demonstrated by covalent labeling of the receptor body using two distinct photoaffinity labeling probes derived from the critical WDN tripeptide (Dong et al., 2006). Of note, similar activation of VPAC1 and calcitonin receptors by homologous endogenous peptide sequences has also been shown to occur, suggesting that such activation may be a common mechanism within the Class II GPCR family (Dong et al., 2006).

Secretin was docked to this preliminary model of the secretin receptor utilizing the seven published photoaffinity labeling spatial approximation constraints. Of particular interest, with the peptide docked this way and given the proposed orientation of the two domains of the secretin receptor, the amino terminus of the peptide was found to be adjacent to the helical bundle in the region it has been independently shown to photoaffinity label (Dong et al., 2004). Additionally, the two spatial approximation constraints newly identified in the current report were also compatible with this docking and orientation. It is encouraging that these independently-derived observations are fully compatible with the working model.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Critical comparison of the currently-proposed model of the mode of peptide docking to the amino-terminal domain of the secretin receptor with those of the astressin/CRF-occupied CRF2ß receptor (Grace et al., 2004) and of the vasoactive intestinal polypeptide-occupied VPAC1 receptor (Tan et al., 2006), shows clear differences. As discussed above, the orientations of the amino-terminal domains with the helical bundle domains of these receptors are all different. The secretin receptor orientation was guided by the WDN motif and was, therefore, fully consistent with its interaction with the relevant region of the receptor body. However, in the CRF2β receptor model, the amino-terminal domain interacts with the transmembrane domain through the "back side" of the model displayed in Figure 8 and Figure 9. As a consequence, Leu^{64}-Gln^{66} of the CRF2β receptor, corresponding to the WDN motif, is facing upward and away from the analogous region of that receptor. In the VPAC1 receptor model, the amino-terminal domain makes contact with the transmembrane domain near the end of the second β-sheet containing residues Ile^{61}-Cys^{63} and Pro^{77}-Gly^{79}. In this orientation, the analogous region to the WDN motif of the secretin receptor is facing sideward. However, there was no clear rationalization for this orientation with respect to the helical bundle.

Comparison of the peptide-binding regions of these receptors also reflects differences. In the current model of the secretin receptor, the peptide is situated between the distal amino terminus, the region containing the majority of the sites of cross-linking, and part of the stable core. Glu^{9}-Ala^{17} of secretin interacts with the tip of the first β-sheet including residues Asp^{49}-Ser^{52}, and the disordered loop comprising residues Lys^{77}-Ser^{80}. Gln^{20}-Val^{27} of secretin interacts with the disordered loop including residues Glu^{31}-Gly^{37} (Figure 10). In the proposed CRF2β receptor structure, regions proposed to represent the peptide-binding site based on chemical-shift perturbations upon antagonist (astressin) binding include Ile^{67}-Thr^{69}, Gly^{90}-Asn^{93}, Glu^{102}-Cys^{103}, and Arg^{112}-Ser^{116}. The peptide-binding region in both the secretin and CRF2β receptor models

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

includes the tip of the first β-sheet and the disordered loop near the palm of the second β-sheet. In the proposed VPAC1 receptor structure, the peptide-binding interface was inferred from three pairs of cross-linked residues, at positions 6, 22, and 24 of the peptide ligand. Here, this region includes receptor residues Gln^{80}-Leu^{84}, Val^{101}-Cys^{105}, and Trp^{110}-Cys^{122}. Notably, this interface would be at the "bottom side" of the model of the secretin receptor displayed in Figure 8 and Figure 9, opposite to the proposed peptide-binding sites in both the currently-proposed model of the secretin receptor and that of the CRF2ß receptor. At present, we do not have sufficient empirical data to confirm whether the alternate models proposed for different receptors reflect divergence in mechanism of action or potential problems in generation of working models.

Another structural motif of substantial interest is the proposed intra-domain salt bridge between Asp^{65} and Arg^{101} of the CRF2ß receptor. Of particular interest, the region of the secretin receptor amino terminus that has been shown to have endogenous agonist activity is centered on the residue analogous to CRF2ß receptor residue Asp^{65}. If such a bond were present and stable in the secretin receptor (it would involve the secretin receptor residues Asp^{49} and Arg^{83}), it might make the endogenous agonist hypothesis less likely as a mechanism of activation of these receptors. In the secretin receptor amino terminus models that were built based on homology with the CRF2ß receptor structure, the salt bridge between Asp^{49} and Arg^{83} actually did not exist in the majority of conformations. Whether the absence of this salt bridge in the secretin receptor models is an artifact caused by an inadequate number of distance restraints in modeling or a genuine feature of the secretin receptor amino terminus remains to be explored. It should be noted, however, that there is no explicit salt-bridge present from the CRF2β NMR distance restraints, while mutation of Asp^{49} to glycine has no negative effect on the biological activity of the secretin receptor (Dong et al., 2006). It is also recognized that solvent-accessible salt bridges are easily broken. Such a breakable salt-bridge between Asp^{49} and Arg^{83} of the secretin receptor

is still consistent with the endogenous agonist hypothesis (Dong et al., 2006). However, in the best conformation of the secretin-occupied secretin receptor currently being proposed the WDN motif is readily accessible to its proposed target of action, even with the salt-bridge incorporated.

It is encouraging that, in the current work, multiple experimentally-derived residue-residue distance constraints were easily accommodated within the molecular model of the secretin receptor. With initial peptide docking accomplished based on the previously-described seven such constraints, it was even more encouraging that the currently-derived two additional such constraints were also then well accommodated. In the best working model, all existing experimentally-derived constraints were fully accommodated and the WDN motif was available for interaction with the body of the receptor.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Acknowledgements

The authors acknowledge the technical assistance of L.A. Bruins and secretarial assistance of E. Posthumus.

References

Abagyan R and Totrov M (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J Mol Biol* **235**:983-1002.

Abagyan R, Totrov, M., and Kuznetsov, D. (1994) Icm - a new method for protein modeling and design - applications to docking and structure prediction from the distorted native confirmation. *J Comput Chem* **15**:488-506.

Abagyan RA and Totrov M (1999) Ab initio folding of peptides by the optimal-bias Monte Carlo minimization procedure. *J Comput Phys* **151**:402-421.

Bisello A, Adams AE, Mierke DF, Pellegrini M, Rosenblatt M, Suva LJ and Chorev M (1998) Parathyroid hormone-receptor interactions identified directly by photocross-linking and molecular modeling studies. *J Biol Chem* **273**:22498-22505.

Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S and Schneider M (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* **31**:365-370.

Bordner AJ and Abagyan R (2006) Ab initio prediction of peptide-MHC binding geometry for diverse class I MHC allotypes. *Proteins* **63**:512-526.

Cao YJ, Gimpl G and Fahrenholz F (1995) The amino-terminal fragment of the adenylate cyclase activating polypeptide (PACAP) receptor functions as a high affinity PACAP binding domain. *Biochem Biophys Res Commun* **212**:673-680.

Cardozo T, Totrov M and Abagyan R (1995) Homology Modeling by the Icm Method. *Proteins* **23**:403-414.

Clore GM, Nilges M, Brunger A and Gronenborn AM (1988) Determination of the backbone conformation of secretin by restrained molecular dynamics on the basis of interproton distance data. *Eur J Biochem* **171**:479-484.

Dong M, Asmann YW, Zang M, Pinon DI and Miller LJ (2000) Identification of two pairs of spatially approximated residues within the carboxyl terminus of secretin and its receptor. *J Biol Chem* **275**:26032-26039.

Dong M, Li Z, Pinon DI, Lybrand TP and Miller LJ (2004) Spatial approximation between the amino terminus of a peptide agonist and the top of the sixth transmembrane segment of the secretin receptor. *J Biol Chem* **279**:2894-2903.

Dong M, Li Z, Zang M, Pinon DI, Lybrand TP and Miller LJ (2003) Spatial approximation between two residues in the mid-region of secretin and the amino terminus of its receptor. Incorporation of seven sets of such constraints into a three-dimensional model of the agonist-bound secretin receptor. *J Biol Chem* **278**:48300-48312.

Dong M, Pinon DI, Asmann YW and Miller LJ (2006) Possible endogenous agonist mechanism for the activation of secretin family G protein-coupled receptors. *Mol Pharmacol* **70**:206-213.

Dong M, Wang Y, Hadac EM, Pinon DI, Holicky E and Miller LJ (1999a) Identification of an interaction between residue 6 of the natural peptide ligand and a distinct residue within the amino-terminal tail of the secretin receptor. *J Biol Chem* **274**:19161-19167.

Dong M, Wang Y, Pinon DI, Hadac EM and Miller LJ (1999b) Demonstration of a direct interaction between residue 22 in the carboxyl-terminal half of secretin and the amino-terminal tail of the secretin receptor using photoaffinity labeling. *J Biol Chem* **274**:903-909.

Dong M, Zang M, Pinon DI, Li Z, Lybrand TP and Miller LJ (2002) Interaction among four residues distributed through the secretin pharmacophore and a focused region of the secretin receptor amino terminus. *Mol Endocrinol* **16**:2490-2501.

Fernandez-Recio J, Totrov M and Abagyan R (2003) ICM-DISCO docking by global energy optimization with fully flexible side-chains. *Proteins* **52**:113-117.

Frimurer TM and Bywater RP (1999) Structure of the integral membrane domain of the GLP1 receptor. *Proteins* **35**:375-386.

Ganguli SC, Park CG, Holtmann MH, Hadac EM, Kenakin TP and Miller LJ (1998) Protean effects of a natural peptide agonist of the G protein-coupled secretin receptor demonstrated by receptor mutagenesis. *J Pharmacol Exp Ther* **286**:593-598.

Gourlet P, Vilardaga JP, De Neef P, Waelbroeck M, Vandermeers A and Robberecht P (1996) The C-terminus ends of secretin and VIP interact with the N-terminal domains of their receptors. *Peptides* **17**:825-829.

Grace CR, Perrin MH, DiGruccio MR, Miller CL, Rivier JE, Vale WW and Riek R (2004) NMR structure and peptide hormone binding site of the first extracellular domain of a type B1 G protein-coupled receptor. *Proc Natl Acad Sci U S A* **101**:12836-12841.

Grauschopf U, Lilie H, Honold K, Wozny M, Reusch D, Esswein A, Schafer W, Rucknagel KP and Rudolph R (2000) The N-terminal fragment of human parathyroid hormone receptor 1 constitutes a hormone binding domain and reveals a distinct disulfide pattern. *Biochemistry* **39**:8878-8887.

Hadac EM, Ghanekar DV, Holicky EL, Pinon DI, Dougherty RW and Miller LJ (1996) Relationship between native and recombinant cholecystokinin receptors: role of differential glycosylation. *Pancreas* **13**:130-139.

Hooft RWW, Vriend G, Sander C and Abola EE (1996) Errors in protein structures. *Nature* **381**:272-272.

Juppner H, Schipani E, Bringhurst FR, McClure I, Keutmann HT, Potts JT, Jr., Kronenberg HM, Abou-Samra AB, Segre GV and Gardella TJ (1994) The extracellular amino-terminal region of the parathyroid hormone (PTH)/PTH-related peptide receptor determines the binding affinity for carboxyl-terminal fragments of PTH-(1-34). *Endocrinology* **134**:879-884.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680-685.

Laskowski RA, Macarthur MW, Moss DS and Thornton JM (1993) Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J Appl Crystallogr* **26**:283-291.

Lisenbee CS, Dong M and Miller LJ (2005) Paired cysteine mutagenesis to establish the pattern of disulfide bonds in the functional intact secretin receptor. *J Biol Chem* **280**:12330-12338.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH and Teller E (1953) Equation of State Calculations by Fast Computing Machines. *J Chem Phys* **21**:1087-1092.

Munson PJ and Rodbard D (1980) Ligand: a versatile computerized approach for characterization of ligand-binding systems. *Anal Biochem* **107**:220-239.

Nederveen AJ, Doreleijers JF, Vranken W, Miller Z, Spronk C, Nabuurs SB, Guntert P, Livny M, Markley JL, Nilges M, Ulrich EL, Kaptein R and Bonvin A (2005) RECOORD: A recalculated coordinate database of 500+proteins from the PDB using restraints from the BioMagResBank. *Proteins* **59**:662-672.

Nemethy G, Gibson KD, Palmer KA, Yoon CN, Paterlini G, Zagari A, Rumsey S and Scheraga HA (1992) Energy Parameters in Polypeptides .10. Improved Geometrical Parameters and Nonbonded Interactions for Use in the Ecepp/3 Algorithm, with Application to Proline-Containing Peptides. *J Phys Chem* **96**:6472-6484.

Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M and Miyano M (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**:739-745.

Pearson RK, Miller LJ, Hadac EM and Powers SP (1987) Analysis of the carbohydrate composition of the pancreatic plasmalemmal glycoprotein affinity labeled by short probes for the cholecystokinin receptor. *J Biol Chem* **262**:13850-13856.

Pham V, Wade JD, Purdue BW and Sexton PM (2004) Spatial Proximity between a Photolabile Residue in Position 19 of Salmon Calcitonin and the Amino Terminus of the Human Calcitonin Receptor. *J Biol Chem* **279**:6720-6729.

Powers SP, Pinon DI and Miller LJ (1988) Use of N,O-bis-Fmoc-D-Tyr-ONSu for introduction of an oxidative iodination site into cholecystokinin family peptides. *Int J Pept Protein Res* **31**:429-434.

Tan YV, Couvineau A, Murail S, Ceraudo E, Neumann JM, Lacapere JJ and Laburthe M (2006) Peptide agonist docking in the N-terminal ectodomain of a class II G protein-coupled receptor, the VPAC1 receptor. Photoaffinity, NMR, and molecular modeling. *J Biol Chem* **281**:12792-12798.

Ulrich CD, 2nd, Holtmann M and Miller LJ (1998) Secretin and vasoactive intestinal peptide receptors: members of a unique family of G protein-coupled receptors. *Gastroenterology* **114**:382-397.

Ulrich CD, 2nd, Pinon DI, Hadac EM, Holicky EL, Chang-Miller A, Gates LK and Miller LJ (1993) Intrinsic photoaffinity labeling of native and recombinant rat pancreatic secretin receptors. *Gastroenterology* **105**:1534-1543.

Zang M, Dong M, Pinon DI, Ding XQ, Hadac EM, Li Z, Lybrand TP and Miller LJ (2003) Spatial Approximation between a Photolabile Residue in Position 13 of Secretin and the Amino Terminus of the Secretin Receptor. *Mol Pharmacol* **63**:993-1001.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Footnotes

M.D. and P.C.-H.L. are co-primary authors.

This work was supported by the National Institutes of Health grant DK46577 (LJM), the National Health and Medical Research Council of Australia (NHMRC) grant 436780 (PMS) and by the Fiterman Foundation. PMS is a NHMRC Principal Research Fellow.

Address for correspondence to: Laurence J. Miller, M.D., Mayo Clinic, 13400 East Shea Blvd., Scottsdale, AZ 85259, Telephone: (480) 301-6650; Fax: (480) 301-6969;
E-mail: miller@mayo.edu

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Figure Legends

Fig. 1. Functional characterization of the Bpa^{21} and Bpa^{23} probes**.** *Left* shows the competition-binding curves of increasing concentrations of unlabeled Bpa^{21} and Bpa^{23} probes to displace the binding of radioligand, ^{125}I-[Tyr^{10}]rat secretin, to membranes from CHO-SecR cells. Values represent percentages of maximal saturable binding that were observed in the absence of competitor. They are expressed as means ± S.E.M. of duplicate data from three independent experiments. The absolute values for maximal binding in these experiments were 3,282 ± 45 cpm. *Right* shows intracellular cAMP responses to increasing concentrations of the Bpa^{21} and Bpa^{23} probes in CHO-SecR cells. There were no significant differences in basal and maximal intracellular cAMP levels in CHO-SecR by both probes and secretin. The basal levels of intracellular cAMP were 1.3 ± 0.4 pmol/million cells, and the maximal levels reached 198 ± 31 pmol/million cells. Values are expressed as means ± S.E.M. of data from three independent experiments, with data normalized relative to the maximal response to secretin.

Fig. 2. Photoaffinity labeling of the secretin receptor. Shown are representative 10% SDS-PAGE gels used to separate the products of photoaffinity labeling of CHO-SecR membranes by the Bpa^{21} (*top*) and Bpa^{23} (*bottom*) probes in the absence and presence of increasing concentrations of competing unlabeled secretin. The labeled secretin receptor with each probe migrated similarly at approximate M_r = 70,000 and shifted to approximate M_r = 42,000 after deglycosylation by endoglycosydate F (EF). Shown also are the densitometric analyses of three similar independent experiments by each probe, with data points expressed as means ± S.E.M. No bands were detected in affinity-labeled non-receptor-bearing CHO cell membranes.

Fig. 3. CNBr cleavage of the labeled secretin receptor. *Left* shows a diagram of the predicted sites of CNBr cleavage of the rat secretin receptor, along with the masses of protein cores of these

fragments. Residues are labeled sequentially from 1, after cleavage of the 22-residue signal peptide (Dong et al., 1999a). *Right* shows a representative autoradiograph of a 10% NuPAGE gel used to separate the products of CNBr cleavage of the secretin receptor labeled with the Bpa^{21} and Bpa^{23} probes. Cleavage of the native receptor labeled with each probe resulted in a band migrating at approximate M_r = 19,000 that shifted to M_r = 10,000 after deglycosylation with endoglycosidase F (EF). The best candidates representing this are the first fragment at the amino terminus of the receptor and the third fragment (*bold circles*).

Fig. 4. Immunoprecipitation of the CNBr fragments of the labeled secretin receptor. *Top* shows a diagram illustrating the theoretical sites of CNBr cleavage of the amino terminus of the HA-tagged secretin receptor construct (SecR-HA37). *Bottom* shows representative autoradiographs of 10% NuPAGE gels used for separating products of immunoprecipitation with monoclonal anti-HA antibody of CNBr fragments from cleavage of HA37-tagged secretin receptor labeled with each of the probes in the presence and absence of competing HA peptide. This provides the definitive identification of the fragment at the most distal end of the amino terminus as the affinity-labeled receptor domain for both the Bpa^{21} and Bpa^{23} probes.

Fig. 5. Endoproteinase Lys-C cleavage of the labeled secretin receptor. *Top* shows a diagram illustrating the theoretical sites of endoproteinase Lys-C cleavage and masses of expected fragments from sequential endoproteinase Lys-C cleavage of the CNBr fragment 1 (Ala^1–Met^{51}) of the secretin receptor. *Bottom* shows that endoproteinase Lys-C cleavage of CNBr fragment 1 from both native and deglycosylated secretin receptor labeled with the Bpa^{21} or Bpa^{23} probe yielded non-glycosylated fragments migrating similarly at approximate M_r = 6,000. This is most consistent with labeling of the amino-terminal non-glycosylated fragment with each of the probes, representing the region of the receptor between residue Ala^1 and Lys^{30}.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

Fig. 6. Photoaffinity labeling and CNBr cleavage of the mutant secretin receptors. *Top left* shows that both the V13M-HA37 and V16M-HA37 mutant secretin receptors were able to be affinity-labeled saturably and specifically by the Bpa^{21} and Bpa^{23} probes. *Top right* shows that CNBr cleavage of the V13M-HA37 and V16M-HA37 mutants labeled with the Bpa^{21} probe yielded a glycosylated fragment with the core migrating at approximate M_r = 9,000 (Arg^{14}-Met^{51}) and a non-glycosylated fragment migrating at approximate M_r = 5,000 (Ala^{1}-Met^{16}), respectively. These data indicate that the site of labeling with the Bpa^{21} probe was between Arg^{14} and Val^{16}. However, CNBr cleavage of these two mutant receptors labeled with the Bpa^{23} probe both yielded fragments migrating at approximate M_r = 19,000 and shifting to approximate M_r = 9,000, indicating the site of labeling being within the fragment Leu^{17}-Met^{51}. *Bottom* shows theoretical sites of CNBr cleavage of the Ala^{1}–Met^{51} fragment of the V13M-HA37 and V16M-HA37 receptor mutants labeled with each of the noted probes.

Fig. 7. Identification of the photoaffinity-labeled receptor residues. Shown are representative radioactivity elution profiles from radiochemical Edman degradation sequencing of the fragments Arg^{14}–Met^{51} from the V13M-HA37 mutant receptor labeled with the Bpa^{21} probe (*left*) and Leu^{17}–Met^{51} from the V16M-HA37 mutant receptor labeled with the Bpa^{23} probe (*right*) . A peak eluted in radioactivity appeared in cycle 2 for the Bpa^{21} probe, representing receptor residue Arg^{15}, and in cycle 5 for the Bpa^{23} probe, representing receptor residue Arg^{21}.

Fig. 8. Ensemble of 20 models of the amino terminus of the CRF2ß receptor. The structurally-stable central core of the CRF2ß receptor amino terminus consisting of residues 59-64, 70-81, and 100-113, is colored in black for backbone carbon and blue for backbone nitrogen. The flanking flexible regions are colored green. The disulfide bonds are shown in yellow. *Left*. The original

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

NMR models. The C_α rmsd of the central core and the flexible regions among the 20 models were 0.84 ± 0.17 Å and 7.7 ± 1.9 Å, respectively. *Right*. The models after refinement in ICM. During refinement, the C_α rmsd change for the central core and the flexible regions were 1.22 ± 0.16 Å and 7.8 ± 1.6 Å, respectively.

Fig. 9. Homology models of the amino-terminal domain of the rat secretin receptor. *Left*. Ensemble of 20 models of the amino terminus of the secretin receptor. The structurally-stable central core, consisting of residues 43-48, 53-64, and 82-93, is colored in black for backbone carbon and blue for backbone nitrogen; the tip of the beta-hairpin, residues 49-52, is colored in orange for backbone carbon and blue for backbone nitrogen, and the flanking flexible regions are colored green. The C_α rmsd of the central core and the flexible regions among the 20 models were 2.11 ± 0.60 Å and 11.3 ± 2.0 Å, respectively. *Right*. Lack of salt bridge between Asp^{49} and Arg^{83} in models of the amino terminus of the rat secretin receptor. Out of the 20 final models of the amino terminus of the secretin receptor, 13 did not contain a salt bridge between Asp^{49} and Arg^{83}. Shown is one representative model.

Fig. 10. Refined conformation of the amino terminus of the secretin receptor. This structure accommodates nine photoaffinity labeling constraints and the three disulfide bonds shown to exist in the secretin receptor (model 1 from Table 2). It also demonstrates the potential availability of the WDN sequence (highlighted) as a possible endogenous agonist ligand that could interact with the body of the secretin receptor. This is illustrated two different ways: the left panel illustrates the entire sequence, while the right panel removes the distal ends of the receptor amino terminus to highlight the peptide-binding surface. *Left*, The amino-terminal domain is represented in yellow, the transmembrane domain in grey. The secretin peptide is colored blue-red from the amino terminus to the carboxyl terminus. The WDN motif and the three solvent accessible,

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

glycosylated Asn^{78}, Asn^{84}, and Asn^{106} are displayed in CPK representation, blue: nitrogen; red: oxygen; white: carbon. Cross-linking residues on the secretin peptide are displayed in xstick, the distance restraints are displayed as dotted lines. Residues labels are blue on the secretin peptide and red on the receptor. *Right,* The final docking pose of the secretin peptide in our secretin receptor model is displayed. The WDN motif is labeled in blue. Red patch represents residues in direct contact with secretin peptide. Orange patch represents residues within 3 Å of secretin peptide.

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 1

Improvement in structural quality of the mouse CRF2ß receptor amino-terminal domain after Monte Carlo simulation, comparison with RECOORD.

	Original ensemble of 20 NMR modes, residue 37 to 122	Ensemble of 20 models after Monte Carlo simulation	545 Original NMR models in RECOORD	545 NMR models after water refinement in RECOORD
WHAT_CHECK Z-scores				
2nd Generation packing quality	-4.3 ± 0.3	-2.2 ± 0.4	-3.5 ± 1.9	-2.5 ± 2.0
Ramachandran plot appearance	-7.1 ± 0.6	-3.6 ± 0.7	-4.6 ± 1.6	-3.4 ± 1.4
χ1/χ2 Rotamer normality	-6.0 ± 0.3	0.9 ± 0.9	-3.7 ± 1.5	-0.9 ± 1.0
Backbone conformation	-11.4 ± 2.1	-8.4 ± 1.1	-4.0 ± 3.0	-3.8 ± 2.7
PROCHECK results				
Most favored regions	43.3 ± 5.7 %	68.4 ± 5.6 %	69.3 ± 14.2 %	76.1 ± 11.3 %
Allowed regions	41.7 ± 5.4 %	28.7 ± 5.4 %	25.4 ± 10.6 %	19.6 ± 8.5 %
Generously allowed regions	11.6 ± 3.8 %	1.9 ± 2.1 %	3.9 ± 3.8 %	2.5 ± 2.1 %
Disallowed regions	3.3 ± 2.2 %	1.1 ± 1.0 %	1.3 ± 1.7 %	1.8 ± 1.8 %

Molecular Pharmacology Fast Forward. Published on May 2, 2007 as DOI: 10.1124/mol.107.035402
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 2

C_α distances between cross-linked residues in secretin and the secretin receptor in three of the best models after grid docking and flexible backbone refinement.

Secretin	Secretin Receptor	Model 1 (Å)	Model 2 (Å)	Model 3 (Å)
Phe^{6}	Val^{4}	12.5	9.2	6.9
Arg^{12}	Val^{6}	4.8	6.0	5.6
Leu^{13}	Val^{103}	15.4	18.9	16.0
Gln^{14}	Pro^{38}	12.1	10.3	12.8
Arg^{18}	Arg^{14}	6.8	6.6	5.5
Leu^{22}	Leu^{17}	6.1	9.1	5.7
Leu^{26}	Leu^{36}	5.2	5.8	6.2
Arg^{21}	Arg^{15}	10.4	10.6	10.3
Leu^{23}	Arg^{21}	12.6	10.8	11.0

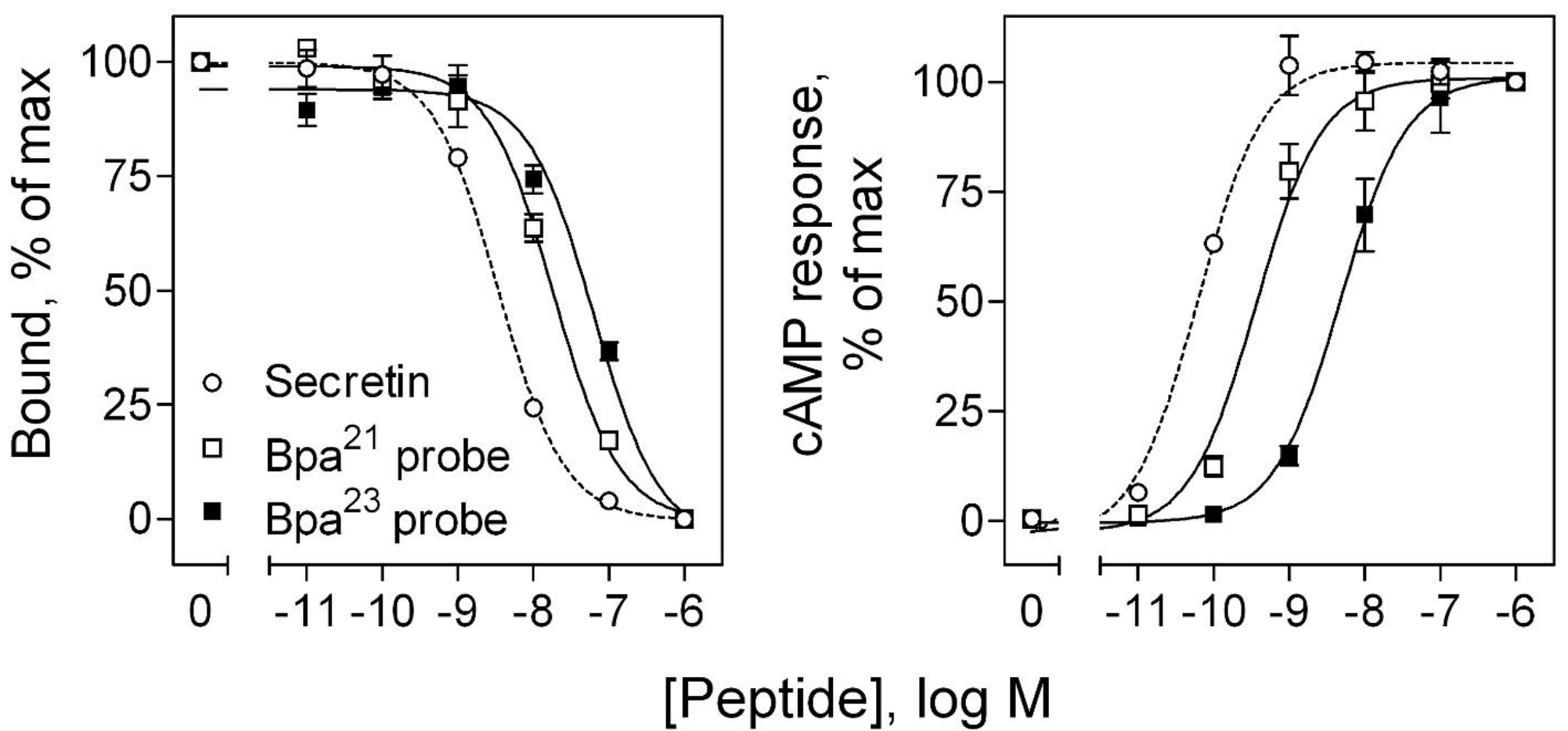
Bound, % of max
100
75
50
25
0
Secretin
Bpa21 probe
Bpa23 probe
0
-11
-10
-9
-8
-7
-6
cAMP response, % of max
100
75
50
25
0
0
-11
-10
-9
-8
-7
-6
[Peptide], log M

Figure 1

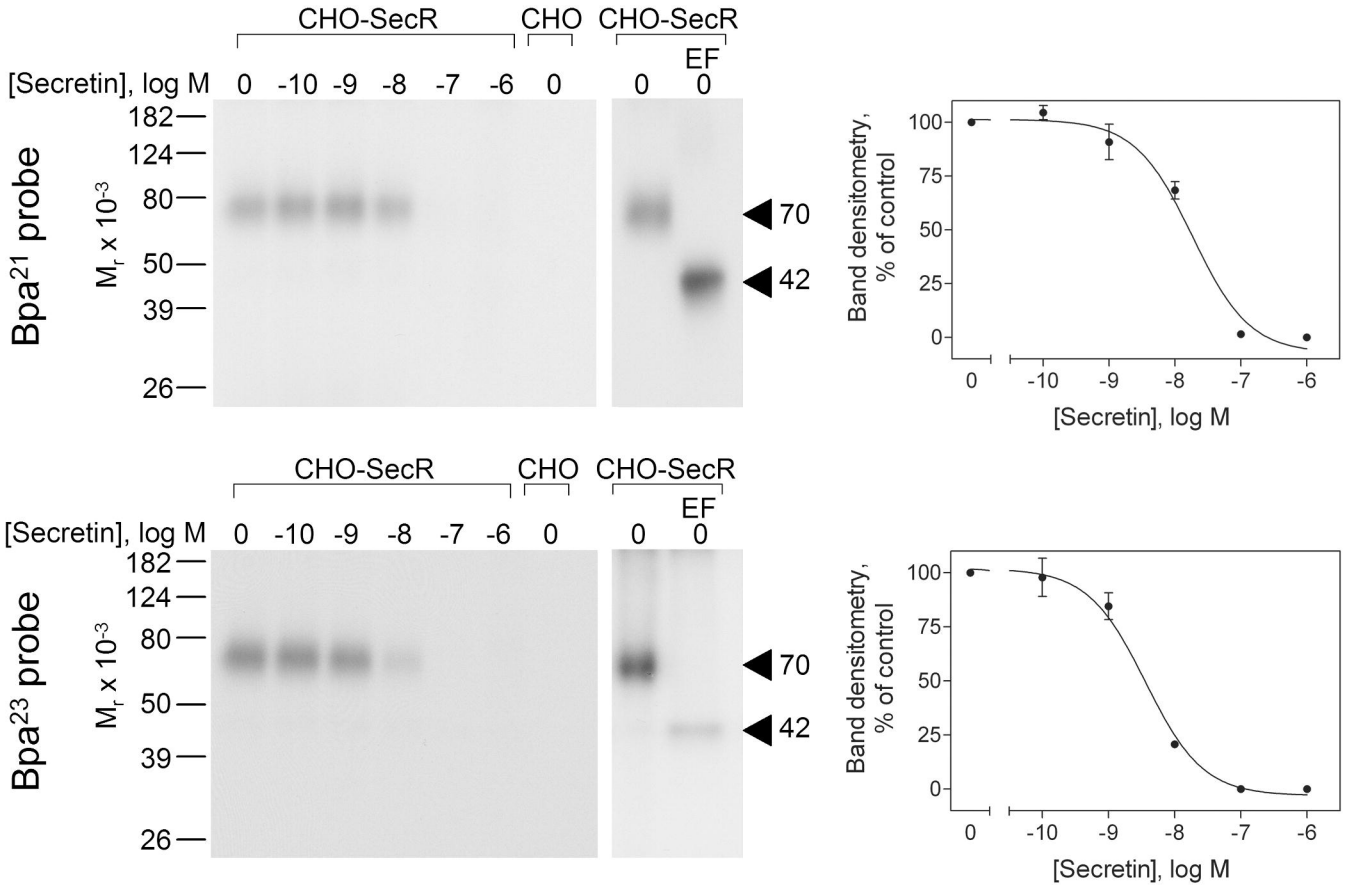
CHO-SecR
CHO
CHO-SecR
EF
[Secretin], log M
0 -10 -9 -8 -7 -6 0 0 0
Bpa21 probe
Mr x 10-3
182
124
80
50
39
26
70
42
Band densitometry, % of control
100
75
50
25
0
[Secretin], log M
Bpa23 probe

Figure 2

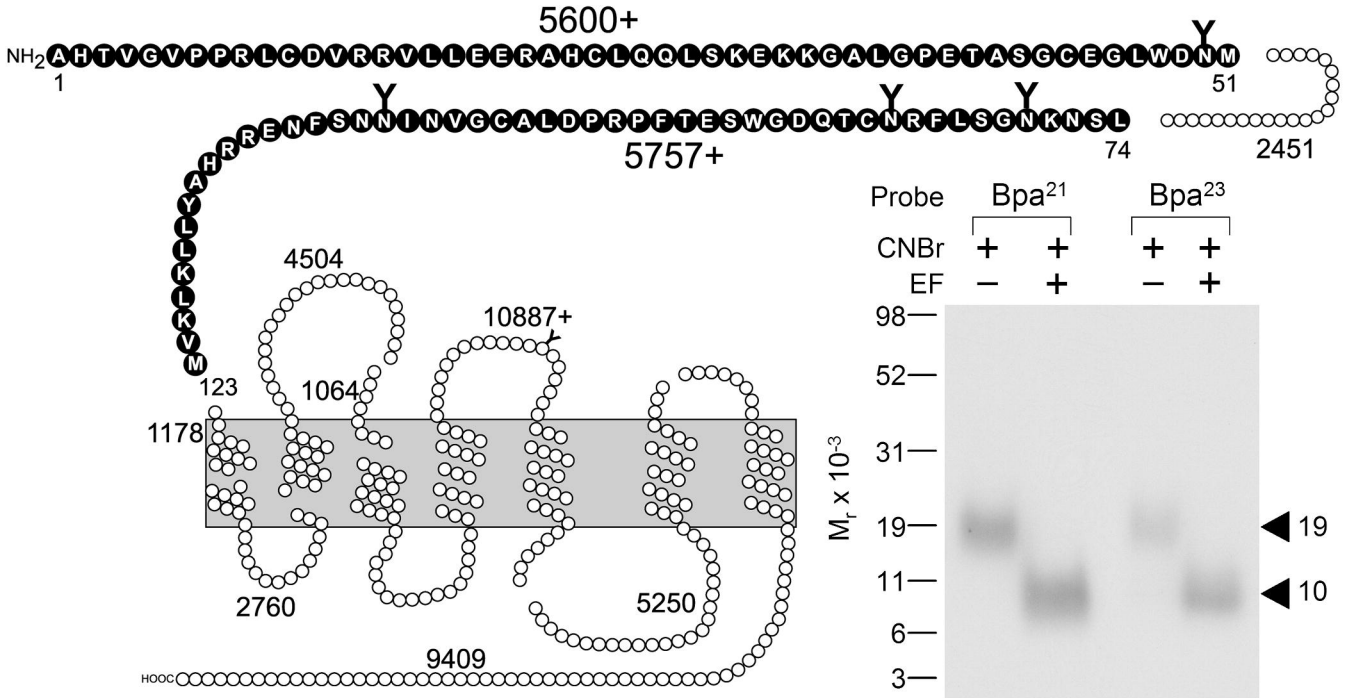
5600+
NH2
1
51
5757+
74
2451
4504
10887+
123
1064
1178
2760
5250
9409
HOOC
Probe
Bpa21
Bpa23
CNBr
EF
98
52
31
19
11
6
3
Mr x 10-3
19
10

Figure 3

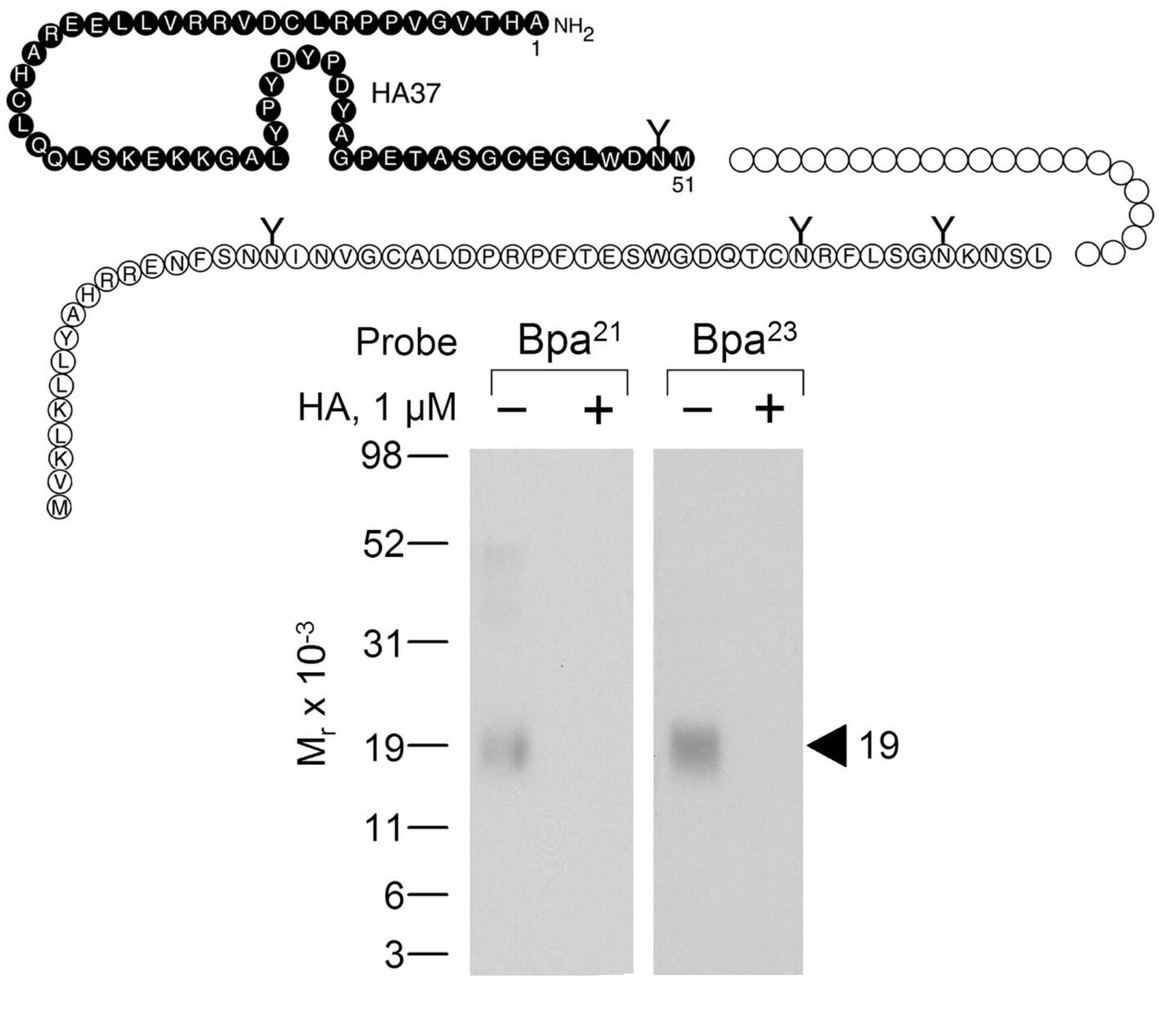
NH2
1
HA37
51
Probe
Bpa21
Bpa23
HA, 1 μM
–
+
–
+
98
52
31
19
11
6
3
Mr x 10-3
19

Figure 4

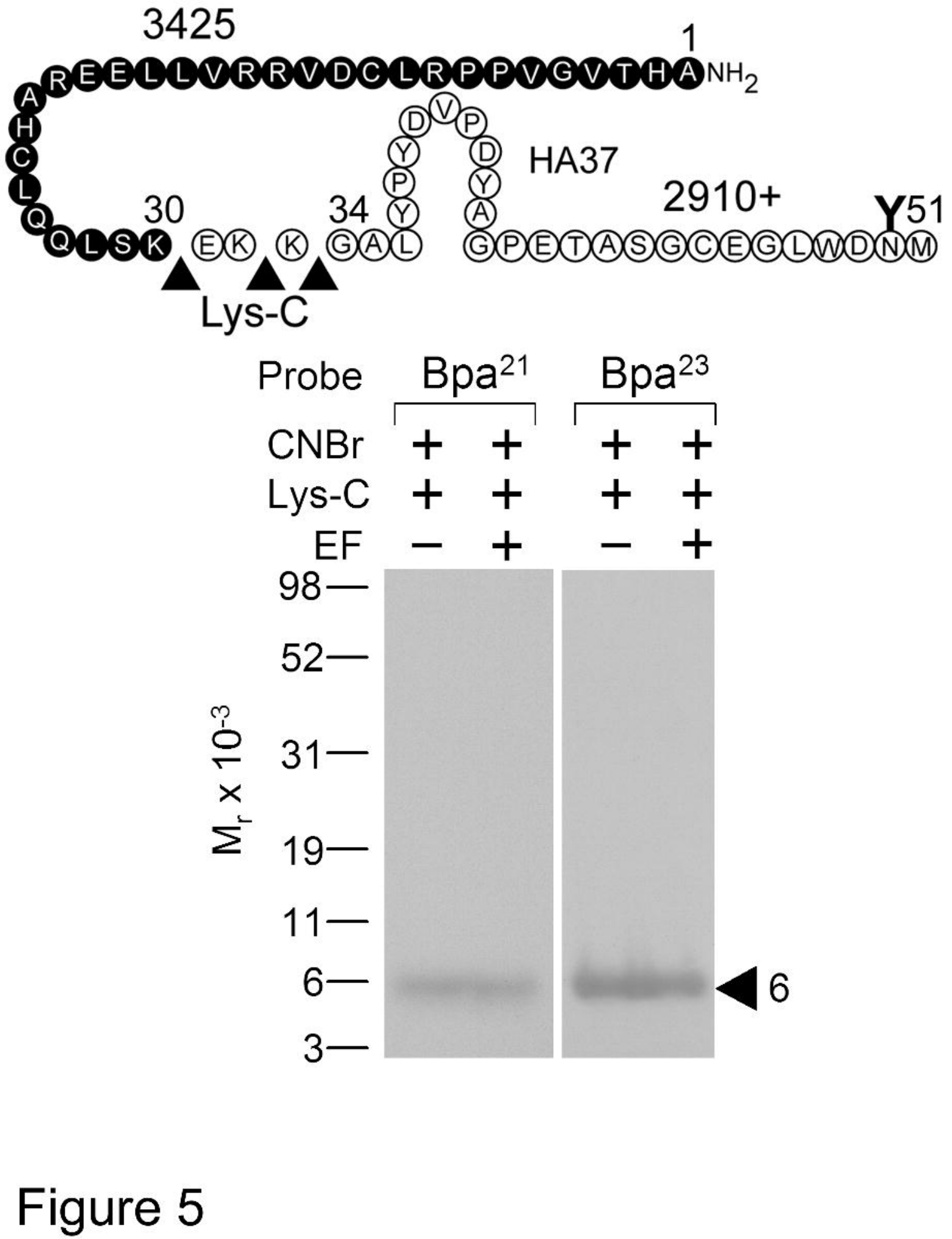
3425
1
NH2
HA37
2910+
Y51
30
34
Lys-C
Probe
Bpa21
Bpa23
CNBr
Lys-C
EF
98
52
31
19
11
6
3
Mr x 10-3
6

Figure 5

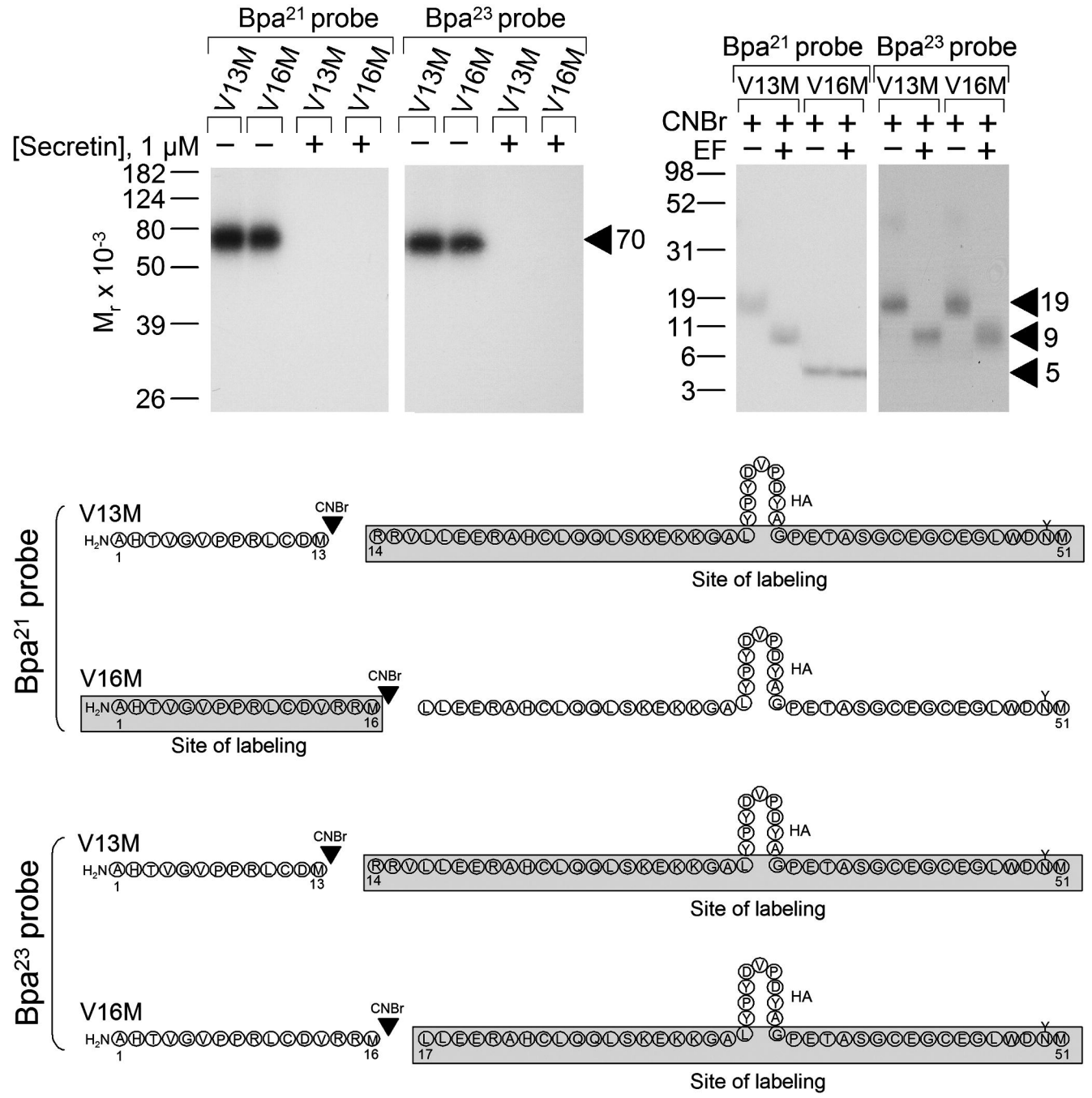
Bpa21 probe
Bpa23 probe
V13M
V16M
[Secretin], 1 μM
182
124
80
50
39
26
Mr x 10-3
70
CNBr
EF
98
52
31
19
11
6
3
19
9
5
HA
Site of labeling

Figure 6

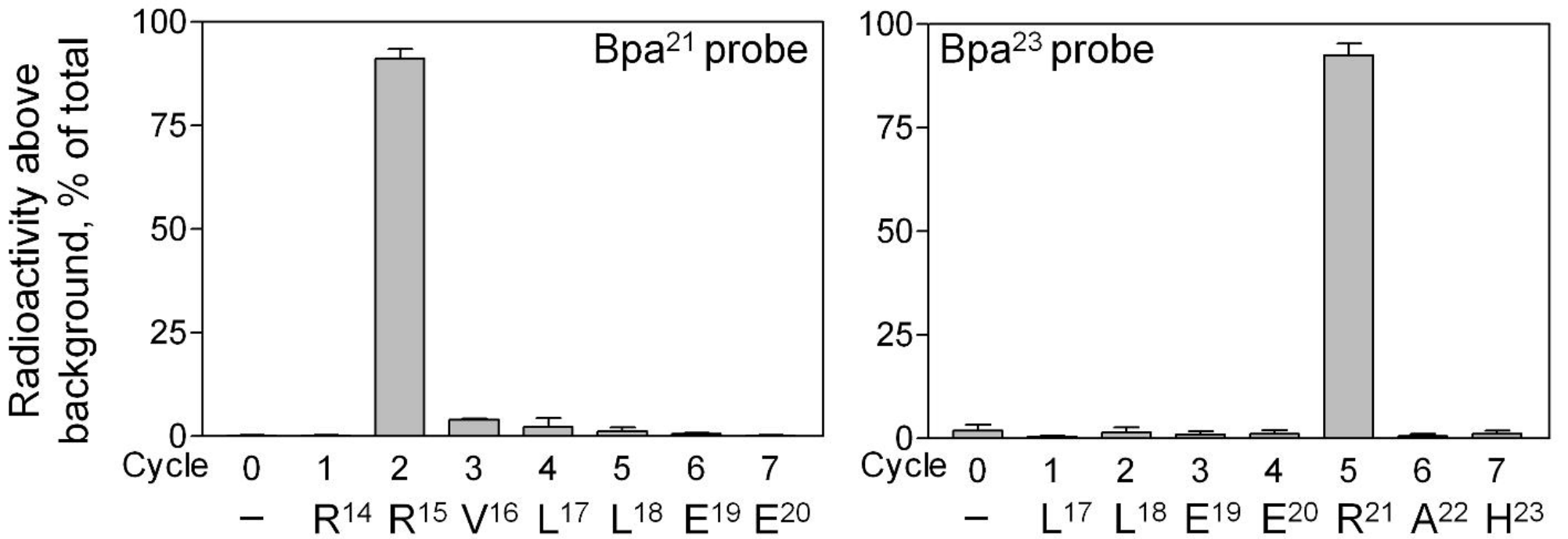
Radioactivity above background, % of total
100
75
50
25
0
Bpa21 probe
Cycle 0 1 2 3 4 5 6 7
– R14 R15 V16 L17 L18 E19 E20
Bpa23 probe
Cycle 0 1 2 3 4 5 6 7
– L17 L18 E19 E20 R21 A22 H23

Figure 7

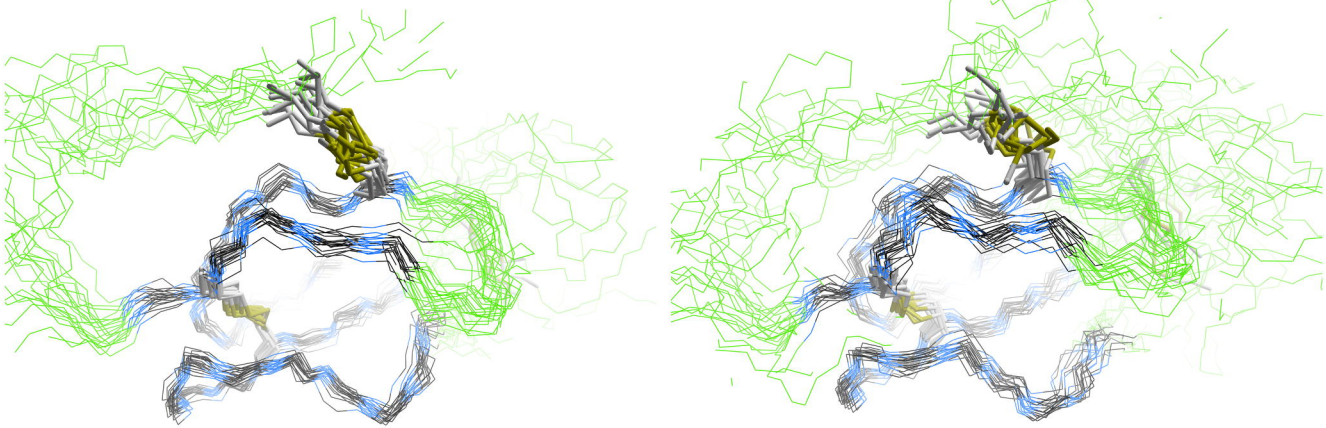

Figure 8

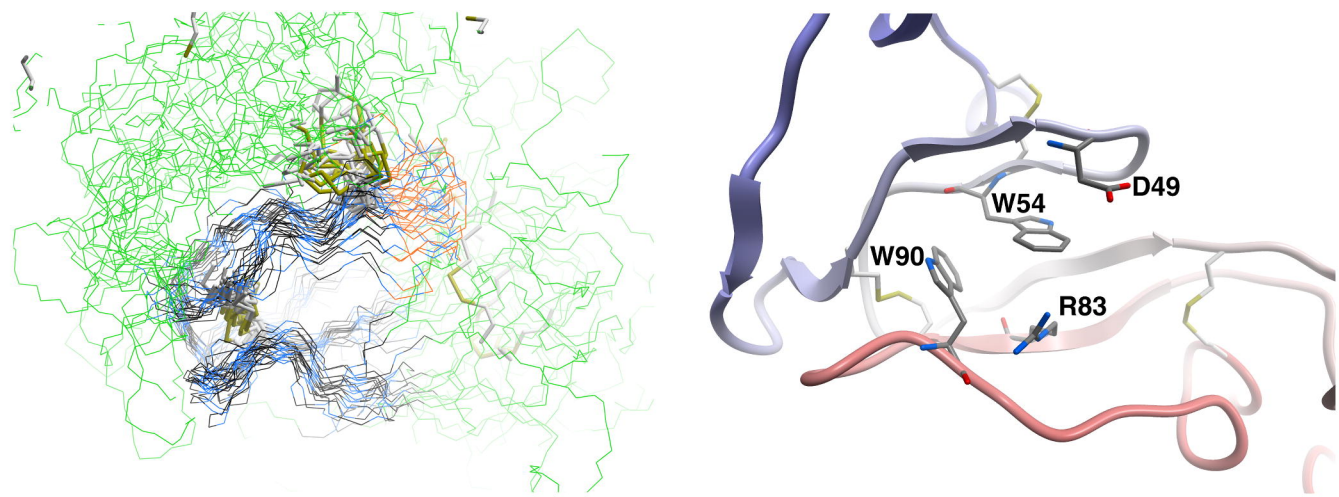

Figure 9

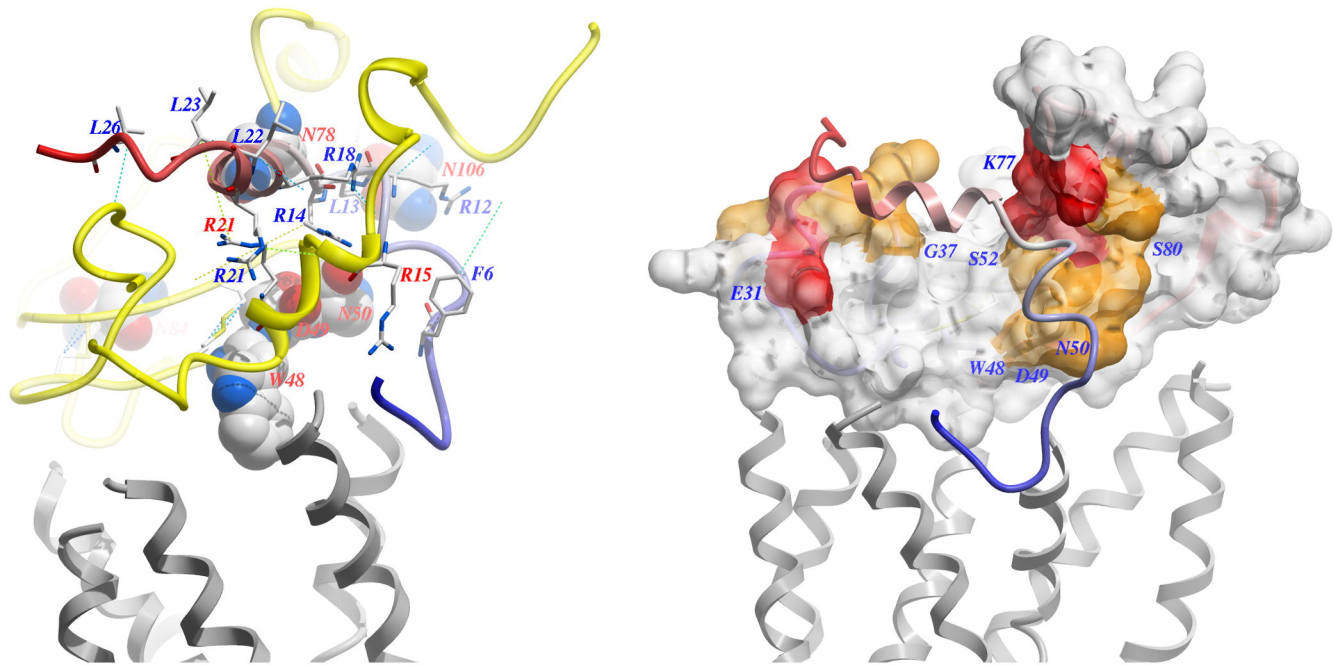
L23
L26
L22
N78
R18
N106
L13
R12
R21
R14
R21
R15
F6
N50
D49
W48
K77
G37
S52
S80
E31
N50
W48
D49

Figure 10