J Comput Aided Mol Des (2012) 26:835–845
DOI 10.1007/s10822-012-9574-x

Predicting the effects of amino acid replacements in peptide hormones on their binding affinities for class B GPCRs and application to the design of secretin receptor antagonists

Jerez A. Te · Maoqing Dong · Laurence J. Miller · Andrew J. Bordner

Received: 20 December 2011 / Accepted: 16 April 2012 / Published online: 11 May 2012

Abstract Computational prediction of the effects of residue changes on peptide-protein binding affinities, followed by experimental testing of the top predicted binders, is an efficient strategy for the rational structure-based design of peptide inhibitors. In this study we apply this approach to the discovery of competitive antagonists for the secretin receptor, the prototypical member of class B G protein-coupled receptors (GPCRs). Proteins in this family are involved in peptide hormone-stimulated signaling and are implicated in several human diseases, making them potential therapeutic targets. We first validated our computational method by predicting changes in the binding affinities of several peptides to their cognate class B GPCRs due to alanine replacement and compared the results with previously published experimental values. Overall, the results showed a significant correlation between the predicted and experimental $\Delta\Delta G$ values. Next, we identified candidate inhibitors by applying this method to a homology model of the secretin receptor bound to an N-terminal truncated secretin peptide. Predictions were made for single residue replacements to each of the other nineteen naturally occurring amino acids at peptide residues within the segment binding the receptor N-terminal domain. Amino acid replacements predicted to most enhance receptor binding were then experimentally tested by competition-binding assays. We found two residue changes that improved binding affinities by almost one log unit. Furthermore, a peptide combining both of these favorable modifications resulted in an almost two log unit improvement in binding affinity, demonstrating the approximately additive effect of these changes on binding. In order to further investigate possible physical effects of these residue changes on receptor binding affinity, molecular dynamics simulations were performed on representatives of the successful peptide analogues (namely A17I, G25R, and A17I/G25R) in bound and unbound forms. These simulations suggested that a combination of the α-helical propensity of the unbound peptide and specific interactions between the peptide and the receptor extracellular domain contribute to their higher binding affinities.

Keywords G protein-coupled receptors · Receptor antagonists · Structure-based design · Peptide modifications · Molecular dynamics simulations · Binding affinity predictions · Homology model

Introduction

Proteins in the G protein-coupled receptor (GPCR) superfamily, which are characterized by a 7-helix transmembrane domain, transfer signals across cell membranes and are the targets of $\sim$30 % of all currently marketed drugs [1]. A subgroup of the GPCR superfamily, class B GPCRs, bind peptide hormones, regulate physiological functions such as exocrine and endocrine secretion and neuronal modulation, and are implicated in multiple human diseases including diabetes, osteoporosis, and psychiatric disorders [2].

Class B GPCRs are distinguished by an N-terminal extracellular domain (ECD), comprised of $\sim$100–160 residues, which plays an important role in binding peptide hormones. Structural and functional studies support a 'two-domain model' of peptide binding and activation, in which

Electronic supplementary material The online version of this article (doi:10.1007/s10822-012-9574-x) contains supplementary material, which is available to authorized users.

J. A. Te · M. Dong · L. J. Miller · A. J. Bordner (✉)
Mayo Clinic, 13400 East Shea Boulevard, Scottsdale, AZ 85259, USA
e-mail: bordner.andrew@mayo.edu

the C-terminal segment of the peptide binds to the receptor ECD, while the N-terminal segment of the peptide interacts with the extracellular surface of the receptor transmembrane domain leading to activation [2, 3]. Generally, N-terminal truncations of the peptides generate competitive antagonists, while C-terminal truncations remain active but with a significant decrease in affinity and potency.

While the structure for the full-length class B GPCR remains undetermined, several X-ray and NMR structures of the receptor ECD with bound peptide were published recently [4–8], and contribute substantially to our understanding of how these receptors bind their ligands. The ECDs of class B GPCRs share common structural features, the 'secretin family recognition fold' (or 'common fold' for brevity), having a core with two anti-parallel β-sheets connected by three disulfide bonds. One of the disulfide bonds connects the α-helix (if present) to the first β-sheet strand. The recognition fold in these structures seems to suggest a common mode of interaction, with peptide binding in a predominantly α-helical conformation and in similar orientations (although the position can deviate by 5–7 Å) sandwiched between the two β-sheets [3].

Experimental alanine scanning of the peptides [9–13] and receptor mutagenesis data are helpful in determining and understanding which residues in the interface of the complex are energetically important for the interaction. Modification of the peptide might also enhance the binding affinity of the peptide for the receptor, which could be useful in developing inhibitors for studying the inactive conformation or as drugs for some members of class B GPCRs. However, extensive peptide synthesis requires significant time and cost. Computational methods that predict the effects of peptide modification on receptor binding affinity can facilitate a more efficient structure-based rational design of such peptide inhibitors.

A number of computational methods that can predict the changes in binding free energy ($\Delta\Delta G$) caused by residue modifications have been developed. These methods vary in the force fields used (e.g. with physics-based potentials or knowledge-based potentials or a hybrid of both) and the scoring function. One of the popular computational tools for predicting binding free energy change due to a modification is the computational alanine scanning protocol [14, 15] in the Rosetta protein modeling program. This program calculates the change in protein–protein binding free energy due to an alanine replacement in the interface. A similar method developed by Bordner and Abagyan [16] used Biased-probability Monte Carlo sampling [17] with the Internal Coordinate Mechanics (ICM) program to predict the effects of single residue changes on protein stability. This approach was developed by fitting a large set of stability data for such residue replacements in monomeric proteins. Thus the method using ICM allows for arbitrary modifications by employing global optimization of a physical energy function to predict the structural effects of the changes followed by scoring with a physics-based energy function.

Although the ICM protocol's application to protein–protein interactions has not yet been tested, a previous study showed that folding of monomeric and molecular interactions of multimeric proteins are governed by the similar chemical and physical forces [18], thus the protocol developed by Bordner and Abagyan [16] can potentially be used for predicting the effects of residue modifications on protein–peptide interactions. Here, ICM was used to predict the effect of single residue changes on the stability of protein–peptide interactions as applied to the binding of peptide hormones to class B GPCRs. In particular, the secretin (Sec) peptide was used as a representative peptide where residues were computationally modified to predict the effect of single residue replacements on its binding affinity for the secretin receptor (SecR), a prototypical member of class B GPCRs. The choice of the test case extends our previous study to enhance the binding affinity of the Sec/SecR complex using lactam bridges [19].

It should be noted that predicting the changes in binding affinity due to a single residue replacement is challenging and is an active area of research. The method of Bordner and Abagyan [16] have a prediction uncertainty of $\sim$1 kcal/mol, which is more than half an order of magnitude difference in binding affinity. When applied to class B GPCRs, a number of other factors could potentially affect the prediction of the changes in binding affinity. Among them, the transmembrane domain, which is not included in the calculations, could interact with the peptide and thereby change the thermodynamics of the interaction. The helicity of the peptide is correlated to the binding affinity, and some modifications may likely affect the peptide's helicity in solution. In addition, a residue replacement could potentially cause a change in the peptide binding site or bound conformation, which is not accounted for in the prediction methods. Another challenge is the use of a homology model in this study to predict these residue changes as there is no available X-ray crystal or NMR structure of secretin bound to the SecR ECD.

In this study, we first showed that the ICM method is capable of predicting changes in the peptide–protein binding free energy and analyzed other effects of the residue changes not captured by the modeling procedure. Specifically, the computational alanine scanning of three crystal structure and two homology models of peptides for class B GPCRs were performed using the ICM method and compared to experimental data and predictions using the Rosetta computational alanine scanning method. Next, we applied the ICM method to the computational design of Sec competitive antagonists by predicting single residue

replacements that are expected to improve Sec–SecR binding affinity and then experimentally tested the top predicted binders. Lastly, we used molecular dynamics (MD) simulations to infer the importance of selected modifications on the average helicity of the unbound peptide and the effects on the interactions between the secretin analogues and the receptor ECD brought about by these residue changes.

Methods

Experimental structures and homology modeling

The peptide/ECD complex structures used here were either downloaded from the Protein Data Bank (http://www.rcsb.org/) [20] or generated from homology modeling. The X-ray crystal structures for glucagon-like peptide 1 with its receptor ECD (GLP1/GLP1R, PDB: 3IOL [4]), the parathyroid hormone with its receptor ECD (PTH/PTH1R, PDB: 3C4M [5]), and glucose-dependent insulinotropic peptide with its receptor ECD (GIP/GIPR, PDB: 2QKH [6]) were obtained from the Protein Data Bank. The ICM program (version 3.6, Molsoft LLC) [21] was used to add the missing hydrogen atoms to the crystal structures and to assign ideal covalent geometry to the residues. A homology model of the peptide hormones Sec and vasoactive intestinal polypeptide (VIP) docked with the N-terminal domain of their cognate receptors (SecR and VPAC2, respectively) were built based on the X-ray crystal structure of the GLP1/GLP1R complex. The ICM program was used for the homology modeling, wherein the disulfide bonds and the backbone of the β-sheets of the ECD were aligned and tethered to the corresponding atoms in GLP1R structure using a quadratic restraint potential, while the sequence of the peptide was aligned and tethered to residues 11–30 of GLP1. The peptide/ECD structure was annealed to the tethers and the structure was refined using energy optimization with Biased-Probability Monte Carlo [17] simulations for 5×10^7 function calls.

Computational ΔΔG predictions

ICM was used to predict the change in peptide–receptor binding free energy (ΔΔG) due to single residue changes in the peptide. Biased-Probability Monte Carlo (BPMC) [17] simulations at T = 700 K were used to optimize the side-chain rotamer angles of each mutated residue (and its corresponding residue in the wild-type peptide) and all contacting residues with side-chain non-hydrogen atoms within 4.0 Å of any side-chain non-hydrogen atoms in the modified residue. The backbone was fixed throughout the simulation and each simulation had 2×10^5 function calls with up to 2,000 steps of conjugate gradient minimization for each accepted Monte Carlo move. An energy function with electrostatic, van der Waals, hydrogen bonding and torsional terms from ECEPP/3 force field [22–24], as well as solvation [25] and side chain entropy [17] terms were used. The lowest energy conformation from each simulation was used to determine the ΔΔG for each residue replaced with an alanine using a protocol described previously [16]. An ICM macro for performing the ΔΔG calculations is included as a Supplementary file. The ΔΔG calculations were performed for all crystal structures and homology models. In addition to the determination of computed ΔΔG values from ICM, the Robetta server (http://robetta.bakerlab.org/alascansubmit.jsp) for computational alanine scanning was also used [14, 15]. The calculated ΔΔG was compared to the experimental ΔΔG obtained from the inhibition constants (K_i) in alanine scanning experiments [9–12]. The experimental $\Delta\Delta G = RT \ln (K_i^{mutant}/K_i^{wild\text{-}type})$, where R is the ideal gas constant and T is the temperature.

Moreover, in order to account for flexibility, ΔΔG predictions were also performed using structures extracted every 2 ns of the last 10 ns of a 30-ns molecular dynamics simulation for Sec/SecR (discussed below). For the Sec/SecR system, each secretin amino acid in positions 16 through 27 (located within the proposed ligand-binding cleft in the ECD) were replaced in the calculation with each of the other nineteen naturally-occurring amino acids. Only replacements in positions 17, 20, 23, 24, and 25 were predicted to result in higher binding energy. These included replacement of Ala in position 17 with Ile, replacement of Gln in position 20 with Glu, Ile, Leu, Met, and Tyr, replacement of Leu in position 23 with Tyr, replacement of Gln in position 24 with Ile and Thr, and replacement of Gly in position 25 with Asp, Ile, Tyr, Arg, and Trp. The predicted ΔΔG values for these peptide modifications are shown in Table 1.

Experimental verification of predicted peptide residue replacements

Peptides

Peptides representing a truncated form of rat secretin eliminating the first four residues at the amino terminus (extending from residue 5 to residue 27 (Sec(5–27)) and 15 analogues of this peptide incorporating amino acid changes in positions 17, 20, 23, 24 and 25 that were predicted to enhance binding affinity were prepared (Table 1). All peptides were synthesized and purified to greater than 98 % purity by GenScript (Piscataway, NJ) and Selleck Chemicals (Houston, TX). Expected molecular masses were verified by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry.

Table 1 Amino acid changes predicted to most improve the binding affinity (ΔΔG) of the truncated secretin peptide (residues 5–27)

Residue	Mutation	Simulation 1		Simulation 2	
		ΔΔG (kcal/mol)	z-score	ΔΔG (kcal/mol)	z-score
Ala^{17}	Ile	−1.31	2.77	−1.44	3.45
Gln^{20}	Glu	−0.71	2.10	−2.36	10.68
Gln^{20}	Ile	−1.17	6.11	−2.62	3.04
Gln^{20}	Leu	−0.65	2.84	−2.54	7.39
Gln^{20}	Met	−0.76	2.64	−2.89	7.56
Gln^{20}	Tyr	0.04	0.09	−0.98	2.15
Leu^{23}	Tyr	−1.48	2.71	−0.52	0.72
Gln^{24}	Ile	−1.16	2.25	−0.81	2.12
Gln^{24}	Thr	−1.07	2.23	−0.58	1.10
Gly^{25}	Asp	−1.08	0.66	−0.58	3.56
Gly^{25}	Ile	−2.47	1.67	−1.95	4.49
Gly^{25}	Arg	−2.72	2.64	−2.41	3.22
Gly^{25}	Trp	−2.09	1.30	−1.37	0.99
Gly^{25}	Tyr	−3.29	1.93	−2.76	2.73

The ΔΔG values shown are average predictions using the ICM method based on six conformations for each of the two independent MD simulations

The [Tyr^{10}]rat secretin-27 analogue that was used as a radioligand in ligand binding assays was synthesized in our laboratory and radioiodinated oxidatively using 1 mCi $Na^{125}I$ (PerkinElmer Inc., Waltham, MA) and exposure for 15 s to the solid phase oxidant, *N*-chlorobenzenesulfonamide (Iodination beads) (Pierce, Rockford, IL), as we described previously [26]. The radioiodinated peptide was purified by reversed-phase HPLC to yield specific radioactivity of 2,000 Ci/mmol [26].

Ligand binding

The secretin analogues listed in Table 1 were examined for their ability to compete for binding of the radioligand, [Tyr^{10}]rat secretin-27, to the Chinese hamster ovary cell line stably expressing the wild type rat secretin receptor (CHO–SecR) [26] using whole cell binding in 24-well tissue culture plates. For this, approximately 50,000 cells were plated into each well and were grown for approximately 72 h prior to the assay. Cells were washed twice with Krebs-Ringers/HEPES (KRH) medium (25 mM HEPES (pH 7.4), 104 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM KH_2PO_4, and 1.2 mM $MgSO_4$) containing 0.01 % soybean trypsin inhibitor and 0.2 % bovine serum albumin, and were then incubated with a constant amount of radioligand, ^{125}I–[Tyr^{10}]rat secretin(1–27) (5 pM, approximately 20,000 rpm), and increasing concentrations of secretin analogues (ranging from 0 to 1 μM) for 60 min at room temperature. Cells were then washed twice with ice-cold KRH medium containing 0.01 % soybean trypsin inhibitor and 0.2 % bovine serum albumin to separate free from cell-bound radioligand. The plate-bound cells were lysed with 0.5 M NaOH and radioactivity was quantified with a γ-spectrometer. Nonspecific binding was determined in the presence of 0.1 μM secretin and represented less than 10 % of total radioligand bound. Competition-binding curves were analyzed and plotted using the non-linear regression analysis program in the Prism software suite v3.0. Binding kinetics were determined by analysis with the LIGAND program of Munson and Rodbard [27].

Molecular dynamics simulations

Molecular dynamics (MD) simulations were used to equilibrate the homology models, to gain insights into the peptide–receptor complexes, and to determine the helicities of the unbound peptides. The GROMACS software package (version 4.5.3) [28] with Amber ff99sb-ildn all-atom parameters [29] was used to perform the simulations. The peptide was solvated in pre-equilibrated TIP3P water model [30] and the appropriate number of ions were added to make the system neutral. The system was equilibrated under a canonical (*NVT*) ensemble for 200 ps using the Berendsen thermostat [31] with a relaxation time of 0.1 ps for temperature control. The system was further equilibrated in an isothermal–isobaric (*NPT*) ensemble for another 500 ps using the Parrinello–Rahman barostat [32] with a relaxation time of 2.0 ps. The reference temperature and pressure for the system were 300 K and 0.1 MPa, respectively. Position restraints (with spring constant of 1,000 kJ/mol nm^2) were applied to all non-hydrogen atoms of the protein complex during the equilibration. All simulations had 2-fs timesteps with the linear constraint solver method [33] used to constrain all bond lengths. The Lennard-Jones and electrostatic interactions were cut-off at 10 Å and the neighbor list was updated every 10 fs. Long-range electrostatic interactions were treated using the particle mesh Ewald method [34] with fourth-order spline

interpolation and 1.6 Å grid spacing. Following equilibration, *NPT* ensemble production runs were performed. Two independent simulations, each lasting 30 ns, were performed for the Sec/SecR complex. Six conformations (every 2 ns from t = 20 ns to t = 30 ns) from each independent MD simulation were extracted and used in the ΔΔG calculation described above. The averaging was performed since multiple conformations were likely to improve the prediction of ΔΔG by accounting for structural flexibility.

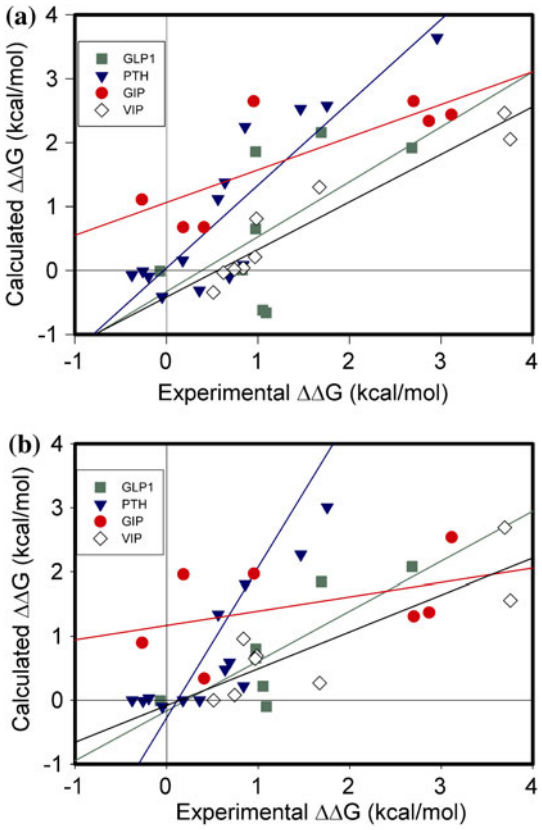
The first simulation was extended to 100 ns (MD1). Secretin analogues in complex with its receptor was obtained by extracting the coordinates of MD1 at t = 60 ns and introducing mutations to particular residues. Another round of equilibration (700 ps) was followed by another 40-ns simulation run for each complex. For the unbound peptides, 100-ns MD simulations were performed. Coordinates from the simulation were saved every 20 ps for analysis of the last 50 ns of the simulation for the unbound peptide and the last 20 ns of the peptide–ECD complex. The analysis was performed using the built-in tools in GROMACS.

Results and discussion

Alanine mutations

The accuracy of the ICM computational method [16] applied to peptide–GPCR complexes was assessed by comparisons between predicted and experimental binding free energy (ΔΔG) values. Prediction were made for alanine replacements of residues 23–34 of GLP1, residues 18–31 of PTH and certain residues of GIP (Fig. 1a), all of which have crystal structures in complex with their respective receptor ECDs [4–6] and also available experimental alanine scanning data [6, 11, 12]. These portions of the peptide hormones are located within the proposed ligand-binding cleft in the ECD. In GLP1/GLP1R, residues 24, 25, and 30 are alanine and so were not included in the ΔΔG calculation. The ΔΔG values predicted from the computational alanine scanning were correlated with the available experimental alanine scanning data [6, 11, 12]. The coefficient of determination (R^2) for GLP1/GLP1R, PTH/PTH1R, and GIP/GIPR were 0.60, 0.77, and 0.57, respectively.

The Robetta alanine scanning server was used to verify the performance of ICM (Fig. 1b). For the peptide–ECD with existing crystal structures, the calculated ΔΔG from the Robetta server for GLP1/GLP1R and PTH/PTH1R were in good agreement with experimental alanine scanning data (R^2 = 0.84 and 0.83, respectively), although the correlation for GIP/GIPR was poor (R^2 = 0.19). Although the ICM protocol performed better for GIP/GIPR, with the given number of available alanine scanning experiments and crystal structures, it would be hard to conclude which method performs better. These results do indicate, however, that even though the ICM prediction method was parameterized from fitting mutation data of monomeric proteins, it is likely applicable to protein–protein interactions and its performance on alanine mutation is comparable to that of the Robetta server.



(c) Coefficient of determination (R^2) from the ΔΔG prediction

Complex	ICM	Rosetta
GLP1/GLP1R	0.60	0.84
PTH/PTH1R	0.77	0.83
GIP/GIPR	0.57	0.19
VIP/VPAC2	0.91	0.58

Fig. 1 Alanine scanning of peptides bound to the ECD of class B GPCRS. *Plots* of calculated versus experimental ΔΔG values (kcal/mol) with calculations performed using **a** ICM or **b** Rosetta for the replacements of non-alanine residues 23–34 of GLP1, 18–31 of PTH, and 20–28 of VIP with alanine. Residues of GIP (19, 20, 22, 23, 26, 27, 30) with available experimental data were also modified. The coefficient of determination (R^2) statistics between the calculated and experimental ΔΔG values for the different complexes are given in **c**. In **b** the point for PTH R20A with a calculated ΔΔG of 8.66 kcal/mol versus the experimental ΔΔG of 2.96 kcal/mol is not shown

While some peptide–ECD complexes have existing crystal or NMR structures, the structures for other complexes are yet to be determined. Here, homology models based on the common interaction mode for peptides binding to class B GPCRs were created. In the modeling procedure, the conserved disulfide bonds and the backbones of the two anti-parallel β-sheets were tethered while allowing the loops and the α-helix to sample the conformational space through a Biased-Probability Monte Carlo simulation in ICM. This is a justified assumption considering that the β-sheets and the disulfide bonds are part of the structurally conserved region while the loop conformations are observed to vary in structure superpositions of different ECD structures within the protein family [3].

Previously, the computational protocol in ICM was used to predict the change in binding free energy for residues 15–27 (except residue 17, which is alanine) in Sec [10]. Notably, the calculations were able to predict the critical residues in Sec, especially the negative effect on binding affinity from alanine replacements of Leu^{19}, Leu^{22}, and Leu^{23} [10], which lie in the binding cleft of the ECD. In addition, the change in binding free energy for residues 20–28 in VIP (Fig. 1a), and certain residues of PACAP/PAC1 (data not shown) that have available experimental alanine scanning data were predicted reasonably accurately. In VIP, the deleterious effects on binding upon the replacement of Tyr^{22} and Leu^{23} with alanine [9] were well predicted. Overall, the $\Delta\Delta G$ prediction using homology models qualitatively agreed with the experimental alanine scanning data [9, 10].

Binding characterization of the Sec(5–27) analogues

Antagonists are desirable drugs for several class B GPCRs and may be useful for understanding inactive conformations of this family of receptors. It is recognized that amino-terminal truncation of natural peptide ligands for this GPCR class results in competitive antagonists [35–37]. Indeed, truncation of the first four amino-terminal residues of secretin eliminates its endogenous agonist activity, but because the secretin amino terminus also contributes important receptor-binding determinants, the resulting peptide (Sec(5–27)) has a binding affinity more than two orders of magnitude lower than that of the natural ligand [36, 38–40].

To improve the binding affinity between Sec(5–27) and SecR, peptide modifications that would stabilize the peptide–protein interactions were predicted using the ICM method. Although high resolution structures of the receptor ECD suggest similarity in structural motifs and ligand binding modes among several members, there are inconsistencies in the absolute sites of ligand binding and in the positioning (5–7 Å deviation) of the ligands in these structures [3], suggesting some variation in binding interactions across this family. Furthermore, alternative binding conformations have been proposed for docking the flexible amino-terminal regions of ligands to these receptors and it was, indeed, due to such differences that we used MD simulations to refine the receptor–peptide interaction of Sec/SecR homology model. Subsequently, six atomic coordinates were extracted from the last 10 ns of each MD simulation, and the change in binding free energy was predicted using ICM and compared to the experimental alanine scanning data (Fig. 2a). As with the original homology model (before MD simulation), the negative effect on binding affinity from alanine replacements of

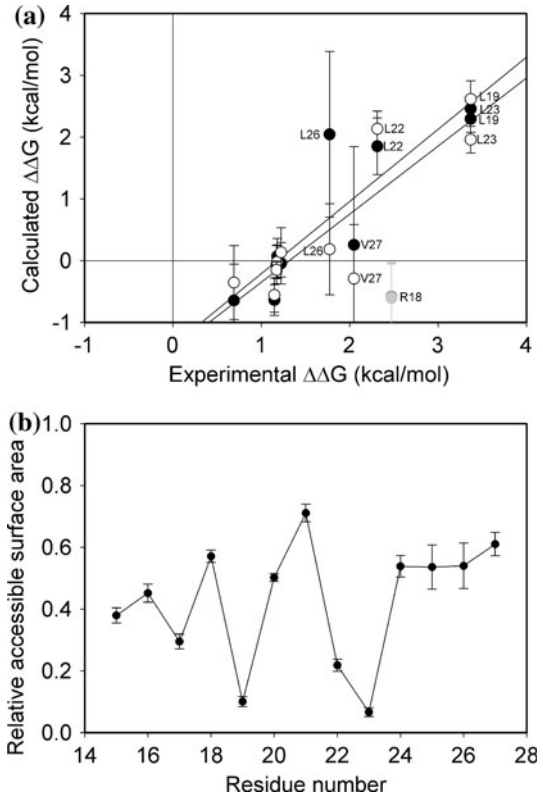


Fig. 2 Computational alanine scanning and solvent-accessible surface area of the Sec/SecR complex. **a** Shown are the averages $\Delta\Delta G$ (kcal/mol) calculated by ICM for Sec alanine replacements using 6 snapshots of the molecular dynamics simulations (every 2 ns from t = 20 to t = 30 ns of two-independent MD simulations represented by *open* and *filled circles*, respectively) for residues 16–27 of Sec/SecR. The coefficients of determination (R^2) for Sec/SecR excluding Arg^{18}, a false negative in the calculation, for these two simulations were 0.78 and 0.77. **b** Shown are the average relative solvent-accessible surface areas for each residue in the peptide. Snapshots taken every 100 ps from t = 20 to t = 30 ns for the same MD simulations as above were used to calculate the SASA values. The relative SASA was determined by dividing the SASA for each peptide residue by the maximum SASA for that residue type, as calculated for the residue in an extended conformation between two glycines. A smaller relative SASA value was associated with a more buried peptide residue in the simulation of the peptide–receptor complex

Leu[19], Leu[22], and Leu[23] were predicted with good qualitative agreement. According to the MD simulations, these three residues were also shown to have lower relative solvent-accessible surface area (SASA) than other residues (Fig. 2b). The changes in Leu[26] and Val[27], which are on the tail end of the C-terminus and were rather solvent-exposed, exhibited the greatest deviation in calculated $\Delta\Delta G$ for the two independent MD simulations. Arg[18], which also has a large solvent-accessible surface area, gave a false-negative result. It is interesting that Arg[18] is three residues away from Asp[15], and the two residues might form a salt bridge that can stabilize the α-helical conformation of the peptide, as discussed below.

While the computational alanine scanning showed reasonable agreement with experimental data, residue replacements with alanine are more predictable in terms of computation [41]; however, the challenge is to predict non-alanine changes, particularly those replacing small residues with large residues. Thus, using the Sec/SecR system, we tried to predict favorable single-point modifications that might improve the binding of Sec(5–27) to SecR by replacing residues 16–27 of Sec with each of the 19 other natural amino acids. Unfortunately, the $\Delta\Delta G$ predictions did not identify mutations for Leu[19], Leu[22], and Leu[23] that might improve binding (data not shown), except for one modification where Leu[23] was replaced with Tyr ([Y[23]]sec(5–27)). Experimentally, however, the ([Y[23]]sec(5–27) analogue significantly decreased binding (Table 2). Although we could not eliminate the uncertainty of the ICM protocol to give false-negative or false-positive predictions, it is possible that certain critical amino acids in the peptide are optimized for binding to the receptor. Previous efforts to modify Leu[19] (which like Leu[23] is assumed to be buried in the binding cleft) by replacing it with a smaller or larger hydrophobic residue were unable to yield peptides that bound better than or comparable to the natural peptide ligand. The replacement of Leu[19] with alanine decreased the binding affinity by >2 orders of magnitude [10]. Modifying residue 19 to Trp also significantly decreased the binding affinity (unpublished results).

Focusing on peptide residues that have contacts with residues in the receptor, we performed 13 additional modifications that have predicted $\Delta\Delta G \leq -1.0$ kcal/mol and z-score > 2 in at least one of the two independent predictions. In this work, peptide analogues of Sec(5–27) incorporating single amino acid changes that were predicted to improve binding affinity were characterized. Their secretin receptor binding affinities are shown in Table 2. Of note, most of the peptide analogues predicted to have improved binding affinities were found to bind with affinities at least as high as that of the parental truncated peptide. In addition to [Y[23]]sec(5–27), [E[20]]sec(5–27) was found to have significantly lower binding affinity than the parental peptide.

It should be noted that the ICM binding affinity prediction is based on a conservative sampling procedure, in which the conformation of the modified residue is optimized within a fixed backbone environment. Methods

Table 2 Primary structures and binding characteristics of Sec(5–27) and its analogues

Peptide name	Peptide sequences (5 10 15 20 25)	Binding affinity IC_{50} (nM)
Sec(5-27)	TFTSELSRLQDSARLQRLLQGLV	890 ± 86
[I[17]]sec(5-27)	TFTSELSRLQDS**I**RLQRLLQGLV	127 ± 15**
[E[20]]sec(5-27)	TFTSELSRLQDSARL**E**RLLQGLV	>1000
[I[20]]sec(5-27)	TFTSELSRLQDSARL**I**RLLQGLV	600 ± 80
[L[20]]sec(5-27)	TFTSELSRLQDSARL**L**RLLQGLV	687 ± 47
[M[20]]sec(5-27)	TFTSELSRLQDSARL**M**RLLQGLV	570 ± 35*
[Y[20]]sec(5-27)	TFTSELSRLQDSARL**Y**RLLQGLV	440 ± 83*
[Y[23]]sec(5-27)	TFTSELSRLQDSARLQRL**Y**QGLV	>1000
[I[24]]sec(5-27)	TFTSELSRLQDSARLQRLL**I**GLV	727 ± 37
[T[24]]sec(5-27)	TFTSELSRLQDSARLQRLL**T**GLV	793 ± 37
[D[25]]sec(5-27)	TFTSELSRLQDSARLQRLLQ**D**LV	363 ± 67**
[I[25]]sec(5-27)	TFTSELSRLQDSARLQRLLQ**I**LV	290 ± 52**
[Y[25]]sec(5-27)	TFTSELSRLQDSARLQRLLQ**Y**LV	660 ± 117
[R[25]]sec(5-27)	TFTSELSRLQDSARLQRLLQ**R**LV	137 ± 17**
[W[25]]sec(5-27)	TFTSELSRLQDSARLQRLLQ**W**LV	227 ± 33**
[I[17],R[25]]sec(5-27)	TFTSELSRLQDS**I**RLQRLLQ**R**LV	23 ± 1**

Shown are the amino acid sequences of amino-terminally truncated rat secretin (Sec(5–27)) and its analogues. Natural residues in the analogues are colored gray, while modified residues are colored black. Shown also are the IC_{50} values expressed as means ± SEM. Two-tailed P value tests were performed to determine the significance of differences using InStat3 (GraphPad Software, San Diego, CA). * and ** indicate values that are significantly different from that of sec(5–27) ($P < 0.05$ and $P < 0.01$, respectively)

including backbone sampling and relaxation [42, 43] might improve predictions of ΔΔG, although this can potentially introduce additional error and so perform worse than the current ICM procedure if structural changes are negligible [44]. Furthermore, the use of a homology model to predict ΔΔG introduces additional uncertainty in the prediction results.

However, seven of the analogues, involving positions 17, 20, and 25, had improved binding affinities relative to that of the parental truncated peptide. Of these, [I^{17}]sec(5–27) and [R^{25}]sec(5–27) had the highest affinities, with each increasing binding affinity by almost an order of magnitude relative to that of Sec(5–27).

The Sec(5–27) analogue that combined both of these amino acid changes, isoleucine in position 17 and arginine in position 25, was also prepared and studied. Figure 3 shows that the effects of each change were additive, with the peptide incorporating both changes, [I^{17},R^{25}]sec(5–27), significantly improving the binding affinity. The affinity of this peptide was approximately an order of magnitude higher than that of the peptides incorporating only one of these amino acid changes and approximately two orders of magnitude higher than that of Sec(5–27).

Molecular dynamics simulations of peptides

Sec/SecR complex

MD simulations of the strongest binding secretin analogues in complex with SecR were performed in order to gain further insight into specific interactions that may explain their high binding affinities. In particular, 40-ns simulations of SecR with three bound secretin analogues ([I^{17}]sec(5–27), [R^{25}]sec(5–27), and [I^{17},R^{25}]sec(5–27)) were performed (Fig. 4a). The root mean square fluctuations

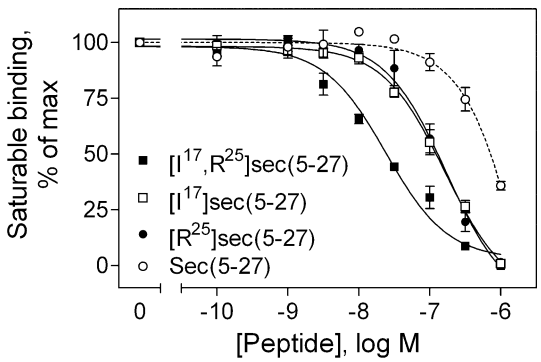


Fig. 3 Binding characterization of Sec(5–27) analogues. Shown are *curves* of increasing concentrations of Sec(5–27), [I^{17}]sec(5–27), [R^{25}]sec(5–27) and [I^{17},R^{25}]sec(5–27) to compete for binding of the radioligand, ^{125}I–[Tyr^{10}]rat secretin(1–27), to CHO–SecR cells. Values represent percentages of saturable binding, expressed as the means ± SEM of duplicate values from a minimum of three independent experiments

(RMSF) per residue of the peptides during the last 20 ns of each trajectory were analyzed. Interestingly, the secretin analogues [R^{25}]sec(5–27) and [I^{17},R^{25}]sec(5–27), but not [I^{17}]sec(5–27), had smaller RMSF, which might indicate stability, on residues that were shown to be critical in alanine scanning experiment such as Leu^{19}, Leu^{22}, and Leu^{23} compared to Sec(5–27) (Fig. 4b). In addition, the analyses suggest that the strongest binder, [I^{17},R^{25}]sec(5–27), has the largest increase in surface area, calculated as $\Delta ASA = ASA_{peptide} + ASA_{receptor} - ASA_{complex}$, buried in the interaction between the peptide and receptor (Table 3). Additionally, this peptide has more favorable van der Waals interactions between residues 15 and 27 of the peptide and residues of the receptor that are within 5 Å of the specified peptide region. Because the nature of the interaction between Sec and the receptor is largely hydrophobic, van der Waals interactions and tight packing at the interface might contribute more to binding affinity. [R^{25}]sec(5–27) also displayed larger ΔASA and more favorable van der Waals energy than Sec(5–27) (Table 3). Interestingly, the interface surface area for [I^{17}]sec(5–27) was slightly lower than Sec(5–27), and the former only has a slightly more favorable van der Waals interaction than the latter.

Unbound secretin in solution

Properties of the unbound secretin analogues can also influence their binding affinities to the secretin receptor. In particular, peptide ligands for other members of the protein family (except for calcitonin, whose helical structure is stabilized by a disulfide bond) are known to undergo a transition from a generally disordered conformation in solution to a more ordered α-helical conformation upon binding their cognate receptors [3]. This is expected to oppose binding due to the lower conformational entropy of the bound peptide. In contrast, modifications that stabilize the helical conformation of the free peptide in solution may reduce this entropic penalty and thereby increase binding affinity. In order to investigate this possible contribution to binding affinity that is not accounted for in the ΔΔG prediction method, simulations of the unbound secretin analogues as well as full-length and truncated wild-type secretin as controls were performed (100 ns each). Analyses from the last 50 ns of the simulation of the unbound peptides revealed some interesting differences. Sec(1–27) was on average more α-helical than the truncated Sec(5–27) (Fig. 5), suggesting that the truncation of residues 1–4 in the peptide might reduce the helicity of the unbound peptide and, therefore, contribute to its lower binding affinity. In comparison to full-length secretin, where the average number of residues involved in a helical structure (α-helix) at residues 15–27 C-termini was 7.26

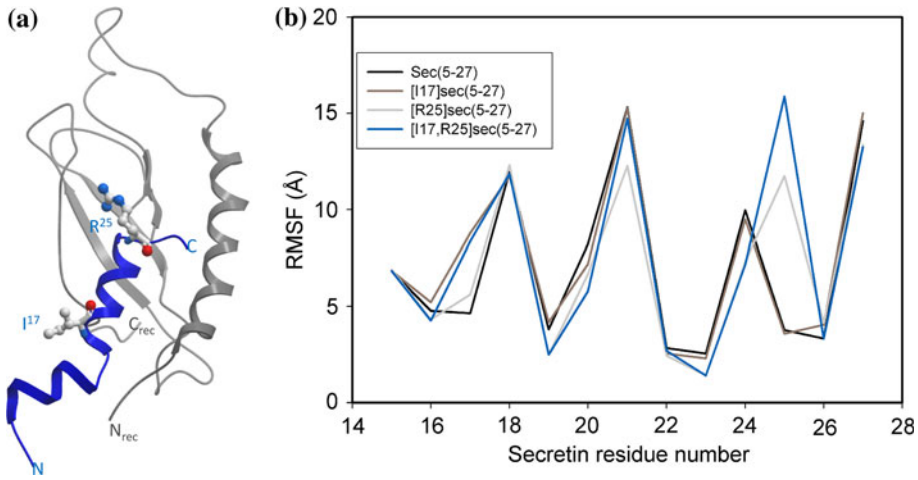


Fig. 4 Molecular dynamics simulation of the Sec/SecR complex. **a** Shown is a representative conformation of [I^{17},R^{25}]sec(5–27) (*blue*) bound to the SecR ECD (*gray*) at t = 40 ns. **b** Also shown is the root mean square fluctuation of the secretin residues during the last 20 ns of the simulation

Table 3 Properties of the secretin analogues in complex and unbound conformations

	Sec(5–27)	[I^{17}]sec(5–27)	[R^{25}]sec(5–27)	[I^{17},R^{25}]sec(5–27)
Hydrophobic ΔASA (Å^2)	230	211	272	289
Hydrophilic ΔASA (Å^2)	738	705	795	830
van der Waals energy[a]	−31.6	−33.7	−44.9	−43.7
Electrostatic energy[a]	−1.0	−0.8	−0.7	−1.1
α-helix[b]	1.73	7.32	1.76	8.54
3_{10}-helix[b]	1.60	0.11	1.43	0.29

Shown are the change solvent accessible surface area (in Å^2) and the van der Waals and electrostatic energy components (in kcal/mol) for the secretin analogue in complex with the ECD of the interaction. Also shown is the average number of residues involved in particular secondary structure for the unbound analogues in solution

[a] Energy components were determined by ICM based on the interaction between residues 15 and 27 in the peptide and its neighboring residues in the receptor (within 5 Å)

[b] α-helix and 3_{10}-helix content were measured using the do_dssp program [46] implemented in GROMACS and reflects the average number of residues within the C-termini (residues 15 and 27) with that particular secondary structure

(no 3_{10}-helical conformations observed), the truncated secretin peptide was less helical, with 1.73 and 1.60 average number of residues involved in α-helical and 3_{10}-helical conformations, respectively. This result is consistent with our previous results obtained using ICM biased-probability Monte Carlo simulations [19], in which the truncated peptide was less helical than the full-length peptide, although the ICM simulations predicted generally higher helicities than the current MD simulations.

The helicity in the C-termini of the Sec peptides (for both Sec(1–27) and Sec(5–27)) was likely due to the salt bridge formed between Arg^{18} and Asp^{15}. An MD simulation of [A^{18}]sec(1–27) was performed to test this hypothesis. A snapshot of the [A^{18}]sec(1–27) mutant at the end of the 100-ns simulation showed a less α-helical C-terminus in the peptide compared to the natural peptide (Fig. 5). The average number of residues involved in a helical structure (α-helix) at residues 15–27 C-termini for the [A^{18}]sec(1–27) were 0 and 0.54, for α- and 3_{10}-helices, respectively. The average Asp^{15}–Arg^{18} separation distance was 2.18 ± 0.43 Å, indicative of a salt bridge, for Sec(1–27); while the Asp^{15}–Ala^{18} separation distance for the modified peptide was a larger 6.70 ± 0.93 Å.

The MD simulations showed that changing alanine in residue 17 to isoleucine may have stabilized the secondary structure of the C-terminus of the peptide and so contributed to improved binding affinity. Simulations of [I^{17}]sec(5–27) and [I^{17},R^{25}]sec(5–27) in solution have on average significantly higher numbers of residues involved in α-helical conformation (Table 3). On the other hand, the average helical content in [R^{25}]sec(5–27) was comparable to Sec(5–27). Is it possible that [I^{17}]sec(5–27) and [R^{25}]sec(5–27) have different effects on the secretin peptide? While we do not eliminate other possibilities for the observed approximate additivity of the binding affinity for [I^{17}]sec(5–27) and [R^{25}]sec(5–27) compared with [I^{17},R^{25}]sec(5–27), it is interesting to speculate that both the unbound form of the peptide and the specific interactions in the peptide–receptor interface could independently contribute to the stability of the complex.

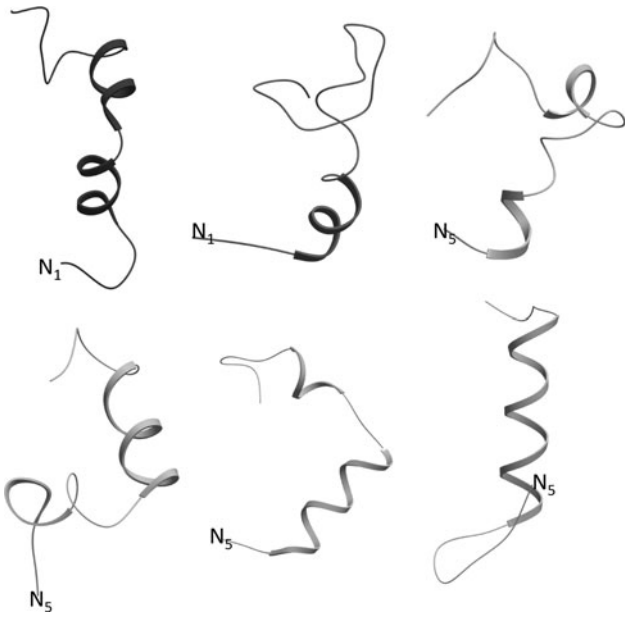


Fig. 5 Snapshot of the unbound peptide after the 100-ns MD simulation. Shown *clockwise* from the *top left* are the conformations of the full-length natural secretin, the full-length R18A secretin variant, truncated natural secretin sec(5–27), and the secretin analogues [I^{17},R^{25}]sec(5–27), [R^{25}]sec(5–27), and [I^{17}]sec(5–27)

In general, these simulations suggest that some modifications in the peptide might affect the helicity of the unbound peptide in solution. Experimentally, peptide hormones are relatively "unstructured," where the helicity of the peptide increases upon binding. Unbound peptides that have some degree of helicity were shown to have higher affinities for the extracellular domain of the receptor [4, 45], most likely because there is less entropic cost for the secondary structure formation. If the findings based on the MD simulations are correct, then the structure of the unbound peptide must be accounted for in order to improve the computational predictions of the binding affinities of peptide hormones for their receptors.

Conclusions

Here, we showed that the ICM method for predicting changes in binding free energy (ΔΔG) was relatively successful when applied to peptides bound to the extracellular domain of class B GPCRs. This method was comparable to Robetta computational alanine scanning in predicting deleterious effects of changing critical residues to alanine for peptides bound to the ECD of several members of class B GPCRs. In addition, the ICM method was applied to designing a competitive antagonist of the secretin receptor by predicting single-point changes in the secretin peptide that have improved binding affinities. Experimental measurements of K_i for the top predicted binders revealed that two of the predicted inhibitors, [I^{17}]sec(5–27) and [R^{25}]sec(5–27), had binding affinities that were almost an order of magnitude higher than that of the original Sec(5–27) peptide. Furthermore, combining these two modifications in [I^{17},R^{25}]sec(5–27) resulted in a peptide with an even greater affinity, approximately two orders of magnitude higher than that of Sec(5–27), demonstrating that these two changes have an approximately additive effect on binding affinity. MD simulations of the secretin analogues both in solution and bound to the secretin receptor suggest that the dominant effects of the two residue changes on SecR binding may be different, with [R^{25}]sec(1–27) stabilizing the interaction with the complex and [I^{17}]sec(1–27) stabilizing the helical conformation of the unbound peptide. These results suggest that a computational method that also predicts the relative α-helix content of the free peptides and then uses these in the calculation of ΔΔG may yield more accurate results for this system. Given the uncertainty introduced by the use of a homology model for the structure of the peptide–receptor complex, the conservative approach used here of predicting single residue modifications may be optimal. However, for peptide–receptor complexes with an available high-resolution X-ray crystal structure consideration of multiple changes at nearby residues, that are not expected to have additive effects on binding, may be a more efficient design strategy. Finally, we note that many other peptide–protein interactions represent potential targets for discovering antagonists for use as molecular probes of these interactions or therapeutics. The computational method presented here should also be applicable to other interactions and can serve as a filter in selecting candidate antagonists for experimental testing.

Acknowledgments The authors thank Mary Lou Augustine and Delia I. Pinon for their excellent technical assistance and the Mayo Clinic Research Computing Services and the Arizona State University Advanced Computing Center for computational resources. This work was supported by the Mayo Clinic and a grant from the National Institutes of Health (DK46577).

References

1. Hopkins AL, Groom CR (2002) The druggable genome. Nat Rev Drug Discov 1(9):727–730
2. Hoare SRJ (2005) Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors. Drug Discov Today 10(6):417–427
3. Parthier C et al (2009) Passing the baton in class B GPCRs: peptide hormone activation via helix induction? Trends Biochem Sci 34(6):303–310
4. Underwood CR et al (2010) Crystal structure of glucagon-like peptide-1 in complex with the extracellular domain of the glucagon-like peptide-1 receptor. J Biol Chem 285(1):723–730

5. Pioszak AA, Xu HE (2008) Molecular recognition of parathyroid hormone by its G protein-coupled receptor. Proc Natl Acad Sci USA 105(13):5034–5039
6. Parthier C et al (2007) Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor. Proc Natl Acad Sci USA 104(35):13942–13947
7. Pioszak AA et al (2008) Molecular recognition of corticotropin-releasing factor by its G-protein-coupled receptor CRFR1. J Biol Chem 283(47):32900–32912
8. Runge S et al (2008) Crystal structure of the ligand-bound glucagon-like peptide-1 receptor extracellular domain. J Biol Chem 283(17):11340–11347
9. Igarashi H et al (2002) Elucidation of the vasoactive intestinal peptide pharmacophore for VPAC2 receptors in human and rat and comparison to the pharmacophore for VPAC1 receptors. J Pharmacol Exp Ther 303(2):445–460
10. Dong M et al (2011) Importance of each residue within secretin for receptor binding and biological activity. Biochem 50(14): 2983–2993
11. Dean T et al (2006) Role of amino acid side chains in region 17–31 of parathyroid hormone (PTH) in binding to the PTH receptor. J Biol Chem 281(43):32485–32495
12. Adelhorst K et al (1994) Structure-activity studies of glucagon-like peptide-1. J Biol Chem 269(9):6275–6278
13. Nicole P et al (2000) Identification of key residues for interaction of vasoactive intestinal peptide with human VPAC1 and VPAC2 receptors and development of a highly selective VPAC1 receptor agonist. J Biol Chem 275(31):24003–24012
14. Kortemme T, Baker D (2002) A simple physical model for binding energy hot spots in protein–protein complexes. Proc Natl Acad Sci USA 99(22):14116–14121
15. Kortemme T, Kim DE, Baker D (2004) Computational alanine scanning of protein–protein interfaces. Sci STKE 219:12
16. Bordner AJ, Abagyan RA (2004) Large scale prediction of protein geometry and stability changes for arbitrary single point mutations. Proteins: Struct Funct Bioinform 57(2):400–413
17. Abagyan R, Totrov M (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. J Mol Biol 235(3):983–1002
18. Cohen M et al (2008) Similar chemistry, but different bond preferences in inter versus intra protein interactions. Proteins: Struct Funct Bioinform 72(2):741–753
19. Dong M et al (2011) Lactam constraints provide insights into the receptor-bound conformation of secretin and stabilize a receptor antagonist. Biochem 50(38):8181–8192
20. Rose PW et al (2011) The RCSB Protein Data Bank: redesigned web site and web services. Nucleic Acids Res 39(suppl 1):D392–D401
21. Abagyan R, Totrov M, Kuznetsov D (1994) ICM-A new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. J Comput Chem 15(5):488–506
22. Momany FA et al (1975) Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. J Phys Chem 79(22):2361–2381
23. Nemethy G, Pottle MS, Scheraga HA (1983) Energy parameters in polypeptides. 9. Updating of geometrical parameters, nonbonded interactions, and hydrogen bond interactions for the naturally occurring amino acids. J Phys Chem 87(11):1883–1887
24. Nemethy G et al (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(15):6472–6484
25. Abagyan R (1997) Protein structure prediction by global energy optimization. In: van Gunseteren WF, Weiner PK, Wilkinson AJ (eds) Computer simulation of biomolecular systems: theoretical and experimental applications. Kluwer, Norwell, MA, pp 363–394
26. Ulrich CD 2nd et al (1993) Intrinsic photoaffinity labeling of native and recombinant rat pancreatic secretin receptors. Gastroenterology 105(5):1534
27. Munson PJ, Rodbard D (1980) Ligand: a versatile computerized approach for characterization of ligand-binding systems. Anal Biochem 107(1):220–239
28. Van Der Spoel D et al (2005) GROMACS: fast, flexible, and free. J Comput Chem 26(16):1701–1718
29. Lindorff Larsen K et al (2010) Improved side chain torsion potentials for the Amber ff99SB protein force field. Proteins: Struct Funct Bioinform 78(8):1950–1958
30. Jorgensen WL et al (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79(2):926–935
31. Berendsen HJC et al (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81(8):3684–3690
32. Parrinello M, Rahman A (2009) Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys 52(12): 7182–7190
33. Hess B et al (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18(12):1463–1472
34. Essmann U et al (1995) A smooth particle mesh Ewald method. J Chem Phys 103(19):8577–8593
35. Pozvek G et al (1997) Structure/function relationships of calcitonin analogues as agonists, antagonists, or inverse agonists in a constitutively activated receptor cell system. Mol Pharmacol 51(4):658
36. Robberecht P, Conlon TP, Gardner JD (1976) Interaction of porcine vasoactive intestinal peptide with dispersed pancreatic acinar cells from the guinea pig. Structural requirements for effects of vasoactive intestinal peptide and secretin on cellular adenosine 3′: 5′-monophosphate. J Biol Chem 251(15):4635
37. Turner JT, Jones SB, Bylund DB (1986) A fragment of vasoactive intestinal peptide, VIP (10–28), is an antagonist of VIP in the colon carcinoma cell line, HT29. Peptides 7(5):849–854
38. Bodanszky M et al (1978) Synthesis and some pharmacological properties of the 23-peptide 15-lysine-secretin-(5–27). Special role of the residue in position 15 in biological activity of the vasoactive intestinal polypeptide. J Med Chem 21(11):1171–1173
39. Gardner JD et al (1979) Interaction of secretin 5–27 and its analogues with hormone receptors on pancreatic acini. Biochim Biophys Acta 583(4):491–503
40. Robberecht P et al (1988) Secretin receptors in human pancreatic membranes. Pancreas 3(5):529–535
41. Potapov V, Cohen M, Schreiber G (2009) Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details. Protein Eng Des Sel 22(9):553–560
42. Benedix A et al (2009) Predicting free energy changes using structural ensembles. Nat Methods 6(1):3–4
43. Yin S, Ding F, Dokholyan NV (2007) Eris: an automated estimator of protein stability. Nat Methods 4(6):466–467
44. Kellogg EH, Leaver Fay A, Baker D (2011) Role of conformational sampling in computing mutation induced changes in protein structure and stability. Proteins: Struct Funct Bioinform 79(3):830–838
45. Runge S et al (2003) Different domains of the glucagon and glucagon like peptide 1 receptors provide the critical determinants of ligand selectivity. Br J Pharmacol 138(5):787–794
46. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22(12):2577–2637