Computational studies on thyrotropin-releasing hormone (TRH) and the TRH-receptor

by

Liisa Laakkonen

A dissertation submitted to
the Graduate Faculty in Biomedical Sciences
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy,
The City University of New York

1995

This manuscript has been read and accepted for the Graduate Faculty in Biomedical Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

August 30, 1995
Date

Dr. Harel Weinstein
Chair of Examining Committee

August 31, 1995
Date

Dr. Terry A. Krulwich
Executive Officer

Members of the Examining Committee:

Dr. Marvin Gershengorn

Dr. Charles Hutchins
Abbot Company
Chicago, Illinois

Dr. Roman Osman

Dr. Massimo Sassaroli

Dr. Stuart Sealfon

THE CITY UNIVERSITY OF NEW YORK

Abstract

COMPUTATIONAL STUDIES ON THYROTROPIN-RELEASING HORMONE (TRH) AND THE TRH RECEPTOR

by

Liisa Laakkonen

Advisor Dr. Roman Osman

The recognition complex between TRH and the TRH receptor has been studied in several ways. A molecular model of the TRH receptor has been built *de novo*, and its stability has been tested by molecular dynamics simulations. The construction has been implemented as an automated protocol. The ligand has been docked inside the receptor model, guided by four direct ligand-receptor interactions that have been identified during the course of the project in collaboration with Dr. Gershengorn's laboratory at Cornell Medical Center. The complex has been studied extensively with Monte Carlo-stochastic dynamics simulations. This is the first time the new computational methodology has been applied to a large complex, and its effectiveness has been examined. The conformations accessible to the ligand and to two constrained analogs have been studied separately in water. A binding conformation of TRH can be proposed based on both the conformational searches and the simulations of the complex.

Acknowledgements

All work described in this thesis has been done under the leadership of Dr. Roman Osman. I am impressed, and hope to have been influenced by his systematic way of working and his respect for chemistry. Our TRH-path has not always been smooth and I do not yet write nice English, but I count friendly survival as a victory.

I am grateful to the members of my advisory committee, Drs. Marvin Gershengorn, Joseph Kushick and Harel Weinstein, for their participation in the project, and Drs. Charles Hutchins, Massimo Sassaroli and Stuart Sealfon for serving as outside examiners of the work. Without the experimental work in Dr. Gershengorn's laboratory, mostly by Dr. Jeffrey Perlman, this project would not have been possible. The smaller giants of the Department of Physiology and Biophysics, Drs. David Garmer, Frank Guarnieri and Chung Wong have been most helpful in explaining and discussing their specialties whenever asked.

Dr. Weinstein I would like to thank additionally for the care he has shown for my well-being and for sharing his strong and beautiful vision of structural biology. Kiitos.

The personal thanks go to people who have made me enjoy the last five years by loving me and letting me love them, and told me stories about all the continents of the world: Waldo Feng and David Kombo; all Thursday guests and lunch company; the pan-American Dominoes. I want to thank Dr. Angeliki Buku for the Metropolitan Opera; Family Eisinger for the American outdoors; all my families in Finland for bread, books and letters; my sister Sirkka Laakkonen for her daily email messages; and my niece Marjatta and nephew Johannes Hyrsky who insist that the summit of life is reached in a Ferris wheel.

My system manager and husband Janne Ravantti I thank both for awk and roses.

Table of contents

List of Tables

List of Figures

1 Introduction

The computational project on the tripeptide hormone TRH (pGlu-His-Pro-NH_2, thyrotropin releasing hormone) and its receptor TRHR started from emerging sequence-related data on the receptor (Straub 1991), without any preceding theoretical work. The project has evolved around questions raised by experimental data. As a result, a wide variety of methods have been used and very different problem areas have been studied. The target of the study is the binding of TRH to the TRH receptor. Because of the complexity of the system, all parts of the binding complex have been studied separately to allow the use of more exact computational methods. Results from the studies of the partial systems have guided the design and analysis on studies of the entire binding complex. The subject matter covered in this work is most logically ordered by increasing size of the molecules, with an according change of the methods used.

The first part of the thesis work deals with the amino acid proline, which is the third and last residue in TRH. The conformational energies of proline were calculated with quantum mechanical methods to address the role of *cis-trans* isomerization of the peptide bond before proline in the action of TRH. The idea of isomerization of the ligand as the activation mechanism of the receptor was not followed in the simulations of the complex, but the work on Pro stands on its own. The results in vacuum are exact, and agree with the closest possible experimental data. Problems were encountered with inclusion of solvation.

In the second part, studies on conformational preferences of TRH and some of its analogs are presented, focusing on two newly synthesized constrained analogs that show very different potencies on the TRHR. Conformational search of peptides is an active research area, and the importance sampling Monte Carlo method used (Guarnieri 1995) represents one of the most recent

advances in the field. The general questions of conformational searching are discussed in Chapter 1.2.

De novo construction of a model of the TRH receptor is the most daring undertaking of this work. There is no structural data on the receptor other than the sequence, and limited knowledge of any homologous protein. The construction is based on a published analysis of conserved physical and biological properties in a large set of related sequences. The construction procedure has been implemented as a protocol for systematic and easy rebuilding, if and when more data on related structures emerge. The receptor structure has been refined by energy minimization and molecular dynamics simulations. General questions of model building are covered in Chapter 1.3.

Collaborative molecular biological and computational studies of the receptor have produced data on three strong direct contacts between TRH and TRHR (Perlman 1994; Perlman 1995). Based on these data the ligand was docked inside the receptor model, and the system was studied with mixed mode Monte Carlo / stochastic dynamics simulations (Guarnieri 1994). As a result, new predictions of the role of specific residues in the receptor can be made, and also a binding conformation of the ligand is proposed. Computational studies of molecular recognition are discussed in the last part of the introduction.

1.2 Study of peptide conformations

The basic assumption in molecular structural calculations is that molecules exist in the state with lowest free energy. This state can be found as the deepest and widest valley of the potential energy surface, which is a function of all the degrees of freedom in the system. The depth of a valley corresponds to the enthalpy of a structure, and the width to its entropy. Locating the global minimum is not easy, because potential energy surfaces have multiple local minima,

often separated from each other by high barriers. Small peptides and proteins differ qualitatively from each other regarding their three-dimensional structure: proteins adopt one well-defined structure while peptides fluctuate between several conformers. Accordingly, the multiple minimum problem has to be rephrased for peptides: finding the global minimum is not the only point of interest, but it is equally important to define the distribution of all thermally accessible states, because several of them are likely to be populated. A direct consequence of several low-energy conformations is that interactions with the surroundings can dramatically change the relative energies of the conformers, and a different conformation may become predominant. This is seen both experimentally and theoretically. For example, the cyclic peptide cyclophilin turns inside out compared to its conformation in chloroform when it binds to its receptor cyclosporin (Wütrich 1991). In proline studies in Chapter 3 it is shown that the structure of AcPro-NH_2 in crystals is not the global minimum in vacuum. The crystal structure, however, is able to form several intermolecular hydrogen bonds, that provide the extra stabilization.

If the quality of the empirical force fields is left undiscussed, the multiple minima problem can be approached with two different strategies. One can try to enhance sampling of the conformational space by either using systematic search methods or random searches, all with the standard force fields. The alternative method is to decrease the effective barriers between different minima, by increasing the temperature or by simplifying the potential energy surface. See Figure 1.1 for a schematic representation of the different methods. Modifications of the force field and enhanced sampling are often used together for best results. See Vásques et al. for a recent review (Vásques 1994).

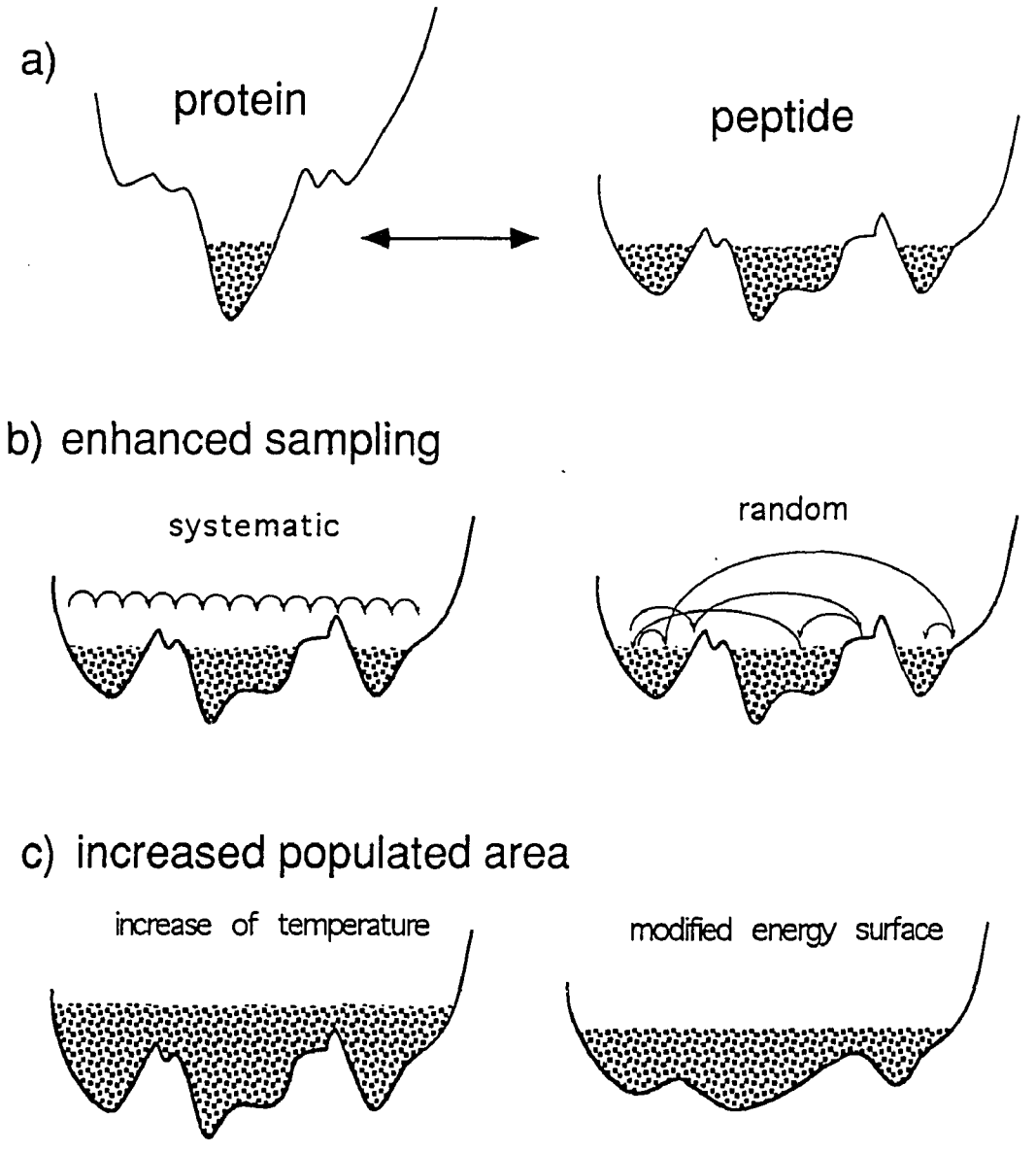


Figure 1.1. A diagram of the potential energy surface of a peptide and of the methods used to study it exhaustively. a) Comparison between a protein with a well-defined 3D structure and a peptide with multiple populated conformations. The shaded area shows thermally accessible states. b) Systematic and stochastic searching. c) The increase of the populated areas and disappearance of barriers by increased temperature and simplified force field.

1.2.1 Enhanced sampling methods

Systematic searches. A systematic grid search is the only method that guarantees to find all low-energy conformers. All relevant degrees of freedom -usually dihedral angles- are changed discreetly over the entire available range, and the rest of the system is minimized. The discretization has to be of high enough resolution, or minima might be overlooked. For most cases, 30° is a good step size for torsional angles; and 12 steps per angle are needed. This is where problems arise. The number of possible states is R^N, where R is the number of values per one angle, and N is the number of independent angles. The number of possible states increases very rapidly with the number of free torsional angles: the simplest protein model, Ac-Ala-NHMe has 2 free torsional angles, and $12^2 = 144$ states, but a blocked arginine with 6 free torsions has $12^6 \sim 3 * 10^6$ states. Exhaustive minimizations are slow, and it is not practical to plan to minimize but a few thousand starting structures, which limits grid searches to backbones of dipeptides. The searches are simplified, if there are constraints that limit the range of angles, or if the system is symmetric, but even in these cases a pentapeptide is the upper limit for this methodology (Saunders 1990).

Random searches. Random searches in the conformational space are performed by Monte Carlo methods (Metropolis 1953). They usually also operate on dihedral angles only. An angle, or a few angles simultaneously, are selected randomly to be changed by a random amount, and the new conformation is evaluated. The most commonly used evaluation method is by Metropolis criteria (Metropolis 1953). If the new conformation has lower energy than the previous one, it is automatically accepted. If not, a random number between 0 and 1 is generated. When the Boltzmann factor of the energy difference between the

structure to be evaluated and the one it was generated from is less than the random number, the structure is accepted; otherwise it is not. The method allows for sporadic increases in energy, by which the molecule can escape from minima. Monte Carlo methods are good in searching the conformational space, and they can handle large systems, but they are relatively slow, because most of the time they sample high energy structures that are rejected by the Metropolis criteria. Millions of trials have to be performed to reach convergence. Commonly several runs with different seed numbers are performed, and the results are combined for enhanced sampling. There are also sampling schemes that target the random trial steps to populated areas of the search space, and consequently increase the efficiency of sampling (Guarnieri 1995).

Molecular dynamics simulations can also be used in searching the conformational space of peptides. The method is uniquely good in following time dependent changes in the immediate environment of a structure. Dynamic simulations, however, do not allow for large deviations from the starting structure during normal simulation times in the order of 100 - 1000 ps. Accordingly, the most interesting dynamical methods in conformational searching are those that apply heat to the system or deform the potential energy surface.

1.2.2 Enhancing the populated conformational space

High temperature, quenching and annealing. Heat in molecular simulations means increased kinetic energy. An insurmountable barrier at 300 K may disappear at 1000 K. For example, the barrier between *cis* and *trans* isomers of peptide bonds for prolines is about 80 kJ/mol, and that between the endo and exo puckered states about 5 kJ/mol. In molecular dynamics simulations at room temperature frequent changes of the puckering are observed (Schmidt 1993), but none of the conformation of the peptide bond. The system has to be heated

up to ~ 2000 K to make it ignore the isomerization barrier (Brünger 1988). It is, however, usually not of interest to study the high temperature effects themselves, but the molecules are cooled back to room temperature, and the role of the high energy simulations is to produce structures in a nonsystematic automatic fashion. The structures can be cooled fast with normal miminization methods. Such quenched dynamics simulations have produced good results for the conformational search of bis-penicillamine enkephalin, compared to NMR data (Petitt 1991). Besides helping to overcome barriers, heat can also have adverse effects on the molecules. In a comparative study of different search methods it was noted, that after minimization, simulations at 1000 K produced many high energy structures (Saunders 1990). An alternative to quenching is gradual cooling, or annealing of the system (Kirkpatrick 1983), by reducing the temperature in small steps, in analogy to slow cooling of a pure liquid to a highly ordered crystalline state. If the changes are infinitesimally small, the system is able to reorganize while it loses energy, and will end in the global minimum. Infinitesimal changes are never possible, and the best cooling scheme seems to depend on the nature of the problem. Annealing can be used both in combination with either Monte Carlo or dynamics simulations.

Systematic simplification of the potential energy surface. A totally different approach is to modify the potential energy surface. This has been done for example by increasing the dimensionality of the system, by stepwise elimination of the barriers, and by simulating a part of the system in the mean field of its neighbors (Crippen 1987; Piela 1989; Roitberg 1991; Vásques 1994).

One method of smoothing potential energy surfaces is the diffusion equation method, DEM (Piela 1989; Kostrowicki 1992; Vásques 1994). In this methodology, the force field is modified gradually by addition of its second derivatives. The addition of second derivatives -negative at maxima and positive at mi-

nima- reduces both the barriers between minima and the depth of the energy wells, leading after repeated additions to a modified force field with only one minimum. The procedure is stopped at this point, and the minimum is located easily by any minimization method. The deformation of the force field is reversed and the position of the minimum is followed back to the original force field. The most critical parts of the method are the assumption that the single minimum of the modified force field includes the original global minimum, and the reliability of the backtracing. The results of DEM are promising, the global minima of alanine dipeptide and Met-enkephalin are located fast and easily (Kostrowicki 1992), and up to a 36-residue peptide can be folded to its native conformation (Vásques 1994).

A method of similar spirit is relaxation of the dimensionality of the system (Crippen 1982; Crippen 1987). The position of each atom in a molecule of N atoms is expressed relative to all its neighbors, and in this N-1 dimensional space energy functions often have only one minumum. The system is forced back to 3D by gradually projecting out components corresponding to lowest eigenvalues. The idea of extra dimensions (4D) has also been combined to temperature and restrained molecular dynamics in NMR refinement (Schaik 1993).

Locally enhanced sampling technique (LES) is designed for determining partial structures, for example side chain positions, in X-ray structure refinement or homolgy modeling (Roitberg 1991). Multiple copies of the residues studied are built in the system, and the energy of the new mean field system is optimized in standard CHARMM force field (Brooks 1983). This does not change the global minimum, and the barriers between local minima are reduced, because the copies are never simultaneously in the high energy positions, and thus the mean energy is lower. The approach has been combined with simulated an-

nealing in conformational studies of tetrapeptides, and a few sidechains of bovine pancreatic trypsin inhibitor BPTI (Roitberg 1991).

There are also methods for structure determination that rely on statistical data of known structures, and refine the combinations of most probable conformations -usually side chains- in a force field. All possible optimization methods have been used, from simulated annealing (Leach 1994) to genetic algorithms and artificial neural networks (Head-Gordon 1993). The goodness of the results depends on the data base.

1.3 Molecular models of G-protein coupled receptors

The receptor for TRH belongs to the superfamily of G-protein coupled receptors (GPCR). Members of the family transduce widely varying signals, carried by light, neurotransmitters and peptide hormones, into cellular responses. GPCRs act by activating G-proteins, which in turn stimulate various effector systems; for example phospholipase C is stimulated by TRH receptor. In addition to functional coupling to the same mediators, these receptors share extensive sequence homology. Sequence analysis suggests that all GPCRs have seven hydrophobic domains with helical periodicity (Donnelly 1989; Dahl 1991; Baldwin 1993), and are therefore considered to be bundles of seven transmembrane helices. The only experimentally resolved structure of a seven helix transmembrane protein, in atomic detail, is that of bacteriorhodopsin (brd) (Henderson 1990). Brd is neither in function nor in sequence similar to the GPCRs. Nevertheless, models of various GPCRs have been constructed based on brd (Hibert 1991; Trumpp-Kallmeyer 1991; Ijzerman 1992; Cronet 1993; Nordvall 1993) in an effort to bring experimental data into a structural framework. More recent data on rhodopsin (rhd) (Schertler 1993; Unger 1995), itself a member of the

GPCR family, have made the use of brd as a template questionable, because a low resolution crystal structure of rhd shows a footprint interestingly different from that of brd. The structure at 9 Å is, unfortunately, too indefinite to serve as a template for homology modeling, and no spatial data has been released.

An alternative approach in modeling is to build structures *de novo* (Findlay 1990; Cronet 1993; Kontoyianni 1993; Ballesteros 1995). This route is less well defined but is feasible even in the absence of an acceptable template. Two assumptions are necessary for building *de novo* models of GPCRs. First, that similar rules govern helix packing in all transmembrane proteins, and guidelines can therefore be derived from the two known structures, bacteriorhodopsin (Henderson 1990) and the photoreaction center (Deisenhofer 1989). Second, one has to assume that similarities in GPCR sequences imply structural likeness, and conservancy data from an aligned set of sequences can be used for structural inferences. The first assumption produces physicochemical rules: helix surfaces facing the lipid are hydrophobic, whereas hydrophilic and charged residues face the interior of the helix bundle (Donnelly 1989; Baldwin 1993; Cronet 1993). The second assumption of structurally similar GPCRs attaches meaning to those sites in the sequences, where either the identity, character, size or H-bonding properties are conserved. In an evolutionary sense, only structurally or functionally meaningful residues are worth conserving. Many of the conserved residues in GPCRs have actually been shown to be important, *e.g.* the highly conserved Asp in helix 2 has a functional role, and Pro in helix 7 structural (See Baldwin 1994 and references therein).

Five steps of model building. There are five steps in constructing a receptor model *de novo*: determining the positions of helical segments in the sequence, forming the helices, orienting them relative to each other, deciding of the relative heights of the helices, and forming the helix bundle. The starting

point is an alignment of related sequences, because all reasoning about conserved features is relevant only within an alignment. It is possible to compare only closely related sequences or to include a wide variety of sequences into the alignment. Aligning varying sequences is hard, but such an alignment is more informative about the general structural features of the family, because all ligand and subfamily specifity is averaged. The alignment is read with two assumptions in mind - that all transmembrane proteins pack similarly and that only functionally or structurally important residues are conserved - and in context of the physical environment of the transmembrane proteins: lipid membranes are hydrophobic, with negatively charged head groups. Accordingly, the faces of the protein that are in contact with the membrane should not include charged or polar residues, but positive charges at the ends of helices are accepted (Ballesteros 1992). The exact identity of the hydrophobic residues should not matter, and thus the outside of the helix bundle should be less conserved than the inside of the bundle. An extension of the importance of hydrophobic faces in packing is a requirement that no insertions or deletions should exist inside the helices, because they shift the periodicity of the helix. Because of the same reason, positions of sequence motifs that may kink and twist the helices should be conserved. This applies at least to prolines, but possibly also to other residues. It is also known, that sequences with repeating prolines, PP and PXP have never been observed in helical conformation (MacArthur 1991). Prolines cannot form hydrogen bonds in the N-terminal direction, and in globular proteins they are preferred as first two residues of a helix, but not likely to exist in the last turn of a helix (Presta 1988).

Sequence based data: helical boundaries and hydrophobic faces. The alignment serves directly for determination of helical boundaries and for orientation of the helices. Simple hydrophobicity plots (Kyte 1982) give first clues

about the positions of the helices. More exact positions can be derived from helically periodic occurence of hydrophobic residues, and of highly conserved residues. The same segments should have no insertions or deletions (Donnelly 1994). Positively charged residues help to determine the intracellular ends of helices (Ballesteros 1992). The degree of conservation in a helix can be quantitated (Donnelly 1989), and with the assumption that most conserved is most buried, it orients the helices. Also, data about residues that are critical to binding of ligands, or crossreact with photoactivated analogs is used. The determination of helix boundaries is exact to 1-2 residues in the intracellular side and to about one turn in the extracellular side (Ballesteros 1995). Most data on conserved residues and on important ligand binding residues agree (Baldwin 1994), and the overall orientation of the helices can be determined.

3D Structure: helices, packing, relative heights. The three other aspects of model building are directly related to 3D structure, and more difficult to derive from the sequences. The structure of the individual helices can be approximated with ideal α-helices, and relaxed with molecular dynamics. It should be remembered, however, that regular helices are an approximation, and that irregularities would change the faces of the helices (Konvicka 1995). Studies on helix packing deal with globular proteins (Chothia 1981), and they give average helix-helix distances and tilts. The distances compare well with bacteriorhodopsin and photoreaction center structures. The extent of the hydrophobic surface and the way it varies along the height of the helix are indicative of the packing and the tilt of the helix. The most problematic issue in packing the helices is their relative vertical positions. In the absence of a crystal structure, the heights could be resolved by distance geometric methods, if there were data on interhelical contacts between every pair of helices. There are two reports of direct interhelical contacts (Rao 1994; Zhou 1994) which is too little to define the relative

heights. The two known interhelical contacts serve best in comparing possible packing models. With this state of data, any choice of relative heights for helices in a model is good, as long as it agrees with other data.

Several *de novo* models of GPCRs have been built (Dahl 1991; MaloneyHuss 1992; Cronet 1993; Zhang 1993), incorporating many of the guidelines mentioned above. The main problem with such models is that the constraints for positioning and orienting helices are weak, and the number of acceptable ways to combine the helices in a bundle is very large. At present, there are no self-consistent tests to probe the validity of these models. The models in their atomic detail do, however, serve well in formulating hypotheses on the roles of specific residues. Such questions cannot be based on mutational data only, but they can be tested experimentally. A reciprocal approach of modeling and mutational work, where theoretical predictions are tested experimentally, and new experimental results incorporated into improved molecular models seems most promising.

1.4 Formation of recognition complexes

Study of ligand-receptor interactions addresses the fundamental biological problem of molecular recognition. Determination of the thermodynamically most favorable modes of binding requires answers to the following questions: Which are the specific interactions that place ligand and receptor productively together? Are there major rearrangements upon binding? How is the ligand guided to the receptor, and how is it released? In computational studies of complexes, three aspects cause most problems: dealing with the flexibility of the molecules, reliability of the energy evaluations and the possible effects of the environment on both.

Several totally different approaches have been used in constructing ligand-receptor complexes: molecular dynamics, Monte Carlo, simulated annealing and distance geometry (Kuntz 1982; Goodford 1985; DesJarlais 1988; Goodsell 1990; Yue 1990; Burt 1991; Cherfils 1991; Miranker 1991; Hart 1992; Cherfils 1993; Di Nola 1994; Eisen 1994; Leach 1994; Teeter 1994; Zhang 1993). Different methods work for different problems and require very different amounts of data. The study of a ligand-receptor complex can mean, for example, looking for a ligand to a known binding site, or searching for the binding site in a known protein structure, or studying the evolution of an already formed complex.

Searching for a good ligand. Search for a ligand to a known binding site in a resolved structure is the normal situation in drug design. For reviews, see Balbes and van Gunsteren (Balbes 1994; Gunsteren 1994). In the simplest and most used approach molecules are docked to each other by rotating and moving a rigid ligand in the rigid assumed binding site, to "locate feasible binding orientations between two structures" (Kuntz 1982). This can be done for a series of compounds from large data bases (Kuntz 1982), or for separate functional groups, that are then combined to one molecule based on the best fitting spatial relationship (Bohm 1992; Eisen 1994). The fitting can be relaxed by allowing some spatial overlap by using soft surfaces. Complexes that are accepted by good geometric fit can also be rated by molecular mechanics interaction energies (DesJarlais 1988). Geometric fit is a necessary condition for a good complex but not sufficient. There is no reason to assume that the ligand fill the whole binding pocket -ions and water can also be involved- or that the shape of the pocket would not change, or that the binding site data, often derived from agonist studies, would aply to chemically different agonists. In fact, data exists arguing against all the three points mentioned. First, even though for large proteins one can assume that the overall structure does not change, the binding pocket

usually adjusts to the ligand (Rydel 1991). Second, water molecules are observed in crystal structures mediating the interactions between the ligand and the protein (Bolin 1982). Third, different binding of agonists and antagonists to the same receptor is demonstrated by the many mutations that affect their binding differentially (Perlman 1992).

Searching for a binding site. Search for the binding site on a protein surface is a different problem, and it requires a fully mobile ligand or probing group. The probe is moved around the protein either systematically or by molecular dynamics or Monte Carlo simulations (Goodford 1985; Goodsell 1990; Miranker 1991; Nicholls 1991; Caflisch 1992; Hart 1992; Di Nola 1994; Leach 1994). The cost of multiple energy evaluations is high, and one way around it is to use affinity potentials instead of exact molecular mechanics force field energies. The affinity potentials are look up tables for different relative positions of all the functional groups that are present in the complex (Goodsell 1990) . Because conformational flexibility not only complicates the geometric fitting, but also increases the number of energetically available states (Lee 1994), it is possible to allow only predetermined conformations for each side chain (Leach 1994). Hot temperature runs (Goodsell 1990; Di Nola 1994) and use of multiple, 1000 - 5000, copies of the ligand to be studied (Miranker 1991) or use of multiple starting structures enhance the searching. Test studies with all methodologies claim success, which is defined as finding the experimentally known complex among the few energetically most favorable structures. False positives are not treated as a problem, but considered as possible different binding configurations (Cherfils 1991). The limiting factor for locating the binding site often seems to be the rigidity of the protein molecule, especially when the structural data are of an uncomplexed protein. Rigid protein structure often excludes possible favorable complexes from further consideration. The correct complexes were not found in

the cases when the protein had distinctly different conformation in complex (Goodsell 1990; Hart 1992; Miranker 1991). This emphasizes the importance of good conformational sampling of all residues in the binding site, and suggests Monte Carlo methods for searching the binding site.

Studying a possible complex. When the goal of the study of a ligand-receptor complex is to follow the evolution of a structure, or to study the feasibility of proposed contacts, flexibility and detailed energy evaluations are necessary requirements. Methods used in these studies are standard molecular dynamics and Monte Carlo simulations, both of which have their advantages and disadvantages. Monte Carlo methods allow for conformational changes and thus do not require the knowledge of exact structures to start with. Most Monte Carlo programs do not, however, allow for bonds or angles to change, and it is usually not possible to include the surroundings into the calculations. Molecular dynamics methods are usually used when the starting structure is reasonably good. Flexibility of both interaction partners is allowed. The energy evaluations from modern empirical force fields are reliable, and adding water or ions is not a problem. Molecular dynamics methods have been used in the study of 5-HT receptor and its ligands (Smolyar 1993; Zhang 1993a; Zhang 1993b).

It is also possible to combine Monte Carlo and dynamic simulations in a new mixed mode method (Guarnieri 1994). This methodology allows for conformational search, because a large part of the complex can be kept active in Monte Carlo steps; and it also avoids many of the collisional high energy conformations, because the entire system is allowed to relax in the dynamics steps. Both aspects are especially valuable in the study of a model structure. See Chapter 6 for an example of this methodology.

2 Background

2.1 Physical properties of TRH

The tripeptide pGlu-His-ProNH_2 can be characterized as a pseudosymmetric molecule of 26 non-hydrogen atoms with three nonfused five-membered heterorings. It has five possible hydrogen bond donors and five acceptors. The 3D structure of TRH is known from a TRH tartrate crystal (Kamiya 1980), and it is shown in Figure 2.1 on the next page. The main determinants of the structure are the dihedral angles in the backbone, and the orientation of the histidine side chain. The bulky rings in the structure stay apart from each other. The rings differ from each other conformationally, the aromatic imidazole of the histidine is planar while the pyrrolidone of the pyroglutamic acid and the pyrrolidine of the proline are C_γ-endo puckered. There is a hydrogen bond in the crystal structure from the N_π of imidazole to the terminal amide oxygen. The crystal structure, however, is not necessarily relevant for the biological functions of TRH, because it is a tightly packed solid state structure, with several intermolecular hydrogen bonds, and no solvent. Furthermore, TRH is protonated in the crystal, and in aqueous solution at physiological pH it is not ($pK_{a(TRH)}$ = 6.25 (Grant 1972)).

Another source of structural information is NMR spectroscopy. There are several NMR studies on TRH (Anteunis 1981; Deslauriers 1973; Deslauriers 1974; Donzel 1974; Feeney 1974; Montagut 1974; Sievertsson 1974; Donzel 1975; Haar 1975; Bellocq 1976; Bellocq 1977; Liakopoulou-Kyriakides 1979; Vicar 1979; Galardy 1982; Unkefer 1983; Mapelli 1989; Mapelli 1990). TRH is dissolved in D_2O or DMSO in these studies, and comparing structures in the two solvents shows how much the peptide can adapt to the environment. The overall backbone conformation inferred from NMR is extended, as in the crystal.

Figure 2.1
The structure of TRH in TRH^{+}-$tartrate^{-}$ crystal

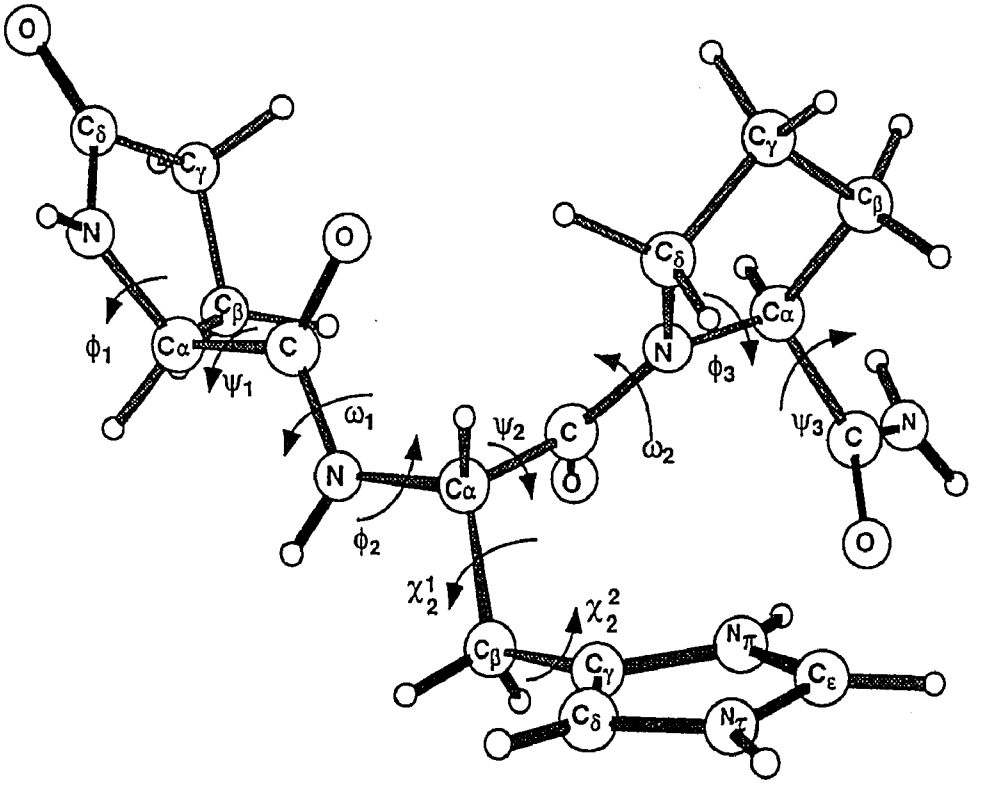


Table 2.1 Torsion angles of TRH

angle	crystal		spectroscopy
ϕ_1	108.0°	=	90° in DMSO[2], 110°-120° in D_2O[5]
ψ_1	146.2°	x	15° in D_2O[5], -20° (aq), 120°(s)[3], Raman spectr.
ω_1	171.3°	=	always trans
ϕ_2	-69.7°	≠	-150° in DMSO[2], -150° or 90° in D_2O[5]
ψ_2	136.6°	x	155° in DMSO[2], 75°-120° or 120-150° in D_2O[5]
ω_2	175.0°	x	180° or 0° in DMSO and D_2O[1,3]
χ^1	-163.1°	x	all rotamers
χ^2	-71.1°	≠	90° in D_2O[5]
ϕ_3	-62.7°	=	-60° – -90° in D_2O[5]
ψ_3	153.5°	x	150°, D_2O[2], -55° in D_2O and DMSO[3], neither direct

= : ϕ_1, ω_1, and ϕ_3 are similar by all methods, ≠ :ϕ_2, χ^2 different, x: ψ_1, ψ_2, ω_2, χ^1, ψ_3 vary
Ref:1. Deslauriers 1973, 2. Donzel 1975, 3. Montagut 1974, 4. Unkefer 1983, 5. Vicar 1979

The values for the dihedral angles are similar to the crystallographic ones, except for $\phi 2$ and $\chi 2$. The real difference, however, is that for five of the ten torsional angles there are two or more values from spectroscopy, reflecting the flexibility of solvated TRH. See Table 2.1.

NMR experiments also propose interresidue interactions. The existence of intramolecular interactions in TRH is suggested by comparisons between TRH and peptides that have only one or two of the residues of TRH, and thus cannot have properties that depend on interresidue interactions. For example, comparison of ^1H NMR spectra of pGlu-NHMe, pGlu-D-His-Pro-NH_2 and TRH suggests, that pGlu does not interact with the rest of the TRH molecule, because changing the group attached to pGlu does not affect its spectrum (Deslauriers 1973; Donzel 1974). The same conclusion is reached from studies of individual ^{13}C relaxation times for hydrogen containing carbons, that show an increased mobility of pGlu relative to the rest of the molecule in D_2O (Feeney 1974). Since pGlu has only one degree of freedom, the rapid relaxation is assigned to the rotation around the ψ_1 angle. The pyroglutamic acid can thus be considered as a conformationally independent residue. Accordingly, the coupling of properties, suggested by NMR, has to exist between His and Pro.

Histidine and proline both have several different structural states. The His-Pro peptide bond can exist in either *cis* or *trans* conformation, as any X-Pro bond (Thomas 1972; Wütrich 1972; Deslauriers 1973; Montagut 1974). Histidine has a free side chain with three different rotameric states (Donzel 1974; Montagut 1974; Vicar 1979; Unkefer 1983), and the imidazole ring is involved in two equilibria: one between the protonated and unprotonated states, with pKa = 6.25 (Grant 1972), and the other between the two tautomers, $N_\tau H$ and $N_\pi H$ (Deslauriers 1973; Montagut 1974). See Figure 2.2.

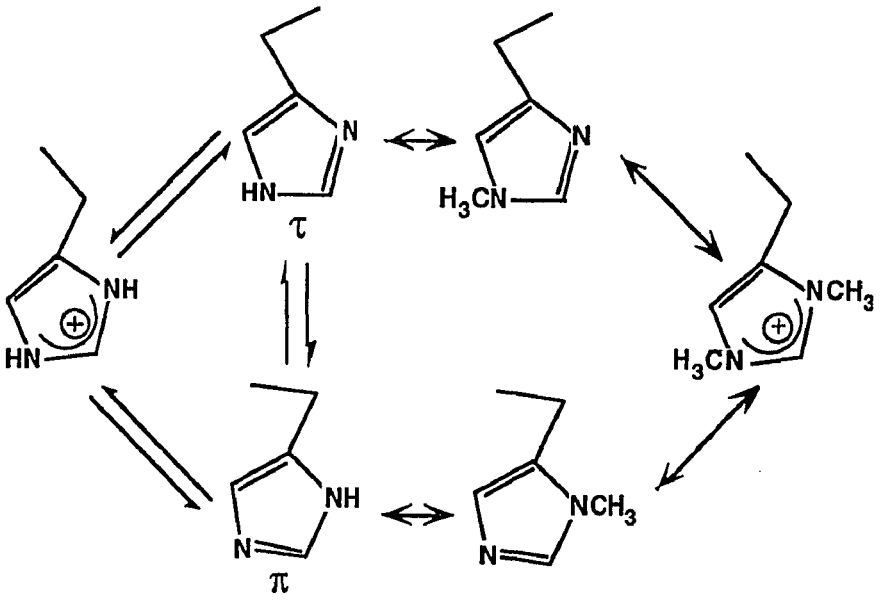


Figure 2.2. Different forms of histidine: τ and π tautomers, protonated imidazolium, and the three possible methylated structures. The paired arrows represent chemical equilibria, and the doubleheaded arrows possible reactions.

All these properties -isomeric, rotameric and tautomeric states- are coupled; changes in the state of histidine affect the *cis-trans* ratio (Deslauriers 1973; Donzel 1974). The deviation of TRH from the behavior of model compounds is larger in DMSO than in water (Donzel 1974).

The *cis/trans* ratio is affected by the nature of groups both preceding and following Pro: a bulky residue, like His, before Pro is shown to increase the *cis*-population (Grathwohl 1976). Counterintuitive to this, the proportion of the *cis* form is much lower in TRH than in Ac-Pro-NH_2, 15% and 25%, respectively, and the population distribution is solvent dependent in TRH, while Ac-Pro-NH_2 shows no solvent dependence (Deslauriers 1973).

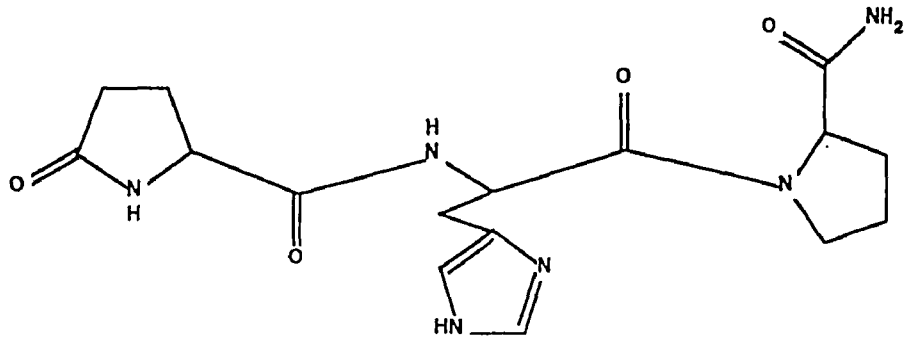
The pK_a of TRH is lower than that of imidazole, 6.25 compared to 7.0, and it is slightly different in the two isomers (Grant 1972). The two tautomers $N_\tau H$-TRH and $N_\pi H$-TRH are equally populated in DMSO (Reynolds 1973; Montagut 1974), but in water $N_\tau H$-TRH is the predominant form, approximately in 4:1 ratio (Deslauriers 1973). The tautomeric equilibrium is slightly different in the *cis* and *trans* isomers in water (Unkefer 1983). Methylation of N_τ or N_π of the imidazole ring eliminates tautomerism and affects the conformation of the side chain. The conformation of N_τMe-TRH is probably like that of TRH, because their ^1H NMR spectra are similar in D_2O and DMSO (Donzel 1974; Bellocq 1976). On the contrary, N_πMe-TRH does not show any of the spectral features of TRH in DMSO, but its ^1H NMR spectrum is a sum of the spectra of its components (Donzel 1974). The existence of the *cis* form in N_πMe-TRH is not known, but the doubly methylated N_τ,N_π-diMeTRH is at least 90 % *trans* (Liakopoulou-Kyriakides 1979).

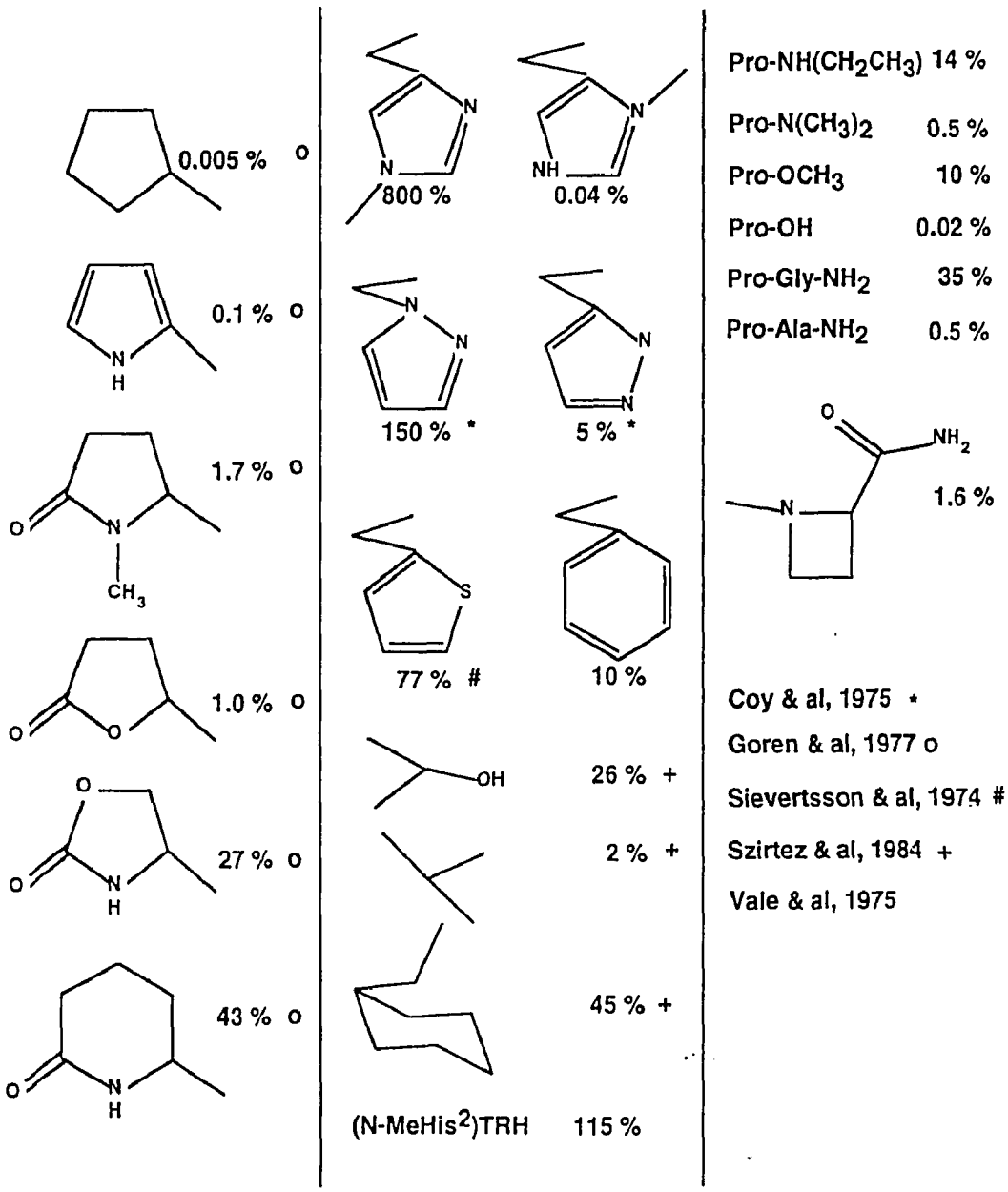
2.2 Structure-activity data of TRH

Hundreds of analogs of TRH have been sythesized and tested for biological activity, and surprisingly, only one of them has higher affinity for the receptor than TRH itself, N_τMe-TRH (Vale 1973; Coy 1975; Goren 1977; Morgan 1979; Nutt 1979; Marshall 1982; Szirtes 1984; Eckle 1985). Methylation of the imidazole π–nitrogen produces an nearly inactive analog (Vale 1973). See Table 2.2 on the next page for a list of TRH analogs.

The binding of TRH to the pituitary receptor is sensitive to pH (Perlman 1992). In acidic conditions the affinity is low and it reaches a plateau around pH = 8. The acidity range of the change agrees with the pKa of His, and suggests

Table 2.2. Structure-activity data on TRH analogs. Numbers shown are fractions of TRH activity.





that TRH binds to the TRHR in the unprotonated form. The effect is not likely to be caused by changes in the receptor, because there are no histidines in the extracellular loops or in the helices. Also, the negatively charged aspartates that could be protonated at pH < 7 were shown not to influence the affinity in a significant manner (Perlaman 1992).

A special property of the TRH system is that all known peptide analogs show the same efficacy. Every analog that binds also activates the receptor fully. It seems that the structural elements needed for binding and activation are the same.

It is not possible to study efficiently the binding and activation of TRH-TRHR complex from the existing structure-activity data. All the three critical aspects of building a pharmacophore are lacking: the antagonist data, the conformational data and the biological data. First, the data on antagonists: there are no partial agonist to TRH, and no peptide antagonists. Benzodiazepines act as competitive antagonists to TRHR, but even the best of them, midazolam, has low micromolar affinity to TRHR, while TRH itself binds with K_d of 1 nM. This means that with the available analogs there is no way to separate structural elements needed for binding from those required for activation, or telling, for example, whether the backbone has a role in TRH action. Neither can one tell, whether the benzodiazepines bind in the same site as TRH. Second, the goal in the synthesis of TRH-analogs has mainly been to prevent enzymatic degradation (Yaron 1993) or to separate the hormonal activity from the central nervous system activity (Szirtes 1984), and not to address structural questions. Accordingly, the modifications have concentrated on the terminal rings. The conformational behavior of TRH has not been the main issue; the first conformationally constrained analogs have been sythesized only recently (Mapelli 1989, Mapelli

1990, Olson 1995), and very few analogs have been studied spectroscopically. Even for the puzzling pair of analogs [τ-Me-His^2]-TRH and [π-Me-His^2]-TRH (See Table 2.2) it is not possible to say, whether the observed large difference in activity is caused by different conformational preferences or elimination of an interaction site with the receptor. There have been several molecular mechanics attempts to bridge conformational data of TRH analogs to measured activities (Belle 1972; Blagdon 1973; Robson 1979; Ward 1987; Garduño-Juarez 1990). They all suffer from the third, and most confusing problem in the structure-activity data: the majority of biological data on TRH analogs comes from whole animal studies, that do not differentiate between affinity and efficacy, nor pay attention to bioavailability or rate of metabolism (Vale 1973).

The existing data on analogs, taken as it is, can be condensed to one condition per residue: 1) pGlu is essential to binding both through its carbonyl and NH-groups, 2) His is permissive and seems to affect the conformation of TRH, and 3) ProNH_2 needs both its CO and NH.

The special advantages that can be achieved from detailed studies are exemplified by 2,4-methanoproline-TRH (MePro-TRH). See Figure 2.3 on the next page for the structure. In contrast to old structure-activity data, that report the biological effect as a percentage activity in whole animals, and the structure as a chemical formula, this analog has been studied thoroughly with both physical and biological methods. The conformational data are from ^1H and ^{13}C 2D-NMR studies (Mapelli 1990), and the binding and activation studies have been done in transfected COS cells (M. Gershengorn, unpublished results). The CH_2 bridge is a structurally small and chemically inert modification, that yet has three major consequences: The bridge constrains MePro-TRH to be at least 95 %

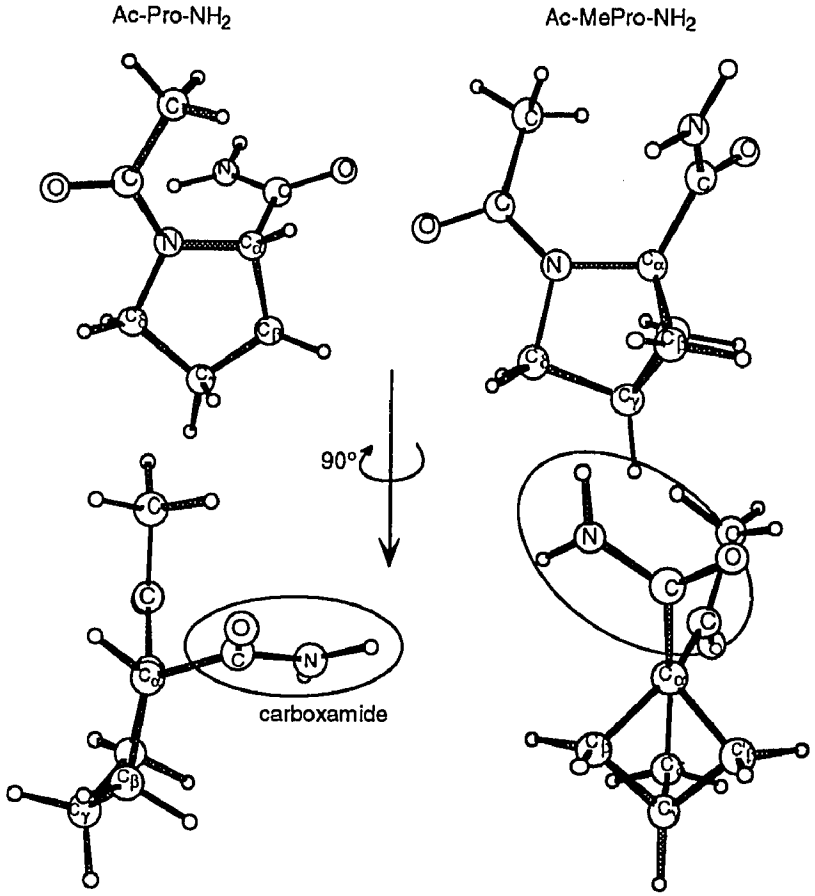


Figure 2.3. Comparison of the structures of Ac-Pro-NH2 and Ac-MePro-NH_2. In the plane of the proline ring the additional methylene group is hardly visible (upper row). The greatly increased bulk of the bridged compound becomes evident from a perpendicular view (lower row). The structures are positioned so that proline ring points to the left, N and Cα of proline overlap, and also the acetyl carbonyl carbon and oxygen overlap maximally. Notice also, how the carboxy termini point to different directions.

trans at the second peptide bond, in contrast to about 75 % in TRH. The position of the carboxy tail differs in MePro-TRH: ψ_3 = -29° in MePro-TRH compared to 150° in TRH. Otherwise both TRHs are in an extended conformation. The volume of the bicyclic ring is much larger than of natural proline, as can be seen inFigure 2.3. The biological behavior of MePro-TRH differs from TRH behavior. MePro-TRH binds weakly to TRHR, with a K_d 11000 times larger than that of TRH. Nevertheless, it produces the full effect of TRH, measured as breakdown of phosphoinositides. Full stimulation of the receptor by MePro-TRH means that the activation step is unhindered.

Three hypotheses can be proposed to explain the reduced binding affinity. One possibility is that the receptor binds preferentially to *cis*-TRH, and the low K_d of MePro-TRH is indicative of diminished *cis*-population. It can also be that the large MePro ring cannot fit inside the binding pocket of the TRH receptor without major rearrangements of the protein. Or, the exact position of the C-terminus is critical to good binding. Each of the different hypotheses suggests specific tests. The first hypothesis could be tested by other analogs that have strong conformational preferences at ω_2. The second hypothesis can be tested by mutating large residues in the binding site to smaller ones. A model of the binding complex can help to identify the residues to mutate. The third hypothesis suggests a Hbond from the carboxy terminus to the receptor.

Study of constrained analogs, with clear conformational preferences, is the best way to deduce the physical conditions for binding and activation. Without all the data, MePro-TRH would only be a bad analog, with 0.01 % of TRH activity.

2.3 Physiological profile of TRH.

TRH has a role in all the three routes of cell-to-cell signaling (Metcalf 1989). It is best known as a hypothalamic hormone, but it also acts as a paracrine signaling agent for example in retina and pancreas, and as a neuromodulator in the central nervous system. TRH is ubiquitous in all the animal kingdom. The endogeneous functions of TRH vary and are mostly unknown (Metcalf 1989). Only the hormonal role of TRH is well defined and understood (Guillemin 1971; Vale 1973; Metcalf 1989). TRH, as the full name thyrotropin releasing hormone says, is a releasing hormone, secreted by the hypothalamus, and acting on the anterior pituitary, where it stimulates the release of thyrotropin (TSH, thyroid stimulating hormone) and prolactin (PRL). Its receptor has been isolated and sequenced (Straub 1991), and TRH release is known to be regulated by thyroid hormones. All biological data related to this work are of TRH acting as a hormone, and concentrate on TRH activating its receptor in the pituitary cells.

There are several reports on endogeneous TRH-like peptides that cross-react with anti-TRH antibodies. They all are of the form pGlu-Xxx-Pro-NH_2, and they have been isolated from tissues that do not have TRH itself, like the male reproductive tract (Cockle 1989; Khan 1992).

TRH is synthesized in the hypothalamus as a long precursor with seven copies of -Gln-His-Pro-Gly-, and is sequentially cleaved into the mature tripeptide (Metcalf 1989). Both the C and N termini of TRH are blocked, which makes it resistant to many common peptidases and proteases. Nevertheless, TRH is rapidly broken down in the body, mainly by pyroglutamase II and proline endopeptidase. The former is highly specific to TRH, the latter is selective for X-Pro

bonds. Of the metabolites, at least cyclo(His-Pro) has its own physiological effects, some of them similar (e.g. inhibition of cholesterol synthesis), some different (*e.g.* inhibition of dopamine uptake) from the effects of TRH (Prasad 1984).

2.4 Sequence related data on TRH receptor

The only structural information available of the TRH-receptor is the primary sequence (Straub 1991). Because of the low abundancy of the protein, no direct physical studies are possible. The sequence is known from three species: man, mouse and rat. For the first 375 amino acids they are ~95 % identical, and none of the differences are found in the putative transmembrane regions. See Figure 2.4 for the mouse sequence.

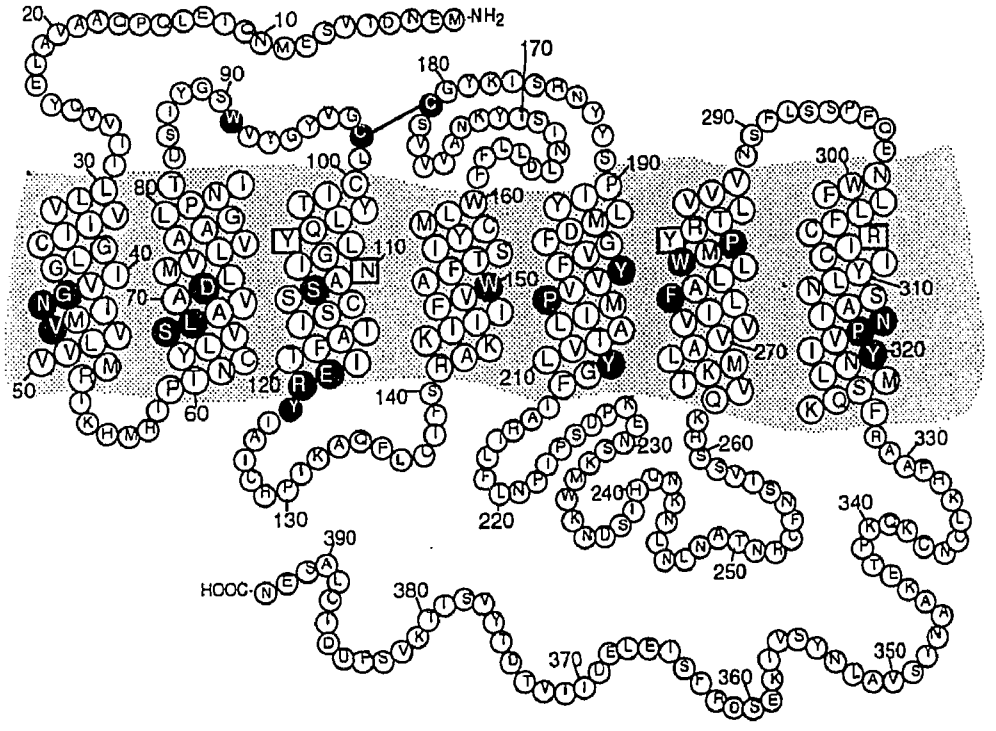


Figure 2.4. The TRHR sequence. Residues that are conserved among most GPCRs are black, and residues that have been shown to interact with TRH are in boxes. They grey area represents the membrane.

The TRHR sequences align well with the rest of the G-protein coupled receptors (Probst 1992), but show some exceptional features. Data on the most important similarities and differences between TRHR and other GPCRs are collected in Table 2.3.

	cons. present	cons. absent	a rare residue in TRHR / reports of mutating the same site
hlx1	G42 (1.49), N43 (1.50), V46 (1.53)		
hlx2	S66 (2.50), L67 (2.51), D71 (2.55)		
hlx3	C98 (3.25), S112 (3.39), 122-124 ERY (3.49-3.51)		Q105(3.32) /(Strader 1987) N110(3.37)
hlx4	W150 (4.50)	P(4.60) subst by W	
hlx5	Y200 (5.47), P203 (5.50)		D195(5.42)
hlx6	F275(6.44), W279 (6.48), P281(6.50)	C(6.47) subst by M	R283(6.52) / (Fong 1994), (Choudhary 1995) Y282(6.51) / (Schambye 1995)
hlx7	N316(7.49), P317 (7.50), Y320(7.53)		R306(7.39) / (Wess 1992), (Suryanarayana 1993)

Table 2.3. Comparison of TRHR to other GPCRs, highlighting similarities and differences. Each residue is identified by its sequential number, and by a systematic number, that is relative to the most conserved residue (X.50) in the helix (Ballesteros 1995).

There are two potential N-glycosylation sites in the N-terminus, several serines and threonines on the intracellular side that could undergo phosphorylation, and the cysteines in the C-terminus known to be palmitoylated in some receptors, linking the receptor to the membrane (O'Dowd 1989) are present. There are also data on the existence of a disulfide bridge between C98 and C179 in extracellular loops 2 and 3 (Perlman, unpublished data).

2.5 Biological aspects of the TRH-receptor

When TRH binds to its receptor, a long cascade of activation reactions starts, which ends in intracellular release of Ca^{2+} and phosphorylation of proteins. A schematic description of the steps that follow the binding event is shown in Figure 2.5.

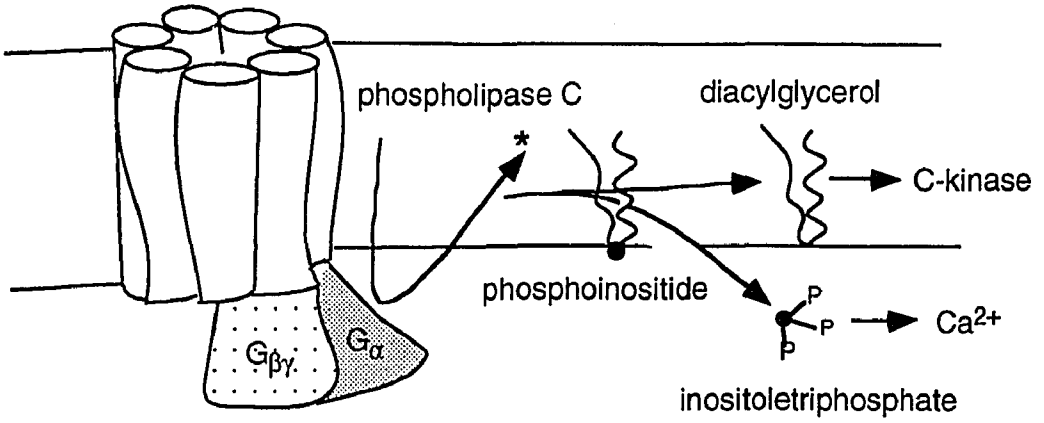


Figure 2.5. Signaling cascade of TRHR. The activated TRH receptor binds to the G-protein, which dissociates into two parts, G_α and $G_{\beta\gamma}$. G_α activates phospholipase C, which breaks phosphoinositide into diacylglycerol (DAG) and inositol triphosphate (IP_3). DAG activates C kinase and leads to influx of extracellular Ca^{2+}, and IP_3 causes release of Ca^{2+}.

Extended ternary complex model. The general understanding of GPCR function has developed in the last few years from a view where the ligand induces an activating change in the receptor to a dynamic picture with different receptor conformations in equilibrium with each other. This means that a second equilibrium between the inactive receptor R and the activated receptor R* had to be added to the classical ternary complex model (De Lean 1980), that includes equilibrium coupling of ligand L, G-protein G and receptor R to LRG complex as the active form of the receptor:

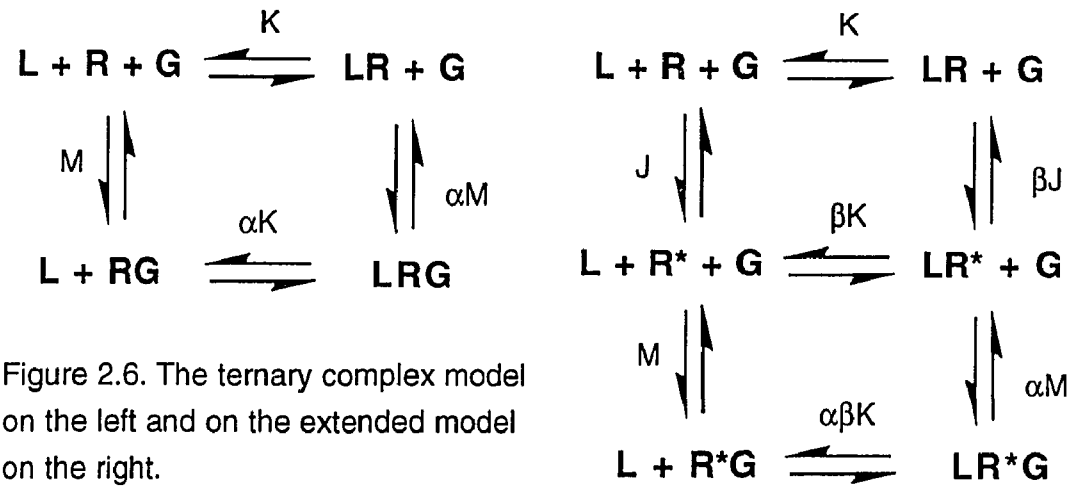


Figure 2.6. The ternary complex model on the left and on the extended model on the right.

The essential feature of the ternary complex model is the allosteric effect the binding of the ligand or the G-protein has on the receptor. Formation of the RG complex enhances the binding of the ligand, and formation of the LR complex enhances the formation of the active LRG complex. In this model the efficacy of a molecule expressed as α in the diagram on the left, is interpreted as its ability to stabilize or destabilize the formation of LRG complex, and is thus directly related to the allosteric effect.

The ternary complex model had to be extended because of data about constitutive activation (Samama 1993). Certain mutations in the third intracellular loop of the β_2-adrenergic receptor produce constitutive activity in the absence of agonist, which can be measured as increase in the basal activity of the receptor. The activity can be further increased by the addition of agonists, but surprisingly it is decreased by some ligands, that are termed inverse agonists. Constitutive activation can be envisioned structurally as removal of a constraining interaction, analogously to rhodopsin where elimination of a salt bridge between helices 3 and 7 causes activity without light (Rao 1994). The physical

structure of the different states of the receptor is not known. The mutations in β_2-adrenergic receptor that caused constitutive activation also increased the affinity of the receptor for agonists, in correlation to their intrinsic activity at the wild type receptor (Samama 1993). The mutations increased the potency of agonists to simulate adenylyl cyclase, and they increased the intrinsic activity of partial agonists; the affinity of antagonists to the mutant receptors did not change. These effects, especially the observed connection between affinity, K, and efficacy, α, and the fact the the changes in affinity are independent of the presence of G-protein, cannot be satisfactorily modeled by the ternary complex model in the left panel of Figure 2.6, but an additional equilibrium between an inactive form of the receptor R and an active form R* has to be added to the model (Figure 2.6, right panel). In the extended ternary model R* can exist in the absence of the ligand, and G-protein can only couple to the activated receptor R*. The efficacy $\alpha\beta$ of a ligand to induce activation is dependent on its stabilization of R* over R, expressed as β in the Figure 2.6 on the right, and on its ability to assist the binding of G-protein to the LR* complex, given by α. The factor β has a positive value for agonists, promoting the formation of LR*; $\beta = 1$ for neutral antagonists; and $\beta < 1$ for inverse agonists, thus stabilizing the inactive form LR. According to this model, the change in affinity of the receptor is directly governed by the factor J, and thus correctly independent of the G-protein. Also, a relationship can be derived between affinity K and β, that is part of the efficacy of a molecule, and an increase of J will increase the apparenty affinity of a ligand relative to its efficacy. Additional support for the existence of several states of the receptor was given by data from transgenic mice (Bond 1995). The mice overexpress myocardial β_2-adrenergic receptors and show maximal activity in the absence of agonist. The activation cannot be blocked by a neutral antagonist but it

can be diminished by inverse agonists. This shows convincinly that activity is dependent on the state of the receptor, not simply its occupancy, and that the ligands affect the equilibrium between the inactive and active conformations.

Data on the TRH receptor. The ability to express functional wild type TRH receptor (Straub 1991) and several mutants in monkey kidney (COS-1) cells (Perlman 1992) has presented an opportunity to study in detail the role of individual residues. Binding, activation and desensitization have been studied. Direct studies of receptor number and location in the cell are complicated by the lack of specific antibodies.

Two point mutations of TRHR, D71A and R283A (Perlman 1992; Perlman 1995), produce totally inactive receptors. Mutation of C335 to a stop codon, on the other hand, produces a constitutively active receptor (Matus-Leibovitch 1995). No mechanistic explanation exists for these observations.

All four aspartates in or close to the putative helices and extracellular loops have been studied by mutations to alanine: D71, D85, D165, D195. Only D71A affected binding by an 8 fold reduction in affinity. Mutation of D71 to A also rendered the receptor inactive (Perlman 1992). Another residue critical for activity is R283 in the sixth helix, and its mutation to A also reduces affinity about 2000 fold (Perlman 1995).

TRHR has been shown to interact with G_q and G_{11}, both by coexpression of the receptor and either one of the G_α's, and by using G_α specific antibodies (Aragay 1992; Hsieh 1992; Lipinsky 1992; Quick 1994). There are also reports about TRHR coupling to G_s, but the evidence is not unequivocal (Paulssen 1992; Quick 1994; Peña 1995). There are no data on effects mediated by $G_{\beta\gamma}$. The coupling to the PLC-pathway is delivered through the third intracellular

loop, but the distal part of the carboxy terminus is not needed for the activation (Nussenzveig 1993).

TRHR shows ligand-induced rapid desensitization (Torjesen 1988), most probably by phosphorylation, in a cell type dependent fashion (Falck-Pedersen 1994). Both agonist-induced and agonist-independent internalization have been observed with fluorescence microscopy (Ashworth 1995). The internalization is inhibited by mutation of the two putative C-terminal palmitoylation sites, cysteines C335 and C337 (Nussenzveig 1993). Part of the receptors is recycled to the cell surface, and it is not known whether internalization leads to direct degradation. Exposure of GH_3 cells to TRH also leads to decrease in amount of TRHR mRNA (Yang 1993), apparently through the PLC pathway.

2.6 Binding studies

Several residues on the extracellular half of the helix bundle have been shown to be important for binding of TRH: Y106, N110 in helix 3, Y282 and R283 in helix 6, and R306 in helix 7 (Perlman 1992; Perlman 1993; Perlman 1994a; Perlman 1994b; Perlman 1995). Whenever it has been possible, the mutant receptors have been tested both for their affinity and activation by the ligand. In some cases the affinity is too low to conduct binding experiments. In all but two cases with both affinity and potency data, changes in potency reflect changes in affinity, suggesting that those mutations did not change the efficacy of receptor activation. Based on this, EC_{50}s can be assumed to reflect K_ds. The testing of proposed interactions has involved reciprocal mutations of the receptor and modifications of the ligand. For example, the existence of a hydrogen bond between Y106 and the side chain carbonyl of pGlu (See Figure 2.7), that

was proposed by our initial modeling study, was tested both by mutation Y106F and by elimination of the carbonyl group from pGlu. These are the minimal changes to eliminate the functional groups needed for hydrogen bonding, and they should not alter the structure drastically. Individually, both structurally small changes cause a 10^5 fold reduction of EC_{50}, whereas testing the analog on the mutant has only a 16 fold additional effect. If the effect of eliminating the Y106 hydroxyl group were unconnected to the effect of eliminating the side chain carbonyl group of pGlu, a 10^{10} fold effect would be expected. The non-additivity is consistent, although not a proof of a direct ligand-receptor contact. It does, however, seem unlikely, that such small changes could cause equally large indirect effects. A major conformational change of the receptor is ruled out by the fact that binding of the specific antagonist chlordiazepoxide is identical in the wild type receptor and in the Y106F mutant. The non-additivity has been used as a criterion for direct interactions also for contacts between N110 and NH of pGlu, R306 and the carboxy terminus of TRH, and between Y282 and His. In these cases unifunctional modifications are not possible, and the effects are not as dramatic as for Y106-pGlu. Additional tests with analogs modified in other parts in the ligand show additive effects. The effect of N110A mutation and modification of pGlu to a lactone are smallest, only 30 fold. The effects of Y282A and R306A modifications, tested with Val^2-TRH and Pyr^3-TRH (Pyr = pyrrolidine) are in the order of 10^5 and 2000, respectively. In contrast to the Y106F mutation, Y282F caused only a 5 fold reduction in binding affinity and potency, and the aromatic ring seems to be the important element of this residue. Accordingly, contacts between TRH and Y106, Y282 and R306 have been used as guides for constructing a ligand-receptor complex, as is described in Chapter 6. The four proposed ligand-receptor contacts are shown in Figure 2.7.

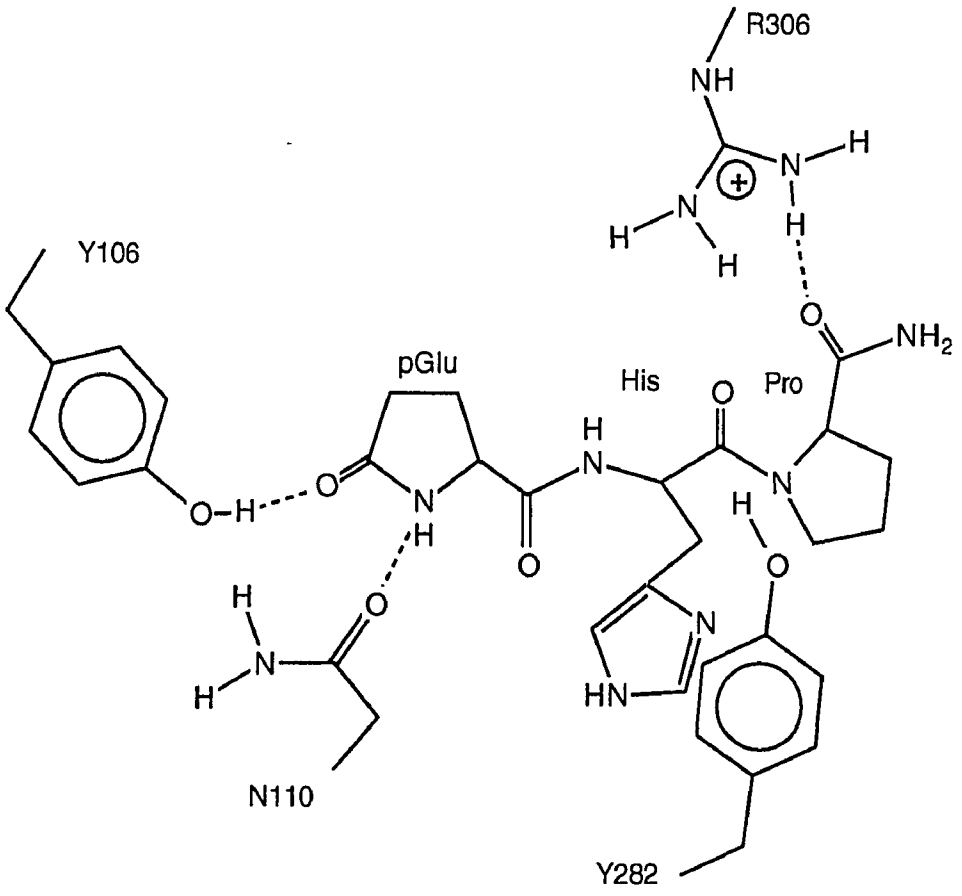


Figure 2.7. The binding interactions between TRH and its receptor. Y106 and N110 form Hbonds to pGlu, Y282 is close to His and R306 forms a Hbond to the terminal carboxylamide.

3 An ab initio study of proline conformations

3.1 Introduction

Proline, with its cyclic side chain, is unique among the common amino acids. The ring structure has three major consequences: First, with C_δ bound to the peptide N there is no amide hydrogen, and no Hbond donation. Second, the ring restricts ϕ-values close to -70°, and proline has only one free backbone torsional angle, ψ. The ring also restricts the conformational space available to the residue before proline (Zimmerman 1977). Third, peptide bonds before proline do not show strong preference for the *trans* conformation in the way other amino acids do, because both *cis* and *trans* forms are equally crowded. *Cis* forms are observed 15-70 % in zwitterionic X-Pro dipeptides in aqueous solution, depending on the nature of X (Grathwohl 1976). All three properties are important in secondary structure formation in proteins. See Figure 3.1 for definitions of the angles.

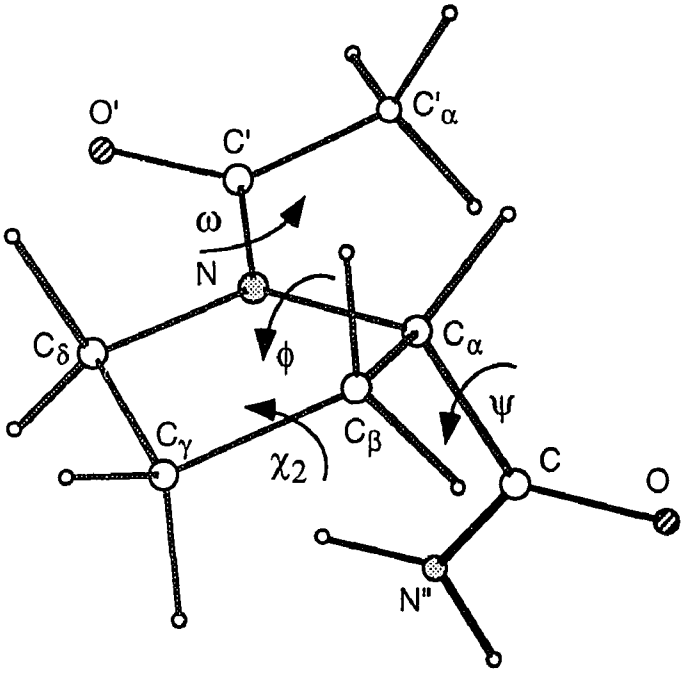


Figure 3.1. Torsional angles of proline. Nitrogens are stippled and oxygens are striped.

Prolines are often conserved in related proteins, which suggests that the special properties of proline are structurally and functionally important (Yaron 1993). Prolines are found disproportionally often in the beginning of α-helices, and they are not present in the last turn of a helix. In the middle of helices they disrupt the Hbonding pattern and kink the helix, which is assumed to have a role in signaling through transmembrane proteins. Prolines define the conformation of many small bioactive peptides (Jorgensen 1988; Schmidt 1993). There is also a whole family of enzymes that specialize in *cis-trans* isomerization of X-Pro bonds in proteins (Fischer 1984; Stamnes 1990), with the isomerization as the rate limiting step in folding (Yaron 1993).

The main free dihedral angle of Pro is ψ. It has three stable states, that define conformations called α_R, P_{II} and C_7, because of their resemblance to right handed α-helix, polyproline$_{II}$ and seven membered ring, respectively. All three states can in principle exist either in *cis* or *trans* peptide bond, giving six possible forms. All the three *trans*-forms are found experimentally in different conditions: Ac-Pro-NHMe crystallizes in α_R conformation (Matsuzaki 1971). In aqueous solution it is α_R or P_{II}, and in nonpolar solvents it is mainly in the C_7 form (Tsuboi 1959; Madison 1980). The less prevalent *cis*-forms have not been studied experimentally in detail. The fact that the predominant form varies from one environment to another suggests that the conformers are close to each other in energy without large energy barriers between them. Some of the conformational data on Pro are contradictory, and different methods give different results: The C_7 form in water is reported either as absent by NMR measurements (Madison 1980) or present as a minor conformer by IR (Higashijima 1977). The same studies show strong solvent dependance of *cis* population, from 4% in CCl_4 to 26 % in water

(Higashijima 1977) or practically no dependence at all (Madison 1980). The *cis-trans* ratios of model peptides N-acetylamide and N-formylamide show little or no solvent dependence, neither experimentally nor computationally (Jorgensen 1988; Radzicka 1988). The goal of this study is to address the *cis-trans* energy difference in different environments and to describe the conformations of *cis*-Pro.

3.2 Methods

The geometry of acetylprolineamides is determined by two free bonds, the peptide bond and the bond between C_α and carbonyl carbon, and the puckering of the ring. These three degrees of freedom are described by dihedral angles ω ($C\alpha'C'NC_\alpha$), $\psi(NC_\alpha CN'')$ and $\chi_2(C_\alpha C_\beta C_\gamma C_\delta)$, respectively. See Figure 3.1. The peptide bond can be in two forms, *cis* or *trans* ($\omega = 0°/180°$), and the C_α–C in three possible forms: α_R ($\psi = -20°$), C_7 ($\psi = 80°$) and P_{II} ($\psi = 160°$). In the pyrrolidine ring C_γ can pucker towards or away from C_α: endo ($\chi_2 = 35°$) or exo ($\chi_2 = -35°$), respectively. All three conformational variables were studied, and they are combined in the name of each structure, *e.g.* *cis*-α_R-exo, *trans*-C_7-endo. A total of twelve forms, modified from the crystal geometry (Matsuzaki 1971) were used as starting structures for optimization.

The energetics of Ac-Pro-NH_2 was calculated *ab initio* with Gaussian 92 (Frisch 1992). Structures were fully optimized at HF/6-31G* level both in vacuum and with a self consistent reaction field representing aqueous solution, and zero point energies (ZPE) were calculated for the optimized structures. MP2 single point energies were calculated at 6-31G* level in vacuum. A scan of the ψ angle was conducted at HF/6-31G* for endo and exo puckered forms of both *cis* and *trans* Ac-Pro-NH_2. All 70 structures were fully optimized.

3.3 Results and discussion

3.3.1 Gas phase minima

The twelve starting structures converged to seven stable minima in gas phase: *trans*-α_R-endo, *trans*-C_7-endo and exo, *cis*-α_R-endo and exo, *cis*-P_{II}-endo and exo. The unstable structures relaxed by rotation of ψ. See Figure 3.2 on the next page. The apparent reason for the instability is close repulsive contacts between the acetyl and amide groups. The optimized structures differ from each other only regarding the conformational parameters studied, otherwise the structures are practically identical: bond distances are similar to 0.005 Å and bond angles to 1.4°. The *trans*- C_7 structures have internal Hbonds between the acetyl CO and amide NH, and they are lowest in energy in gas phase. The *cis*-endo structures are stabilized by 2.5 - 4.0 kJ/mol compared to the corresponding exo structures, and *trans*-endo by 8 kJ/mol. See Table 3.1 for relative energies and Table 3.2 for a more complete energy listing.

	HF+ZPE 6-31G*, vac	HF+ZPE 6-31G*, aq	MP2+ZPE 6-31G*, vac
trans-α_R-exo	16.33 (0.2)	4.52 (12.9)	22.17 (0.0)
trans-C_7-exo	8.59 (3.1)	8.28 (2.9)	8.16 (3.7)
trans-C_7-endo	0.00 (95.4)	0.00 (78.1)	0.00 (95.9)
cis-α_R-exo	15.50 (0.2)	18.87 (0.0)	19.02 (0.0)
cis-α_R-endo	11.50 (1.0)	16.05 (0.1)	14.62 (0.3)
cis-P_{II}-exo	23.83 (0.0)	10.44 (1.2)	28.68 (0.0)
cis-P_{II}-endo	20.12 (0.0)	7.07 (4.7)	24.76 (0.0)

Table 3.1. Relative energies (kJ/mol) of stable conformations of Ac-Pro-NH_2. The number in parenthesis is the fractional population, vac stands for vacuum and aq for the reaction field.

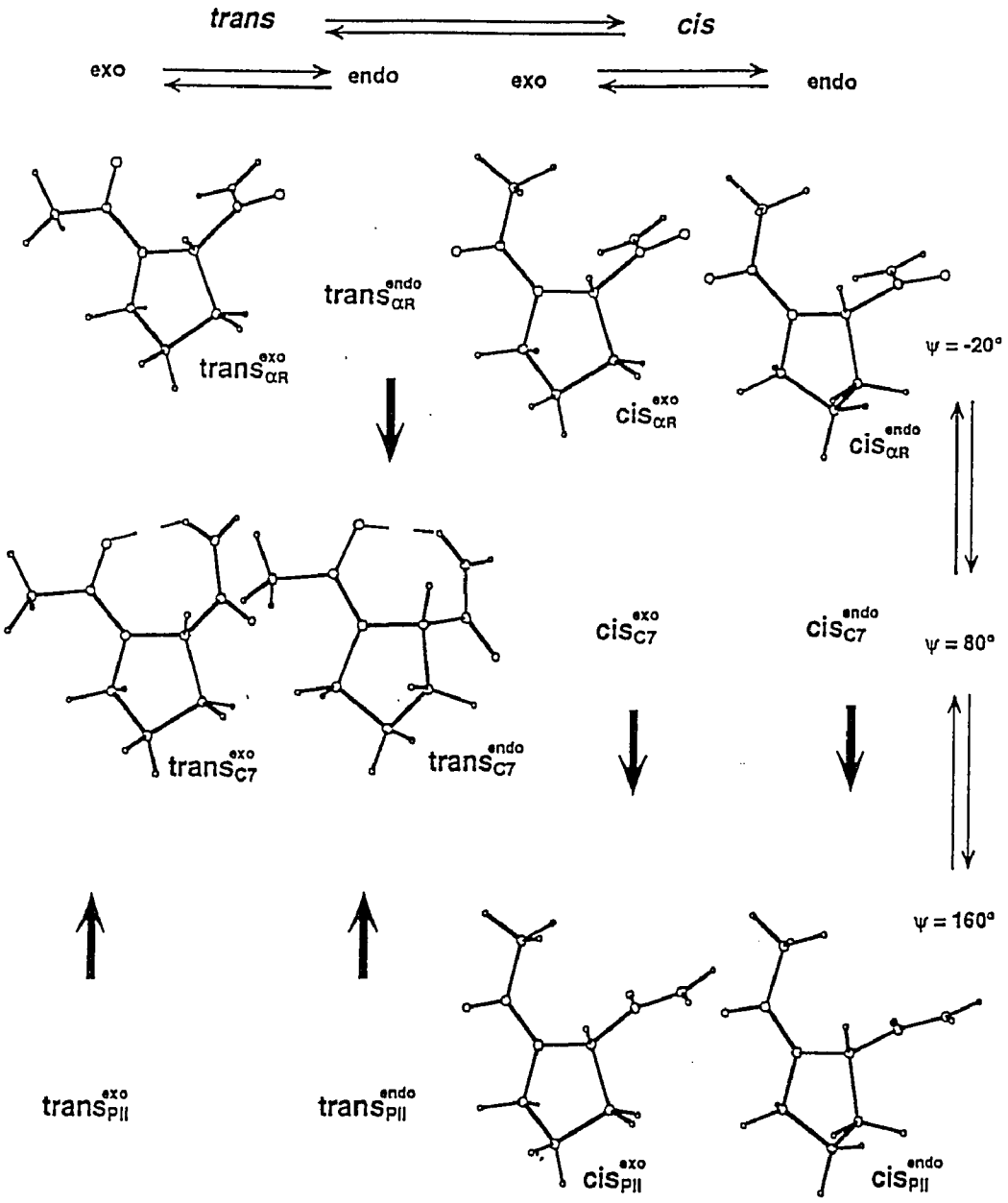


Figure 3.2. Stable conformers of Ac-Pro-NH_2 in vacuum at HF/6-31G* level. The three structures on the left are in *trans* conformation and the four on the right are *cis*. The endo and exo puckering of the structures is marked. In the right margin of the picture the values for the ψ angle are shown. The five names without a structure in the picture represent combinations of torsional angles that did not produce stable structures, and the arrows show where those starting structures optimized.

		angles		μ			E	
		vac	aq	vac	aq		vac	aq
t-α_R-exo	ω=	-171.16/	-174.68	6.36	8.03	HF/6-31G*=	-530.7261/	-530.7327
	ϕ=	-69.37/	-68.72			ZPE/6-31G*=	0.2107/	0.2108
	ψ=	-21.22/	-27.71			MP2/6-31G*=	-532.2985/	-532.3118
	χ2=	37.63/	35.49					
t-C_7-exo	ω=	-175.17/	-173.39	3.74	4.79	HF/6-31G*=	-530.7299/	-530.7322
	ϕ=	-81.34/	-80.87			ZPE/6-31G*=	0.2116/	0.2115
	ψ=	86.39/	66.80			MP2/6-31G*=	-532.3052/	-532.3092
	χ2=	31.77/	28.55					
t-C_7-endo	ω=	-172.92/	-171.70	3.62	4.62	HF/6-31G*=	-530.7329/	-530.7351
	ϕ=	-85.76/	-83.54			ZPE/6-31G*=	0.2115/	0.2114
	ψ=	76.56/	63.30			MP2/6-31G*=	-532.3084/	-532.3121
	χ2=	-37.76/	-37.25					
t-P_{II}-endo	ω=	/	178.45		5.78	HF/6-31G=		-530.7305
	ϕ=	/	-72.71			ZPE/6-31G=		0.2107
	ψ=	/	155.42			MP2/6-31G*=		-532.3056
	χ2=	/	-25.88					
c-α_R-exo	ω=	7.15/	4.24	2.05	3.20	HF/6-31G*=	-530.7260/	-530.7268
	ϕ=	-75.65/	-68.72			ZPE/6-31G*=	0.2108/	0.2106
	ψ=	-23.46/	-32.04			MP2/6-31G*=	-532.3000/	-532.3022
	χ2=	36.73/	35.35					
c-α_R-endo	ω=	10.64/	5.50	1.30	1.95	HF/6-31G*=	-530.7277/	-530.7281
	ϕ=	-90.16/	-83.40			ZPE/6-31G*=	0.2109/	0.2108
	ψ=	-6.99/	-14.89			MP2/6-31G*=	-532.3020	-532.3029
	χ2=	-38.03/	-36.00					
c-P_{II}-exo	ω=	-4.81/	-6.62	7.08	8.65	HF/6-31G*=	-530.7224/	-530.7302
	ϕ=	-58.09/	-61.76			ZPE/6-31G*=	0.2105/	0.2107
	ψ=	158.21/	172.04			MP2/6-31G*=	-532.2961/	-532.3124
	χ2=	35.84/	33.29					
c-P_{II}-endo	ω=	-3.15/	-4.46	7.03	8.59	HF/6-31G*=	-530.7238/	-530.7314
	ϕ=	-72.16/	-72.57			ZPE/6-31G*=	0.2105/	0.2108
	ψ=	161.71/	175.31			MP2/6-31G*=	-532.2976/	-532.3139
	χ2=	-37.06/	-36.39					

Table 3.2. Structures and energies of the optimized AcProNH$_2$. Energies are in atomic units, angles in degrees and dipole moments μ in Debyes. In the names of the conformers, t means *trans* and c *cis*, vac means vacuum and aq the reaction field.

All relative energy differences in vacuum are larger at MP2 level than at HF level. This is in accordance with well documented MP2 overstabilization of H-bonded conformations relative to non-Hbonded ones at normal sized basis sets (Hu 1993; Cramer 1994). The correct energetics can be reached with very extensive methods and large basis set, *e.g.* triple excitation coupled clusters with a tailored triple zeta basis set (CCSD(T) and TZ2P+f) (Hu 1993). That kind of calculations are not feasible for a system as large as AcProNH$_2$. Accordingly, energy comparisons at HF level seem more reliable.

The most stable structure in vacuum is *trans*-C_7-endo. See Table 3.1 for HF+ZPE/vac column. *Trans*-C_7-endo is the only structure with relative population over 5 % in vacuum. There is no gas phase data on Ac-Pro-NH$_2$, so the closest comparison is to experimental data from nonpolar solvents: the C_7 structure is the predominant one in chloroform, but minor amounts of other forms are also observed (Higashijima, 1977; Madison 1980; Tsuboi 1959). It is especially worth noticing that in contrast to earlier calculations (Hodes 1979; Madison 1980) *trans*-P_{II} structures were unstable, in spite of repeated optimization trials with slightly different starting structures and calculation of forces at every point.

The experimental data on *cis*-population of Ac-Pro-NHMe in nonpolar solvents is contradictory, both very low (0-4%) (Higashijima 1977) and intermediate populations (20%) (Deslauriers 1973; Madison 1980) have been inferred from NMR experiments. The reason for different results might be that intensities from ^{13}C NMR spectra (Deslauriers 1973; Madison 1980) are a much less reliable source for relative amounts than ^{1}H-intensities (Higashijima 1977), as there are several competing relaxation pathways. The *cis*-peaks are small, and thus prone to misreading. The calculated *cis-trans* energy difference is 11.5 kJ/mol, which corresponds to 1 % *cis*, in agreement with Higashijima & al. (Higashijima 1977).

The α_R-endo is the most stable *cis*-conformer. In vacuum the computational results are in good agreement with the available experimental data.

Crystal effect. The structure *trans*-α_R-endo in vacuum calculations is practically identical to the crystal structure (Matsuzaki 1971). The rms deviation between the heavy atoms from the crystal structure is 0.12 Å, when C_γ that puckers differently and the C-terminal methyl group of the crystal are excluded. The effect of the crystal environment is best estimated by comparing *trans*-C_7-exo, the minimum energy structure in vacuum, to the crystal-like structure. In both structures, the acetyl carbonyl and the amide hydrogen are involved in hydrogen bonding. *Trans*-C_7-exo forms one internal Hbond between the two, and in crystal *trans*-α_R-exo the same groups form H-bonds to two neighboring molecules. Ac-Pro-NHMe in C_7-form could not form a network of intermolecular interactions, as it does not have any free proton donors left. The calculated energy difference between the two forms is 16.33 kJ/mol. The extra stabilization has to come from the packing in crystal. Studying the crystal it is noticed that in the hydrogen bonding chains of molecules the dipole moments are oriented along the chain axis, and the neighboring chains are antiparallel (Matsuzaki 1971). This clearly stabilizes the crystal. Also, the calculated and crystallized molecules differ in the C-terminus, crystal being a N-methyl amide and the calculated structure an amide. N-methyl groups of the crystal come close to each other ($r_{C\text{-}C}$ = 3.8 Å) and this favorable van der Waals interaction could also stabilize the α_R form in the crystal.

It is instructive to notice, how the nonisotropic environment of crystal is able to provide over 16 kJ/mol energy to the molecule. A more complex environment, like a protein, can be estimated to stabilize different peptide conformers by much larger amounts of energy.

3.3.2 Solution minima

Reoptimizing the previously stable structures with a reaction field representing water (Wong 1991a; Wong 1991b; Frisch 1992; Wong 1992a; Wong 1992b) did not cause any drastic structural changes. Solvation stabilized all structures and changed the relative energies of the conformers. This was expected, as the conformers differ from each other mainly by the relative orientation of acetyl and amide groups, and thus have very different dipole moments. See Table 3.2. The previously missing *trans*-P_{II} was stable at 6-31G but not at 6-31G* level. Disturbingly, the C_7 structures are still lowest in energy. This is clearly wrong: there should be little, if any of the internally Hbonded *trans*-C_7 in water. The solvent model in Gaussian 92 is clearly not able to stabilize the open structures relative to the C_7. If the *trans*-C_7 is considered an unsuitable case for the method, and based on experimental data eliminated from the comparison (Higashijima 1977; Madison 1980), the energetics agree with the rest of experimental data. One *trans* form is present, *trans*-α_R-exo. It is the lowest energy structure, representing 69 % of the population. The energetic order of the *cis* structures is changed by solvation, and *cis*-P_{II} is now more stable than *cis*-α_R. There are no experimental data on the *cis* conformers. The *cis-trans* energy difference is reduced to 2.5 kJ/mol, which corresponds to 25 % of the lowest energy *cis* conformer, and 32 % of all *cis* conformers combined. This is very close to the experimentally observed 25 % (Deslauriers 1973; Higashijima1977; Madison 1980).

Madison's CD spectroscopic results tell (Madison 1980) that two AcPro-NH_2 conformers regarding ψ_3 should be present in water : P_{II} and α_R. This can be reconciled with the calculated data in two ways: Either the HF/6-31G data are correct, and *trans*-P_{II} is really a stable structure; or the observed P_{II} population is

cis-Pro. From CD it is not possible to resolve the conformation of the peptide bond. The results from HF/6-31G level with the reaction field, omitting C_7 structures, give 68.8 % *trans*-α_R-exo, 7.5 % *trans*-P_{II}-exo, 23.0 % cis-P_{II} structures and less than 1 % of *cis*-α_R structures. The two possibilities both fit well to the experimental data.

3.3.3 Scan of ψ

One possible way to improve the unsatisfactory solvation results would be to try another continuum solvation model. The solvation model in Gaussian 92 is not a sophisticated one. It does not take into account the form of the molecule but approximates it by a sphere with an adjustable radius. Other solvation models, for example in AMSOL (Cramer 1991) and the Generalized Born / Surface Area model (Still 1990), used in Macromodel (Mohamadi 1990), calculate the surface of the molecule. Adding solvation effect from a good solvation model to a potential energy surface from Gaussian should give better conformational energetics. Because of the ongoing interest in the laboratory for Conformational Memory calculations by capabilities incorportated in MacroModel (See Chapter 4), the use of Generalized Born / Surface area model would be desirable. The potential energy values of the existing proline parameters in MacroModel are not in full accordance with the proline calculations described above. For example, the relative energies of the *cis* minima are changed, and the low energy area in *trans* around ψ=70 is much narrower. Both endo and exo forms of *cis* and *trans* Pro were fully optimized for several values of the ψ angle, for a total of 70 points. See Figure 3.3 on the next page for the potential energy curves. The corresponding curves from MacroModel are also in Figure 3.3. In principle, if the force field parameters

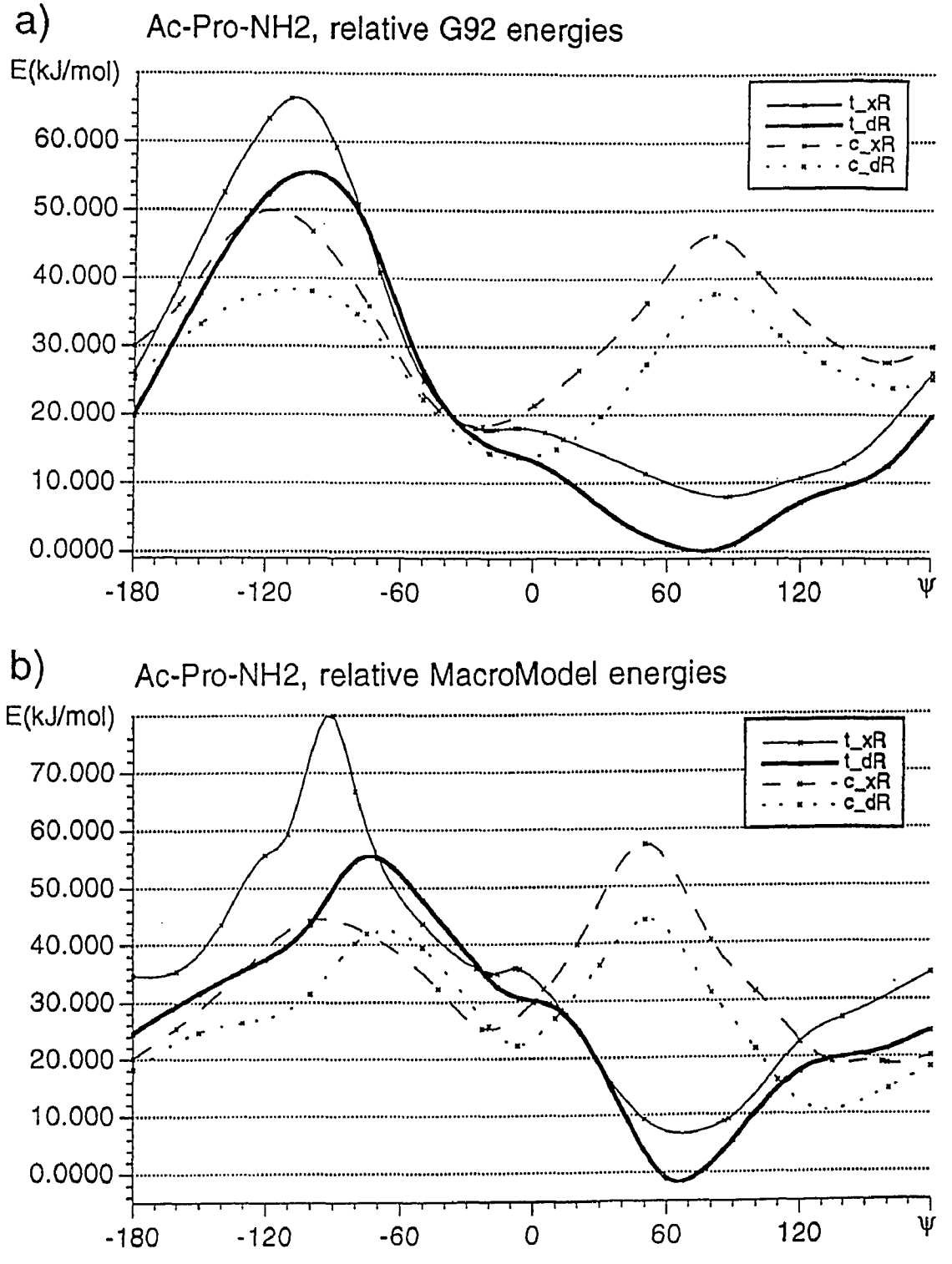


Figure 3.3. Potential energy curves of $AcProNH_2$ as a function of the torsion angle ψ in vacuum. a) Gaussian92 results, optimized at HF/6-31G*, b) Macro-Model results, on structures optimized with Gaussian92.

were optimal, the two methods should produce close to identical curves. MacroModel offers the capability to build substructures with user defined parameters. The reoptimization of the torsional parameters for proline is complicated by the simplified form of the torsional energy equation in MacroModel:

$$E(tors)= v_1(1+\cos\psi) + v_2(1-\cos 2\psi) + v_3(1+\cos 3\psi) + v_1 + v_2 + v_3. \quad (4)$$

No phase γ is allowed. The authors of the program are in the process of updating the equation to allow the free phase. With the full Fourier equation, trial fitting of the curves has posed no problems. This part has to wait untill the next release of MacroModel will be finished.

It is possible that an extra torsional term that couples ω and ψ is needed for a simultaneous good fit of the four curves. In Figure 3.3, the curves for *cis* and *trans* forms differ mostly at $\psi = 70°$, where *trans* forms have a minimum and *cis* forms have a maximum. The reason for the difference is clear: in *trans* forms there is an internal Hbond at $\psi = 70°$, and in *cis* forms there is repulsion between the CH_3 of the N-terminal acetyl group and the NH_2 of C-ter-minal amide. A modified torsional angle for Pro has been defined by Karplus for the same reason, in a study of *cis-trans* isomerization of a proline dipeptide (Fischer 1994).

As a preparation for the reoptimization of the MacroModel torsional parameters, atomic charges were calculated for the stable conformers of $AcProNH_2$. This was done by fitting point charges to the electrostatic potential with the CHELPG algorithm (Breneman 1990) in Gaussian 92. The charges and their averages are presented in Table 3.3. The charges of non-polar hydrogens are added to the heavy atoms, to which they are bound, because AMBER* uses united atom representation.

	t-α_R-x	*t*-C_7-x	*t*-C_7-d	*c*-α_R-x	*c*-α_R-d	*c*-P_{II}-x	*c*-P_{II}-d	ave
N	-0.378	-0.463	-0.392	-0.278	-0.255	-0.453	-0.411	-0.376
C_δ	0.135	0.154	0.092	0.128	0.094	0.159	0.158	0.132
C_α	0.130	0.190	0.153	0.084	0.081	0.160	0.048	0.121
C_β	0.053	0.019	-0.035	0.061	-0.008	0.014	0.011	0.016
C_γ	0.021	0.038	0.108	0.016	0.068	0.037	0.058	0.049
C	0.440	0.781	0.842	0.776	0.885	0.915	0.976	0.858
N"	-1.012	-1.067	-1.118	-0.949	-1.028	-1.204	-1.182	-1.080
NH1	0.375	0.456	0.472	0.365	0.396	0.454	0.450	0.424
NH2	0.440	0.444	0.454	0.429	0.439	0.492	0.475	0.453
O	-0.649	-0.618	-0.623	-0.642	-0.671	-0.643	-0.645	-0.642
C'	0.749	0.780	0.758	0.665	0.652	0.764	0.765	0.733
O'	-0.642	-0.676	-0.670	-0.619	-0.616	-0.631	-0.633	-0.641
C_α'	-0.52	-0.038	-0.041	-0.036	-0.038	-0.065	-0.069	-0.048

Table 3.3. CHELPG charges for the stable conformations of $AcProNH_2$. In the names of the molecules, *t* stands for *trans* and *c* for *cis*, x for exo and d for endo. The atom names are as in Figure 3.1: N", NH1 and NH2 are the atoms of the amide, and C', O' and C_α' of the acetyl group.

Solvation energies of the seven stable gas phase structures were calculated with these charges in MacroModel. See Table 3.4.

	HF+ZPE 6-31G*, vac	MacroModel solvationE	G92(vac) + MM solvation
trans-α_R-exo	16.33	-14.38	1.95
trans-C_7-exo	8.59	-1.49	7.1
trans-C_7-endo	0.00	0.00	0.00
cis-α_R-exo	15.50	-19.69	-4.19
cis-α_R-endo	11.50	-19.08	-8.58
cis-P_{II}-exo	23.83	-33.72	-9.89
cis-P_{II}-endo	20.12	-32.87	-12.75

Table 3.4. Relative solvation energies (kJ/mol) of stable conformations of Ac-Pro-NH_2 including solvation from MacroModel.

Encouragingly, the solvation energies for the internally Hbonded structures are much smaller than for the open strucures. *Trans*-C_7 is no more the global minimum.

3.4 Summary

Quantum chemical calculations are the most exact computational approach to study the structure and properties of molecules. Calculated conformational energetics of $AcProNH_2$ are in good agreement with the experimental data in vacuum, but the reaction field in Gaussian 92 can not handle solvation properly. If the internally Hbonded structures are eliminated from comparison, the results are good.

Prolines are so important for the properties of proteins that obtaining a reliable force field presentation of them would be valuable, especially for the study of peptides. The strategy to calculate the solvation effects by adding a detailed continuum solvation stabilization to a good gas phase energy curve has proved to be succesfull (Still 1990). The proline parameters are worth pursuing, although the work could not be finished now.

4 Conformational search of TRH and its analogs in water

4.1 Introduction

The goal of studying the conformational space of TRH and some of its analogs in this work is not primarly to describe the behavior of the tripeptide, but to try to distinguish effects caused by the different conformational preferences of the compounds from the possible changes in contacts with the receptor. Several other researchers have also tried to define the binding conformation of TRH (Garduño-Juarez 1990; Olson 1993) but based on whole animal data.

The common strategy in studying ligand-receptor interactions is to modify functional groups of the ligand to form analogs. This is often combined with mutations of the receptor. The observed differences in the binding of analogs are interpreted in terms of interactions by the changed functional groups. In fact, the results of biological experiments show a net effect of all differences between the parent compound and the analog, both those that are intrinsic to the ligand, and those that arise from the changed ligand-receptor interactions. One can try to dissect the two effects by first studying the conformational behaviour of the analogs separately, either by physical studies or by calculations. If the populated torsional spaces are similar in two structurally related analogs, the observed effects are undoubtedly caused by different interactions. In the opposite case, when the conformations differ, it is not possible to tell conformational effects apart from the altered contacts. It is also important to remember that a protein provides a totally new anisotropic environment for the ligand. The size and form of the binding pocket limit the possible forms of the ligand, and contacts with the receptor can provide large amounts of energy to change the preferred conformation. Comparisons between structurally variant analogs with different de-

grees of freedom are complicated by the fact that changes can compensate each other, and often what counts for binding and activation is the overall form, or the relative positions of a few functional groups relative to each other.

A special impetus to search the conformational space of TRH was given by a pair of analogs, that are covalently constrained to be *trans* at the second peptide bond. The analogs were synthesized to address the question about the importance of *cis* and *trans* isomers of TRH. See Figure 4.1 for the formulas.

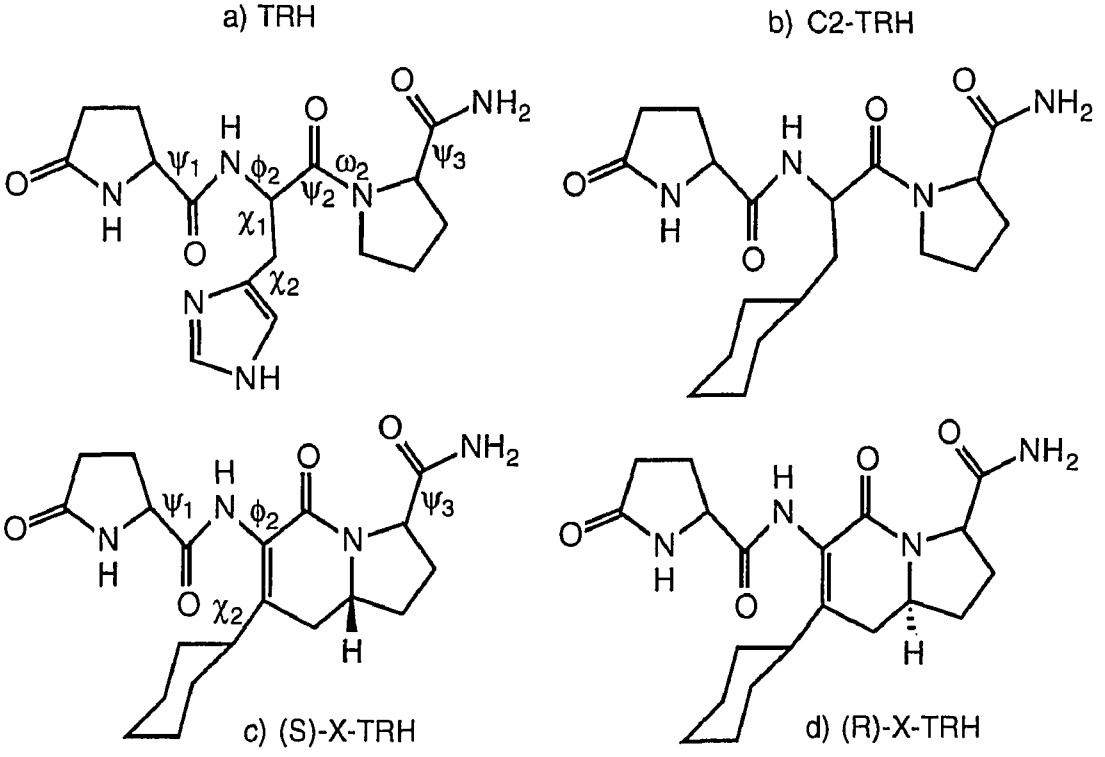


Figure 4.1. Chemical formulas of a) TRH, b) cyclohexyl-Ala^2-TRH, C^2-TRH, c) (S)-3-amino-pyroglutamyl-9-carboxamido-4-cyclohexyl-3,4-dehydropiperolid-2-one, (S)-XTRH and d) (R)-3-amino-pyroglutamyl-9-carboxamido-4-cyclohexyl-3,4-dehydropiperolid-2-one, (R)-XTRH.

The C^2-TRH has the imidazole of histidine substituted by a cyclohexyl ring. The constrained derivatives (S)-XTRH and (R)-XTRH have a methylene bridge between C_β of the second residue and C_δ of Pro, and differ from each

other in the stereochemistry at C_δ. The two analogs are diastereomers. There is also an additional double bond between C_α and C_β of residue 2. The bridge eliminates two free torsional angles, $\psi 2$ and $\chi 1$, and constrains ω_2 to *trans* conformation. The double bond transforms a previously sp^3 hybridized C_α to a sp^2 C_α'. The new fused ring system is nearly planar, with only puckering at the bridge carbon and in the proline ring.

The ability of the three compounds, C^2-TRH, (S)-XTRH and (R)-XTRH to bind and to activate TRHR has been tested on COS-cells (M. Gershengorn, unpublished results). The ratio of K_i/EC_{50} is same for all compounds, indicating that they differ only in affinity for the receptor. They activate the receptors fully, but with greatly different EC_{50}s: TRH - 0.65 nM, C^2-TRH - 750 nM, (R)-XTRH - 29 000 nM and (S)-XTRH - 33 nM. The bridge clearly does not prohibit binding or activation, but its sterochemistry has a 900 fold effect.

The assumptions in comparing the compounds are that C^2-TRH binds in the same way as TRH, and changes from C^2-TRH to XTRHs display increased rigidity of these analogs. The vastly different potencies of structurally nearly identical (S)-XTRH and (R)-XTRH suggests that the difference is conformational: (S)-XTRH is mainly confined to a conformation that binds optimally to the receptor, and (R)-XTRH mostly populates conformations that cannot bind to the receptor. Based on this assumption, determining the conformation of (S)-XTRH would give a good model for the bioactive conformation of TRH itself.

4.2 Methods

The structures of C^2-TRH, (S)-X-TRH and (R)-X-TRH were built in MacroModel (Mohamadi 1990) and optimized with the AMBER* forcefield (Weiner

1981; Weiner 1984; McDonald 1992), each in the two possible puckered forms of the proline ring. All conformations have the rest of the structure equatorial to the cyclohexyl ring. The two constrained analogs are very similar but they prefer oppositely puckered prolines. See Table 4.1 for the optimized initial structures in water, and their relative energies.

	(S)-XTRH		(R)-XTRH		C^2-TRH	
	exo	**endo**	**exo**	endo	exo	**endo**
$\psi 1$	175.5	175.5	162.1	160.5	167.2	167.9
$\phi 2$	72.8	73.3	57.0	55.8	123.3	123.4
$\psi 2$	170.8	168.8	-165.3	-165.4	-152.3	-153.2
$\chi 1$	2.3	1.3	0.3	-0.3	65.9	65.8
$\chi 2$	65.3	66.0	118.0	118.0	51.1	51.1
$\varpi 2$	160.3	167.9	-172.2	-163.6	-178.8	177.9
$\psi 3$	133.7	150.4	145.0	159.0	135.8	146.5
θ	27.2	-38.6	36.2	-28.7	36.9	-36.9
E(kJ/mol)	29.8	10.6	0.0	5.2	1.5	0.0

Table 4.1. Initial minimized structures and relative energies of the analogs. The most stable conformation of each structure is shown in bold face. The angles are defined in Figure 4.1.

The conformational spaces of the lower energy conformers of each structure, and of *cis* and *trans* TRH were studied by two stage simulations, recently described by Guarnieri and Wilson (Guarnieri 1995). The molecules were first subjected to Monte Carlo simulated annealing (Kirkpatrick 1983) in MacroModel (Mohamadi 1990) with the Generalized Born / Surface Area solvation (Still 1990). Data about the values of the torsional angles of low energy structures are collected from the runs, and combined to so called conformational memo-

ries. These conformational memories guide the second set of simulation to sample only the populated areas in the torsional space, where the probability of reaching low-energy conformations is high. The search space is reduced usually to less than half of the original torsional space. See Figure 4.2 to notice, for example, how ϕ_2 shows well localized populations. Potential energy surface is unchanged, but its unpopulated areas are not sampled.

Each initial simulation consisted of 16 cooling cycles, with the new temperature $T_{n+1} = 0.9 * T_n$, the previous temperature. The cooling started at 800 K and was decreased in 10 steps to 310 K. Each temperature was simulated for 50 000 steps and each total run consisted of $8 * 10^6$ MC-steps. The variables in the simulations were the free torsional angles, four in the constrained analogs and six in the open compounds. At each step, two angles were changed simultaneously by a random value between -180° and 180°. The bond lengths and bond angles were kept constant.

In the second part of the simulation, the conformational memories guided the collection of 1000 structures from a Monte Carlo simulation of 500 000 steps at 310 K. The structures were clustered according to their pairwise rms differences with the program Xcluster (Shenkin 1994) for structural comparisons.

4.3 Results

The distributions of the torsional angles of C^2-TRH and XTRHs as a function of temperature are shown in Figure 4.2 on page 54. The three compounds are in separate colums. Angular values that are free in C^2-TRH but covalently constrained in the bridged analogs are shown in the bottom of the Figure. Each 3D graph has temperature as one axis (800 K - 310 K), angle values as another

(180° - ⁻180°) and the fractional population as the third axis. Data of *cis* and *trans*-TRH in similar format are shown in Figure 4.3 and 4.4 on the following pages. The Figures 4.2, 4.3 and 4.4 show preferred and unpopulated values separately per each angle. It is not possible to reconstruct the structures of the peptides from these graphs, because the probabilities may be conditional, as determined by the covalent structure of the analogs. For example, certain values of $\chi 1$ and $\psi 3$ together cause the carboxy terminus to clash with the backbone carbonyl of His. See Table 4.2 for the angles that define the structural families.

Only two torsional angles in the constrained analogs are not affected by the extra ring. These angles are the first and the last torsions: ψ_1 that governs the relative position of pGlu, and ψ_3 that rotates the carboxamide tail. All compounds resemble each other for ψ_1. It prefers values around 180°, although the distributions are clearly wider for the open compounds than for the constrained ones. *Cis*-TRH and C^2-TRH also have minor populations at 0°. The ψ_3 angle has a major population at 160° in all compounds and a minor population at -20°. (S)-XTRH shows also an intermediary value, 70°.

Differences between the open and constrained compounds are seen in the angles that span the altered area around Cα of the second residue. The angles ϕ_2 and χ_1 differ most profoundly between the constrained analogs and the free compounds. In TRH, ϕ_2 shows a broad distribution around 180°, and it also has a wide range of values in C^2-TRH, but centered at 160°. In the bridged analogs the ϕ_2 distribution is bimodal, with peaks at ±60°. The height of the vertical axes in the XTRHs graphs demonstrates, how well localized these angles are. The two populations are present in both compounds at hot temperatures, but converge to $\phi 2$ = -60° in (S)-XTRH and $\phi 2$= 60° in (R)-XTRH at 310 K.

Figure 4.2. Populated conformational spaces of C^2-TRH and XTRHs.

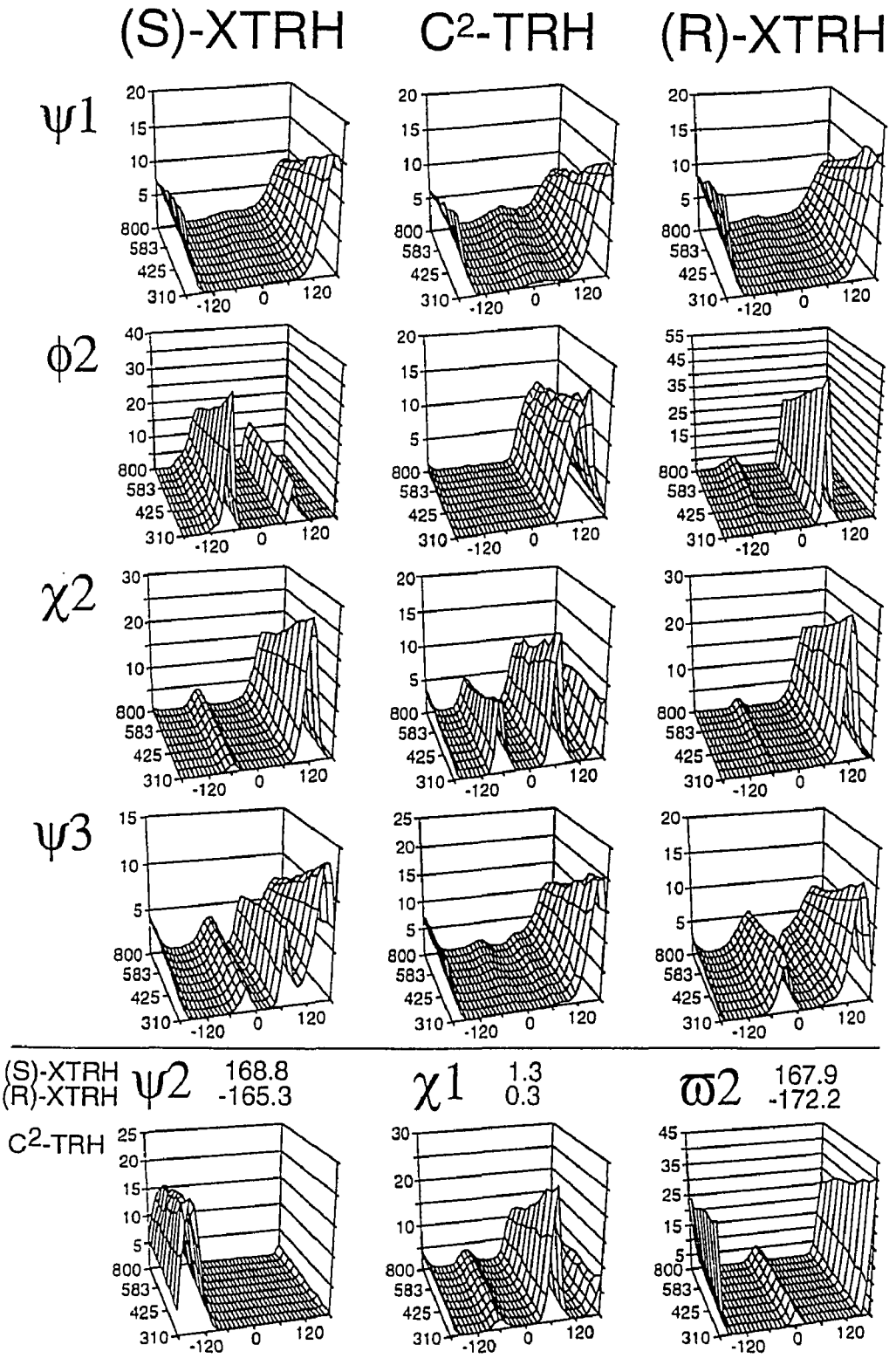


Figure 4.3. Populated conformational space of *trans*-TRH.

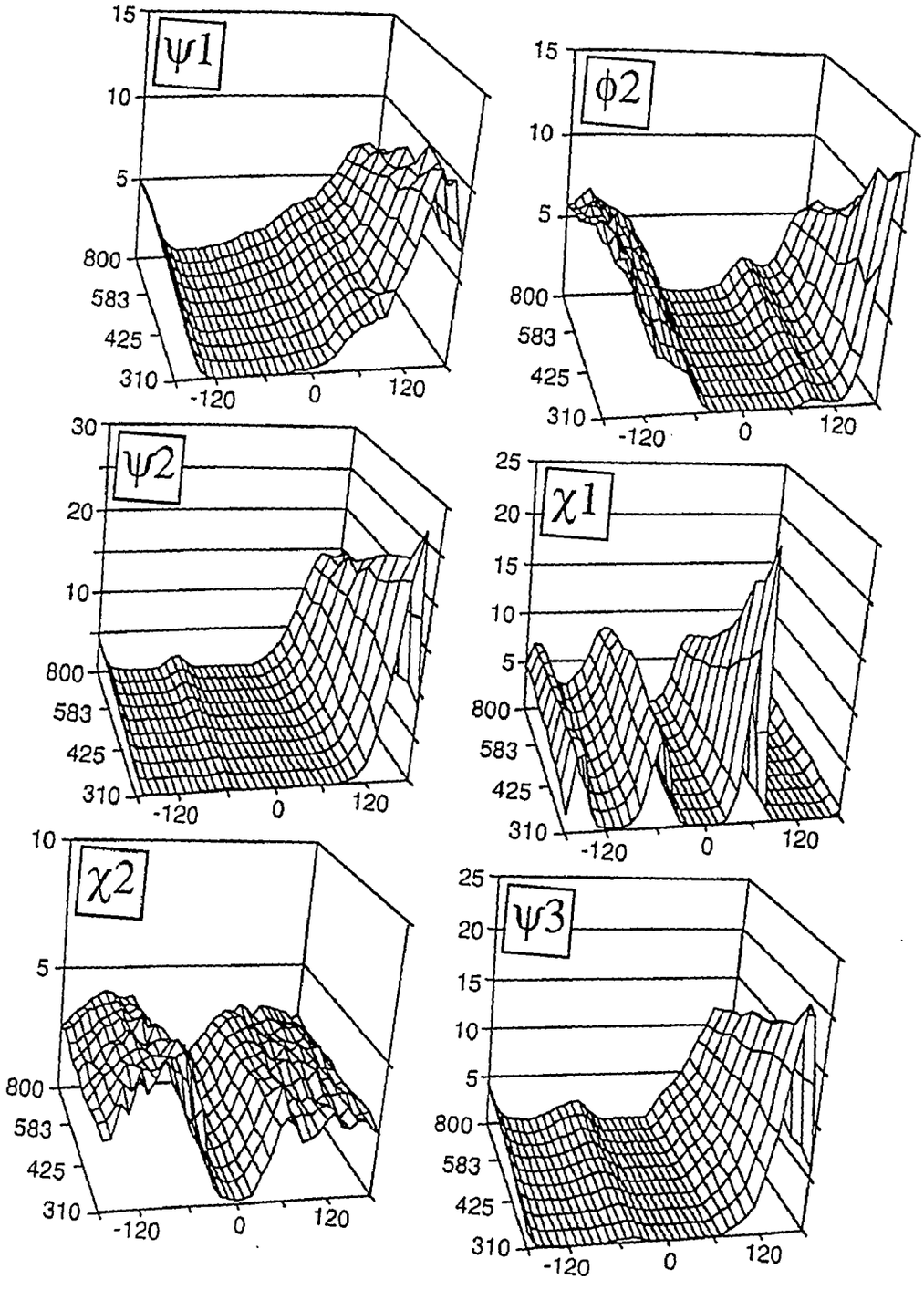
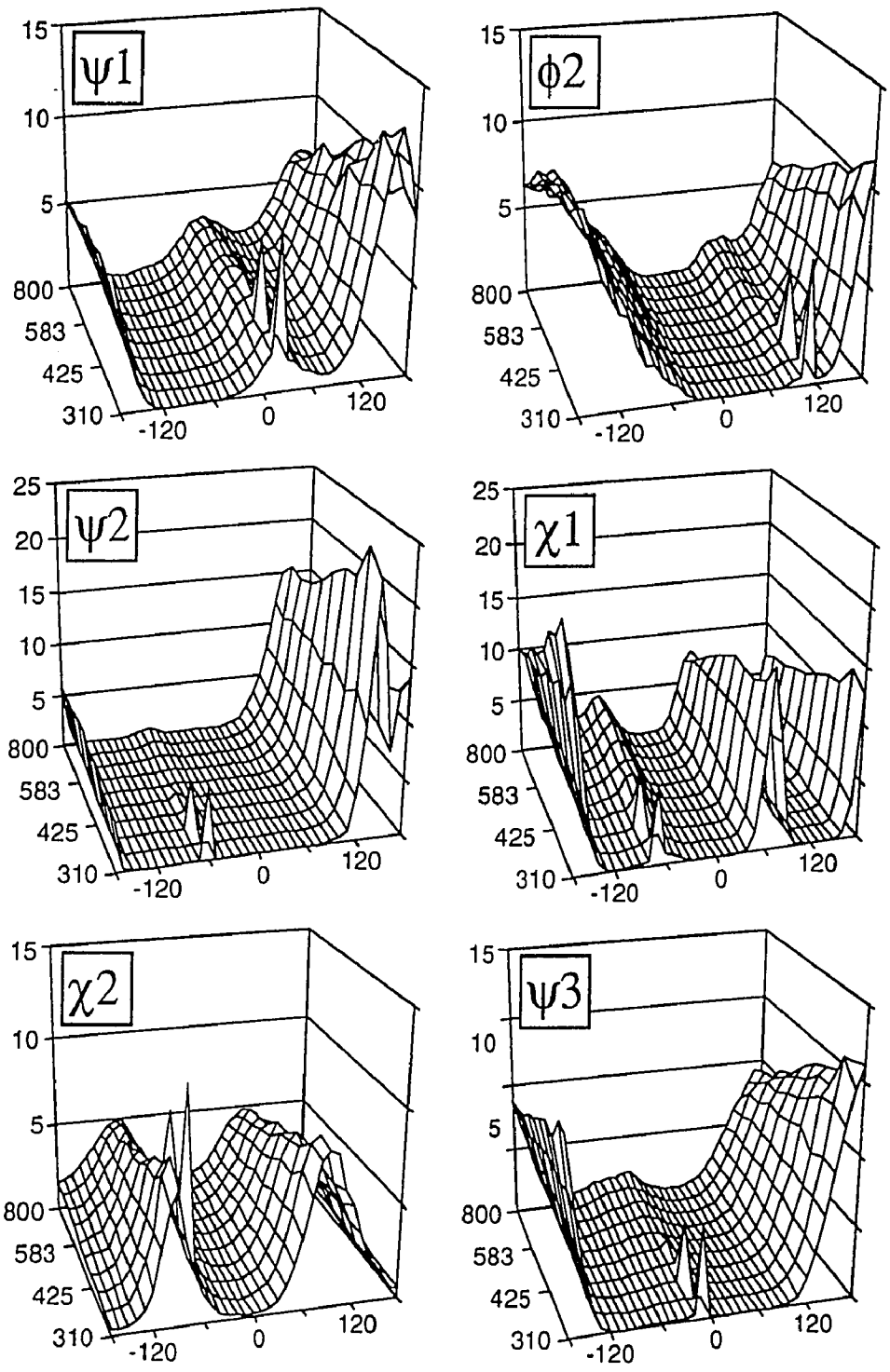


Figure 4.4. Populated conformational space of *cis*-TRH.



The χ_1 angle shows a trimodal distribution in the free compounds, as is expected for a single bond between two sp^3-carbons. The relative populations differ from *trans* to *cis*, in accordance with experimental data (Unkefer 1983). In XTRHs the χ_1-angle is over a double bond, and accordingly the system is planar, $\chi_1 = 0°$. The tilt of the ring of the second residue is governed by χ_2. The difference between the aromatic imidazole and the aliphatic cyclohexyl is seen in the χ_2 distributions, which are bimodal for the planar His ring and trimodal for the tetrahedral connection to the cyclohexyl ring. The χ_2 distribution is exceptionally soft in *trans*-TRH, only a small fraction of values around $\chi_2 = 0°$ is unpopulated. In the constrained analogs χ_2 is well localized to values around 120°. The ψ_2 angle behaves similarly in all compounds. In the constrained analogs the peptide bond, ω_2, is close to 180°. In C^2-TRH this angle was kept as a variable, and a small population of *cis* form is seen. For TRH the two conformations were studied separately.

One thousand structures of each compound were collected at 310 K, as described before. The reason to go from the angular distributions back to structures is to see which combinations of angles are found together. The angular distributions dissect the differences between the compounds and help to target the questions about conformational differences, as for ϕ_2 in this work. Comparison of structures, on the other hand, tells more about the overall differences and similarities. The structures were clustered according to their pairwise rms deviations into conformational families (Shenkin 1994). The families and the angles that define them are presented in Table 4.2 on the next page. Only one value for each angle is given in the Table 4.2, but as the previous graphs 4.2, 4.3 and 4.4 show, none of the torsions is strictly localized. All torsional values that belong to

a continuous distibution are represented by the one number shown in the Table. The angles are defined in Figure 4.1.

	ψ_1	ϕ_2	ψ_2	χ_1	χ_2	ψ_3	% total	
trans-TRH	160	160	170	50	-160	160	88	
	160	-160	160	-70	-100	160	6	
	160	160	160	-160	50	160	5	
	60	100	-60	-70	-80	160	<1	
cis-TRH	10	v a r	y i n g	v a l	u e s		3	
	160	180	160	60	100	180	21	
	160	180	160	180	70	180	35	74
	160	180	160	180	-80	180	39	
C^2-TRH	-10	140	-160	60	60	180	<1	
	180	120	-140	-60	60	180	<1	
	180	120	-140	60	±60	180	94	
	180	120	-140	180	±60	180	4	
ω_2=-5	180	120	-120	60	±60	180	2	
(S)-XTRH	160	-60	169	1	120	160	82	
	160	-60	169	1	120	80	9	91
	160	-60	169	1	120	-30	<1	
	160	60	169	1	120	160	8	8
	160	60	169	1	120	80	<1	
(R)-XTRH	160	60	-165	-0	120	160	89	
	160	60	-165	-0	120	80	<1	100
	160	60	-165	-0	120	-30	11	

Table 4.2. Populated conformations of TRH, C^2-TRH and XTRH. The last column shows some families combined for easier comparison.

Trans-TRH, C^2-TRH and the constrained analogs all have one strongly predominant conformational family, covering about 90 % of the sample. *Cis*-TRH has three major conformational families, that differ in the position of histidine. As can be seen from Table 4.2, the differences between the conformational families of the constrained analogs are caused by variations of ϕ_2 and ψ_3. For the open compounds, the dividing angles are ϕ_2, ψ_2, χ_1 and χ_2.

The two compounds with smallest structural difference are the constrained analogs. Both populate mainly one well defined conformation, that is characterized by the value of ϕ_2. For (S)-XTRH, structures with ϕ_2 = -60° dominate, and for (R)-XTRH, only structures with ϕ_2 = +60° are observed. The compounds also have different angular distributions for ψ3, but only values around 160° are found in the collected structures. The difference between $\phi_2 = \pm 60°$ structures can be appreciated from Figure 4.5. If pGlu rings of the two compounds are superimposed, the effect of ϕ_2 = -60°-> +60° is to turn the rest of the molecule by 120°. The predominance of ϕ_2 = +60° structures in (S)-XTRH is shown in Figure 4.6. In trial runs (R)-XTRH was once found in ϕ_2 = -60° conformation but never in the final sample of 1000 structures. In Figure 4.2 the ϕ_2 = -60° conformer of (R)-XTRH is seen at hot temperatures. The population of ϕ_2 = -60° form in (R)-XTRH is at least 1000 fold lower than in (S)-XTRH.

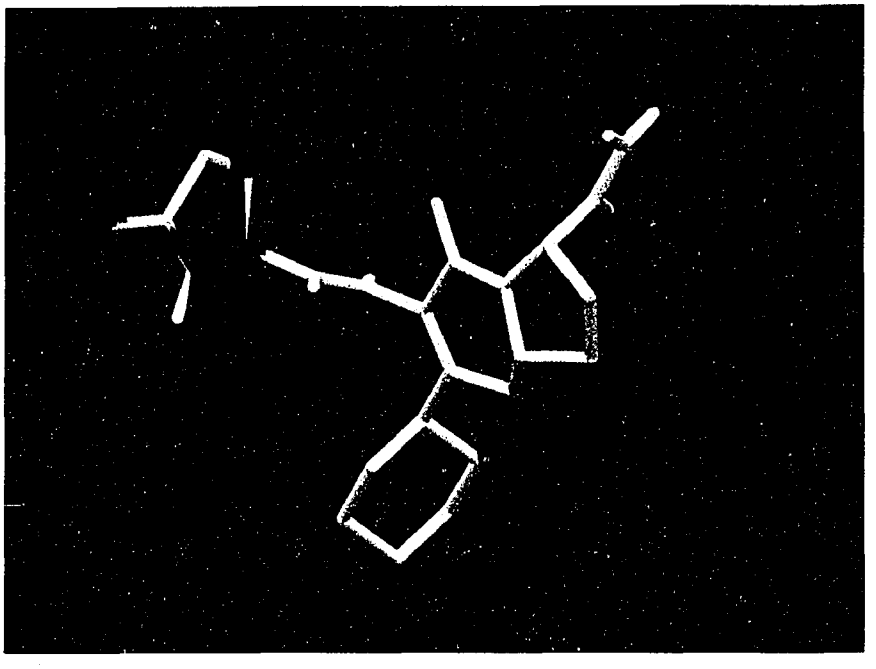
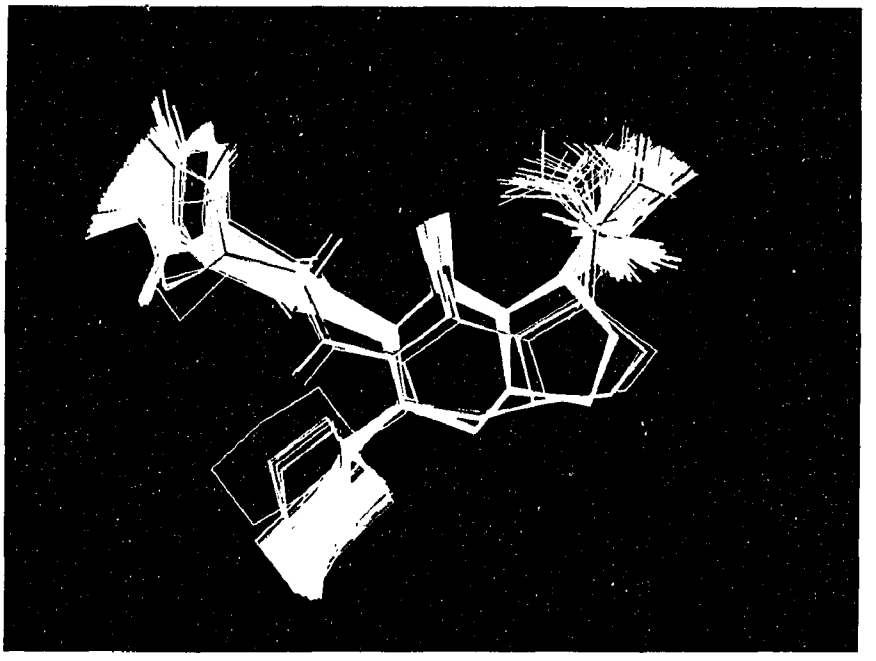


Figure 4.5. The dominant structures of (S)-XTRH (white), ϕ_2 = +60°, and (R)-XTRH (red), ϕ_2 = -60°, superimposed at pGlu.

a)



b)

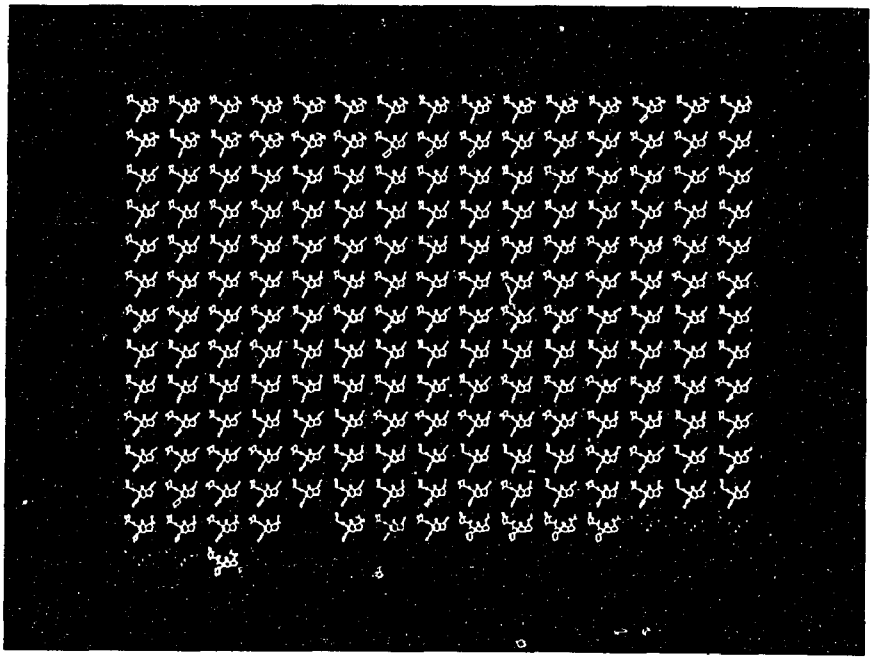


Figure 4.6. a) 200 structures of (S)-XTRH superimposed. Each color marks one conformational family. b) The same structures separated, to highlight the population ratios.

The main difference between the open compounds and the constrained ones is not in the structures themselves but in increased variability of the free angles, as can be seen also from the width of the population peaks in Figures 4.2 - 4.5. *Trans*-TRH and C^2-TRH have one major conformational family. *Cis*-TRH has three major conformational families, as shown in Figure 4.7 below. The variability is also seen in Figure 4.8, that shows 200 *cis*-TRH structures superimposed. The general path of the backbone is the same in all conformations, but the positions of both N and C-termini, and of the His side chain, are scattered. This is in sharp contrast to the well defined structures of (S)-XTRH in Figure 4.5. The same result is emphasized by superposition of the *cis*-TRH structures at pGlu. Figure 4.8 shows one member of each conformational family of *cis*-TRH superimposed at pGlu. The positions of $HisProNH_2$ cover 180°.

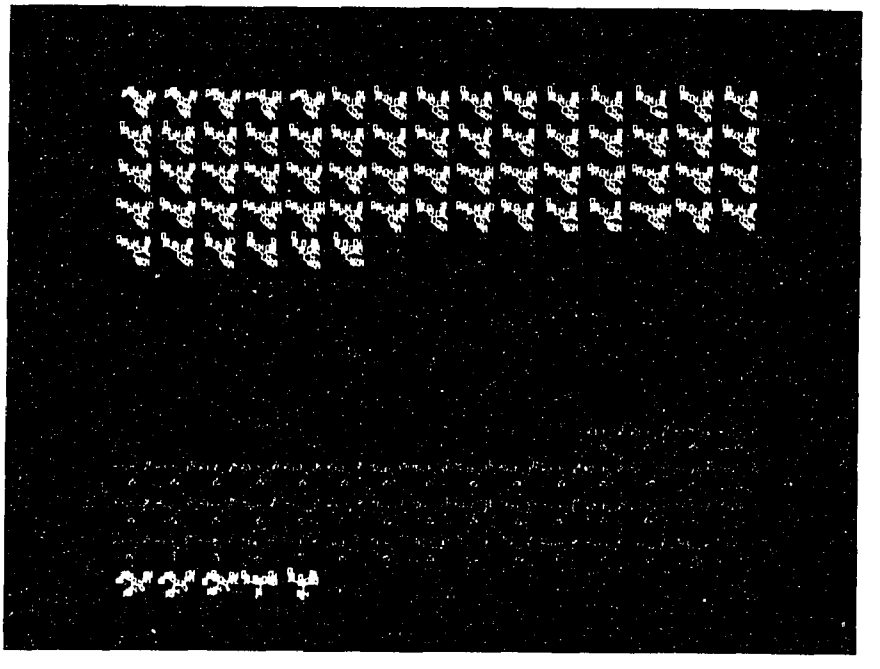
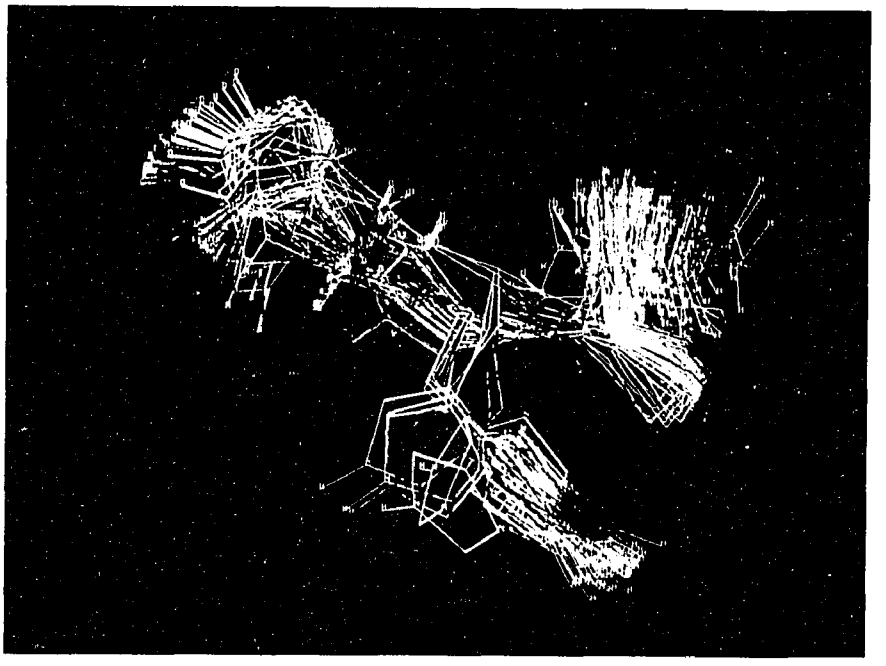


Figure 4.7. Population distribution of a sample of 200 structures *cis*-TRH.

a)



b)

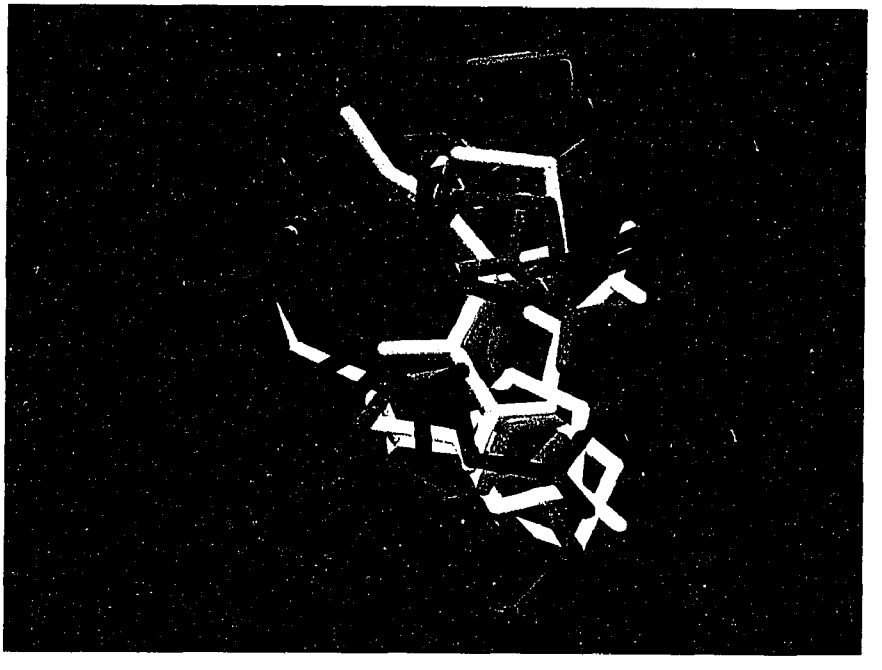


Figure 4.8. a) 200 structures of *cis*-TRH superimposed for minimal rms deviation. Each color marks one conformational family. b) One member of each family superimposed at pGlu.

Comparison of XTRHs to TRH itself is complicated by the double bond of the bridged compounds. The effect of the bond is seen from comparison of the angles ϕ_2 and χ_1 of the structures in Table 4.2. Because of these differences it is not possible to compare the values of the torsional angles directly to each other, but the whole structures have to be considered. All conformational families of TRH found with the conformational searches were compared to the major conformation of (S)-XTRH, with an emphasis on the superposition of the termini. Surprisingly, the best fitting structure is a *cis* form ($\chi_1 \approx 60°$, $\chi_2 \approx 100°$). It fits well for pGlu and His, and in an acceptable way for the carboxy tail. Only the position of the Pro-ring is different in (S)-XTRH and the *cis*-TRH. See Figure 4.9 for a superposition of the two. Similar superpositions of *trans*-TRH with both XTRHs are shown in Figure 4.10 on the next page.

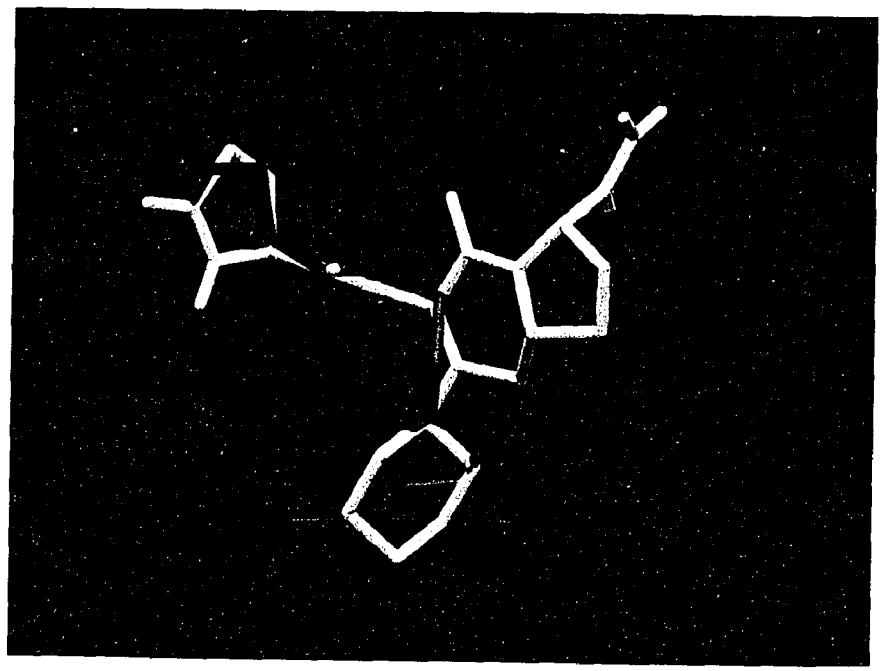


Figure 4.9. Superposition of *cis*-TRH (green) and (S)-XTRH (white).

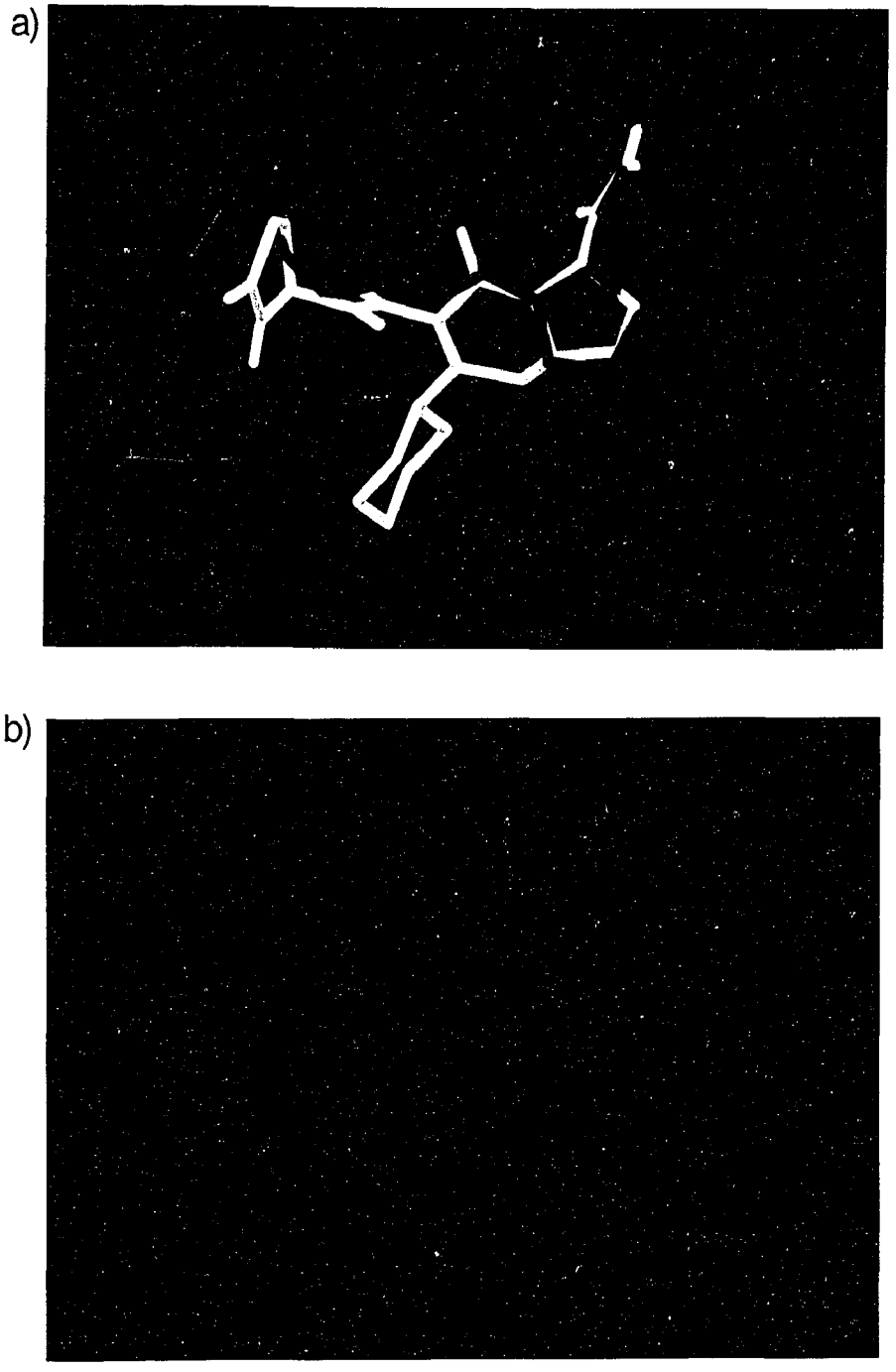


Figure 4.10. a) *Trans*-TRH (blue) superimposed on a) (S)-XTRH (white) and b) (R)-XTRH (red).

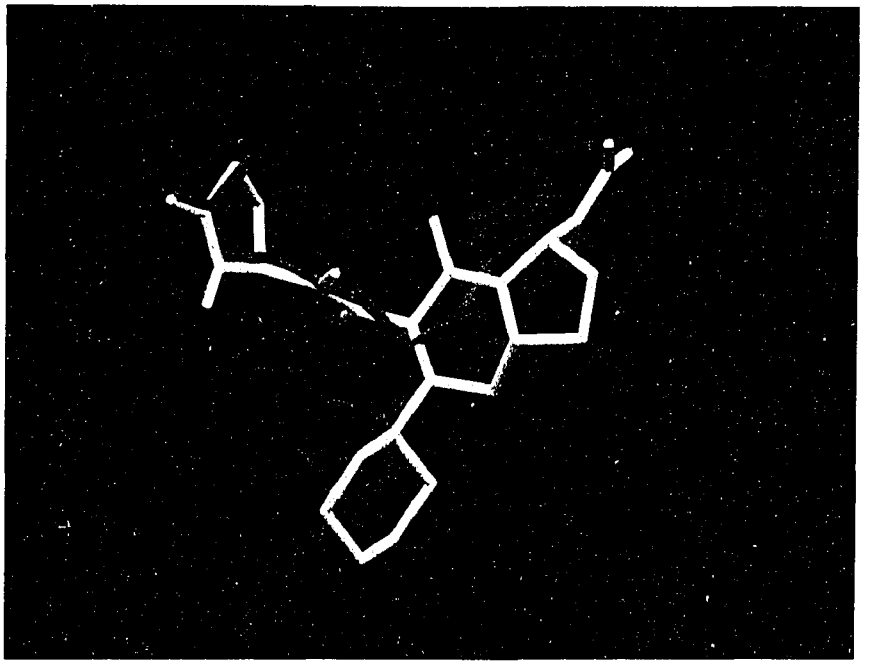
4.4 Discussion

The XTRHs differ from each other only at C_δ of proline. The stereochemical difference causes the angle ϕ_2 to adopt a different position in the two compounds, and the conformational difference reduces their affinity 900 fold. The effect can in principle be caused by two different mechanisms: it is possible that both conformers can bind to the receptor, but $\phi_2 = 60°$ does not fit well to the binding pocket, and either misses some contacts or causes steric clashes. The other possibility is that only the $\phi_2 = -60°$ conformer binds to the receptor, and the observed affinity difference is a reflection of population ratios. The two possibilities could be distinguished experimentally. If the difference is caused by the availability of $\phi_2 = -60°$ conformer, no mutations in the receptor are expected to enhance the K_d of (R)-XTRH. If, on the other hand, the binding site has to adjust to accomodate (R)-XTRH, mutation of a bulky residue in the binding site into a smaller one should make the two analogs bind nearly similarly. The data on MePro (See Figure 2.3) suggests that the binding site is crowded. Examination of the proposed binding site inside the receptor model (See Chapter 6) proposes I109, I116 and L278 as the most likely candidates to change. If the difference in binding affinities of XTRHs is indicative of a missing ligand-receptor contact, the two analogs should bind equally badly to Y106F or R306A.

(S)-XTRH shows so little conformational flexibility in water that it can be assumed to behave similarly inside the receptor. If it is so, its good affinity means that it is constrained to a conformation that binds optimally to the receptor, and it is a good model for the bioactive conformation of TRH. The 50 fold decrease from the affinity of the native TRH most probably reflects differences in the structure of the two analogs inside the binding pocket.

Results from comparisons of TRH conformations to (S)-XTRH were totally unexpected. The analogs were specifically synthesized to study the *trans*-TRH, and yet the conformation of TRH in water that best corresponds to (S)-XTRH is a *cis* form. The energy difference between the major *trans*-form and this *cis*-conformation is small, only 2 kJ/mol. The comparison becomes more interesting, when results from the simulations of the entire binding complex (Chapter 6) are compared to the (S)-XTRH. Figure 4.11 shows the comparison of (S)-XTRH and the *trans*-TRH from the receptor. The fit is good, and now TRH has ϕ_2 = -80°. This tells that the same overall conformation and the same relative positions of the two important amide groups can be reached in several ways. This receptor-bound conformation of TRH is not found in the conformational families in water, although the ϕ_2 = -80° and ψ_2 = -60° areas are not totally unpopulated in Figure 4.3. The energetic difference between the two *trans* forms in water is 52 kJ/mol. This is a large amount of energy, but well within the range of ligand-receptor interaction energies seen in simulations. It is clear from these studies that conformational data of flexible peptides in solution cannot be unequivocally transferred to the receptor, but studies of the actual complex are needed. In addition to exploring the possibilities for the receptor to induce and stabilize *trans*-TRH in α-helical conformation, the possibility of *cis*-TRH being the binding conformation has to be reconsidered, and the behavior of the *cis*-TRH complexed to the TRH receptor to be investigated. There is no data about *cis*-TRH binding the receptor. This is not surprising, as so little is known about the conformational preferences of TRH analogs. Data about MePro-TRH -very little *cis* form, low affinity- is consistent with the idea of the receptor binding preferentially *cis*-TRH, although it is also possible that the bulky MePro-TRH cannot adjust easily to the α-helical conformation the receptor would prefer for the *trans* form.

a)



b)

Figure 4.11. Comparison of the *trans*-TRH from the ligand-receptor simulations (yellow) to a) (S)-XTRH (white) and b)to the dominant *trans*-TRH in water (blue).

5. Construction of a model of the TRH receptor

5.1 Introduction

A molecular model of the TRH receptor is necessary for computational studies of the binding pocket, and it is also instrumental in analysing the experimental studies and in designing new experiments. Scarcity of data prohibits homology modeling. A molecular model of TRH receptor has been built *de novo*, based on a comprehensive analysis of a large set of GPCR sequences by Baldwin (Baldwin 1993). The analysis predicts position, orientation and tilting of helices of a generic GPCR, by addressing the conservancy of physical properties among the receptor sequences, and also the amount of lipid exposed surface area at different heights of the protein. The prediction assumes that the transmembrane helices have the same structure in all GPCRs; that helices pack in consecutive order; that the helix packing can be based on predicted lipid exposure; and that the helices are α-helical. The predicted GPCR structure agrees well with the known rhodopsin footprint (Schertler 1993; Unger 1995). Templates for bacteriorhodopsin and GPCR structures were derived from the data. The templates were scaled by existing bacteriorhodopsin structure (Henderson 1990). A model of TRH receptor was built on the GPCR template by assigning the relative heights to the helices according to brd, and fitting realistic helices to the template. The TRH receptor model is consistent with experimental results. Models of bacteriorhodopsin (brd) and rhodopsin (rhd) were built with the same procedure as test cases. The constructed brd structure is essentially equal to the real one (Henderson 1990), and the rhd model can accomodate known specific Hbonding data (Rao 1994).

5.2 Methods

5.2.1 Forming the templates

Baldwin's prediction of helix packing (Baldwin 1993) is expressed as three cross sections, representing the intracellular, middle and extracellular parts of a model GPCR. The data show the positions of conserved residues in GPCRs, and the extent of buried surface in each cross section. See Figure 5.1 for an example.

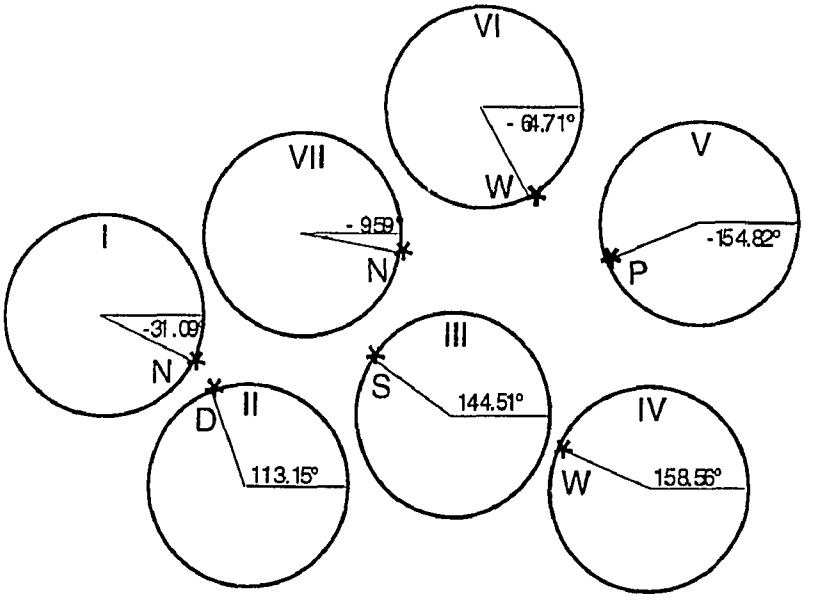


Figure 5.1. Diagrammatic view of the cross section at the middle height of the receptor, showing the conserved residues that are used for orienting the helices. With the systematic numbering scheme (Ballesteros 1995), the residues are N1.50, D2.25, S3.39, W4.50, P5.50, W6.48 and N7.49.

The superposition of the predicted slices matches well with the known rhd footprint. The cross sections could be combined to a three-dimensional (3D) model of helix axes in GPCRs, if the cross sections were scaled to atomic distances and data of the vertical positions of the slices were added. Fortunately, the article also shows brd in similar representation, and that has been used to complete the missing aspects of a GPCR model.

Bacteriorhodopsin. The picture of brd (Baldwin 1993) was scanned at 300 dpi resolution. The center of helix 4 was defined as the x,y-origin in every slice, and thus the z-direction coincides with the axis of helix 4. This agrees with the prediction that helix 4 be the shortest of the transmembrane helices, and should therefore be perpendicular to the membrane in order to span it entirely. Midpoints of helices and their positions relative to helix 4 were defined from the scanned data.

The data from the brd cross sections were converted to molecular dimensions by comparison to the known brd structure. Six interhelical distances, chosen to span the length and the width of the bundle (between helices 1-2, 1-5, 2-4, 3-5, 4-6, 6-7), were determined from the scanned data and from the brd structure. The distances in the crystal were measured between helical axes, determined by the program dials_and_windows (Sklenar 1989; Swaminathan 1990). Because it is not exactly clear which parts of the receptor correspond to the cross sections, two sets of distances were measured in the crystal structure: one between top, middle and last points and the other between 1/5, 1/2 and 4/5 heights of helical axes. Both actual 3D distances and 2D distances projected on the plane of the cross section were measured, generating four sets of distances for each level of the protein. Distances in each of the four sets were divided by the corresponding distances in the picture, the ratios were averaged within the sets, and the average with smallest variation was chosen to be the conversion factor. All values were similar but the ratios to 3D-distances varied more than those to projected 2D-distances, presumably because the helices in the crystal are axially displaced relative to each other. The conversion factor 0.8913 cm/Å $\pm$ 0.0059, measured between 1/5, 1/2 and 4/5 heights of the axis points in 2D, was used to scale the data. The small variance of the conversion factor suggests that the picture is an accurate representation of brd structure.

GPCR model. The cross sections of the GPCR prediction were scanned and measured similarly, and the data of helix axis positions were scaled by the conversion factor defined from brd data. A critical aspect of the GPCR model, that cannot be derived from the projected data (Baldwin 1993; Schertler 1993; Unger 1995), is the vertical positions of the helices. The three planes have been positioned 10 Å apart from each other, because that is the average distance in brd from the middle to 1/5 or to 4/5 of the helix height. The centers of all helices could either be placed at the same level or the relative displacements from brd could be used. The latter option was chosen, but helix 4 was additionally lowered by 2.5 Å, as proposed by Tuffery et al. (Tuffery 1994). It has to be stressed that this is an arbitrary initial positioning, and the vertical displacements of the helices will be refined by forthcoming data on specific interhelical interactions. The completed GPCR template consists of twenty-one points, three per helix, and it defines the helix axes in space. The template is shown in Protein Data Bank format (Abola 1987) in Table 5.1 on the next page.

5.2.2 Fitting helices to the template

Realistic helices (see below) were fitted to the template. The helices were oriented according to Baldwin. The conserved residues used for orientation were chosen from middle height of the helices, to minimize the effect of tilting on orientation. The selected residues and their angles relative to positive x-axis are shown in Figure 5.1.

The construction of receptor models on a template has been implemented as a set of unix shell scripts and gawk programs (available upon request) that use the program Charmm (parameters 22β) (Brooks 1983) for all coordinate manipulations and energy minimizations. The first script forms a helix of a given sequence with representative side chain dihedrals (McGregor 1987). It

changes the puckering of proline to *exo*, because that is the dominant conformation in helices (Milner-White 1992) and because in initial modeling of helices the *endo* form caused sterical clashes between Pro C_δ and backbone CO three residues before it (X_{i-3}). Proline kinks are built in, as defined by Sankararama-

```
HEADER    GPCR axes (A-G), three points per each helix      GPCR   1
REMARK    AXI is a dummy residue with one atom A            GPCR   2
REMARK    each axis starts from the extracellular side      GPCR   3
ATOM      1   A   AXI A   1      -26.078  11.786   7.690    GPCR   4
ATOM      2   A   AXI A   2      -25.831   7.919  -2.310    GPCR   5
ATOM      3   A   AXI A   3      -25.552   4.114 -12.310    GPCR   6
ATOM      4   A   AXI B   1      -20.789   1.856   3.135    GPCR   7
ATOM      5   A   AXI B   2      -19.025   0.124  -6.865    GPCR   8
ATOM      6   A   AXI B   3      -17.323  -1.516 -16.865    GPCR   9
ATOM      7   A   AXI C   1      -10.610   1.516   4.796    GPCR  10
ATOM      8   A   AXI C   2       -9.312   3.341  -5.204    GPCR  11
ATOM      9   A   AXI C   3       -8.074   5.104 -15.204    GPCR  12
ATOM     10   A   AXI D   1        0.000   0.000   7.500    GPCR  13
ATOM     11   A   AXI D   2        0.000   0.000  -2.500    GPCR  14
ATOM     12   A   AXI D   3        0.000   0.000 -12.500    GPCR  15
ATOM     13   A   AXI E   1        2.815  11.817   0.746    GPCR  16
ATOM     14   A   AXI E   2        2.258  12.188  -9.254    GPCR  17
ATOM     15   A   AXI E   3        1.856  12.467 -19.254    GPCR  18
ATOM     16   A   AXI F   1       -7.239  18.221   0.894    GPCR  19
ATOM     17   A   AXI F   2       -7.950  17.478  -9.106    GPCR  20
ATOM     18   A   AXI F   3       -8.662  16.736 -19.106    GPCR  21
ATOM     19   A   AXI G   1      -15.561  13.054   5.821    GPCR  22
ATOM     20   A   AXI G   2      -16.620  11.539  -4.179    GPCR  23
ATOM     21   A   AXI G   3      -17.540  10.085 -14.179    GPCR  24
CONECT    1    2                                            GPCR  25
CONECT    2    3                                            GPCR  26
CONECT    4    5                                            GPCR  27
CONECT    5    6                                            GPCR  28
CONECT    7    8                                            GPCR  29
CONECT    8    9                                            GPCR  30
CONECT   10   11                                            GPCR  31
CONECT   11   12                                            GPCR  32
CONECT   13   14                                            GPCR  33
CONECT   14   15                                            GPCR  34
CONECT   16   17                                            GPCR  35
CONECT   17   18                                            GPCR  35
CONECT   19   20                                            GPCR  37
CONECT   20   21                                            GPCR  38
```

Table 5.1. GPCR template in pdb format. Each axis point is represented by a dummy residue AXI, that consists of one massless, chargeless atom A. The helices are called by letters A to G, with A=1, B=2 etc. The fist point of each helix is in the extracellular side. Notice, that helix 4 / helix D lies on the z-axis.

krishnan et al. (Sankararamakrishnan 1990), with the following backbone torsional angles (ϕ,ψ): Pro_i(-57.2°, -43.9°), X_{i-1}(-55.3°, -50.6°), X_{i-2}(-75.9°, -42.5°), X_{i-3} (-68.9°, -37.9°). Bond angles that differ from standard values are: $C_\alpha(X_{i-2})$-$C(X_{i-2})$-$N(X_{i-1})$ = 119.0°, $C(X_{i-2})$- $N(X_{i-1})$-$C_\alpha(X_{i-1})$ = 122.2° and $N(X_{i-1})$-$C_\alpha(X_{i-1})$-$C(X_{i-1})$ = 112.7°.

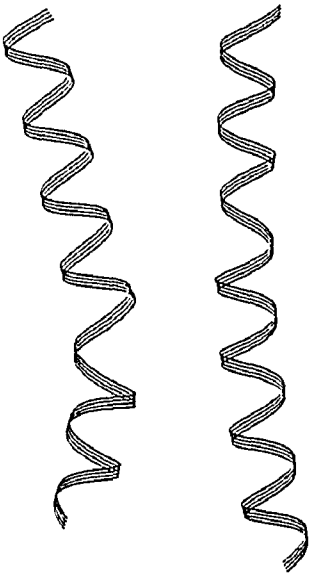


Fig 5.2. An example of a helix without a Pro (right), and a helix with a Pro-kink with the above mentioned angular values (left).

The constructed helices are minimized stepwise, first keeping the backbone frozen, then optimizing all degrees of freedom. This protocol leads to stable helices and prevents side chain - backbone Hbonds from forming. The second script combines the optimized helices into a bundle. It superimposes midpoints of helical axes obtained from dials_and_windows on those of the template and fits the axis points to the line that passes through the template. The third script rotates the helices around the lines fitted to their axes to angles defined for selected residues. Figure 5.3 on the next page shows a diagram of the stepwise fitting procedure.

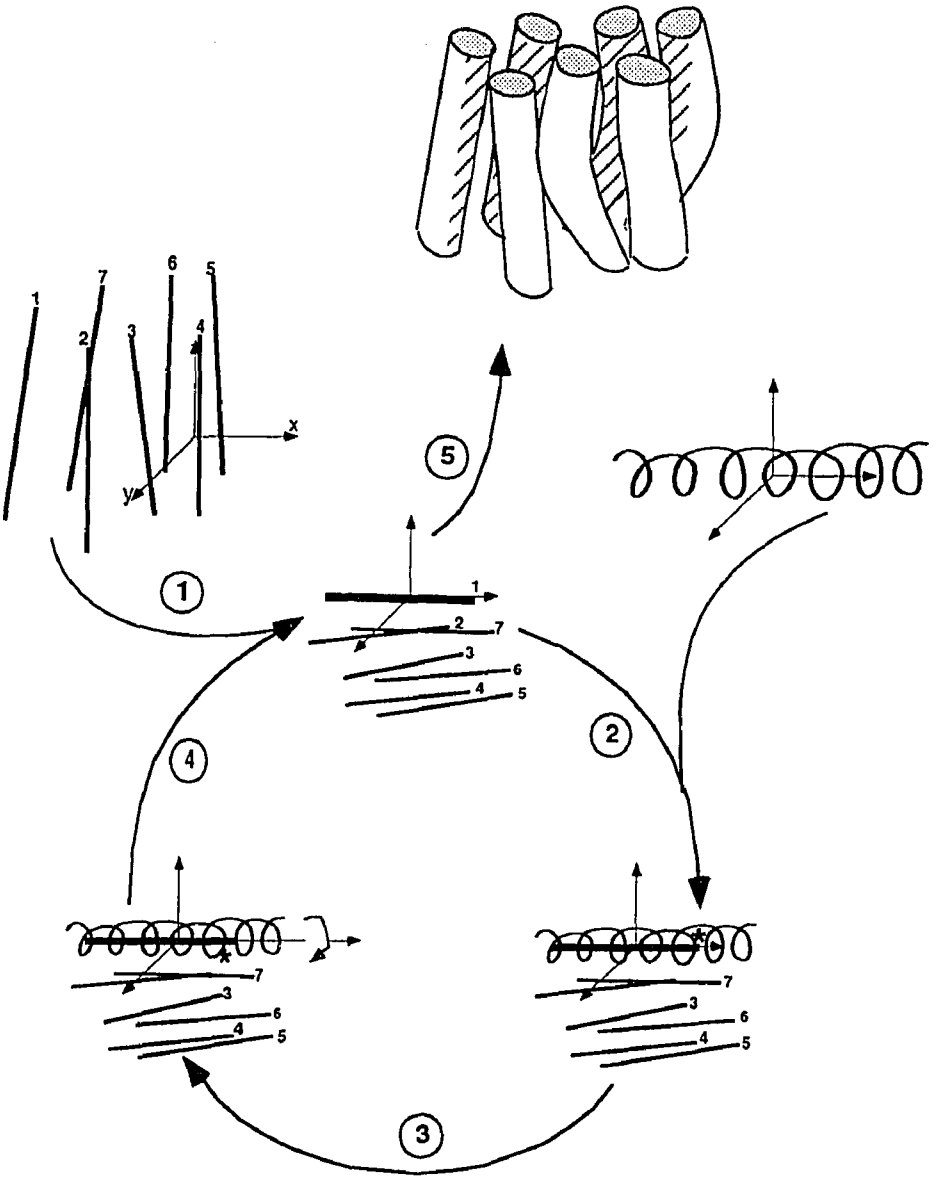


Figure 5.3. Fitting helices to the template. 1) The template is oriented so that helix 1 lies on x-axis, extracellular side in the positive direction. 2) Optimized helix 1, lying on the x-axis, extracellular side in the positive direction, is superimposed on the template. Their midpoints coincide, and the line that fits best the axis points of helix 1 matches the line that passes through the guide for helix 1 in the template. 3) Helix 1 is rotated around x-axis so that the chosen conserved residue points in the desired direction. 4) The template with helix 1 is now oriented to position the guide for helix 2 on the x-axis. The cycle is repeated by adding (step 2) and rotating the next helix (step 3), and reorienting the template (step 4) for the addition of a helix until the entire bundle is constructed. 5) The finished bundle is rotated so that the guide of helix 4 coincides with the z-axis.

5.3 Results and discussion

5.3.1. Testing the templates

Bacteriorhodopsin. The most direct test of the construction is to compare helical axes of the existing brd structure to those determined from the cross sections. Both were treated as 21 points, and the distances between the points were the same in both sets, to focus the comparison to in-plane positions of helices. The rms difference between the constructed and real brd axes is 1.1 Å. The positions of the axes match very well, and the observed difference is caused by the fact that the real axes are kinked, whereas the constructed ones are nearly straight lines. It can be concluded, that the transformation of data from Baldwin's picture to a backbone template was succesfull.

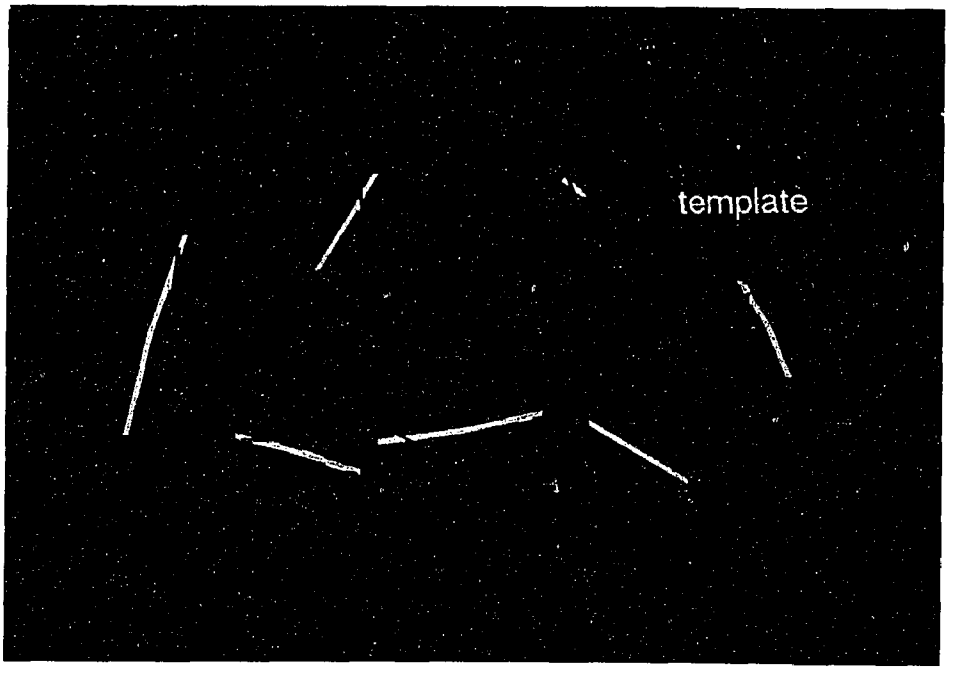


Figure 5.4. Comparison of the constructed brd template to the real brd axes.

The addition of helices to a template by the second script was tested by fitting authentic brd-helices to the constructed brd-template. The resulting structure is practically identical to the real brd: rmsd for Cα is 1.3 Å, and rmsd for all atoms is 1.4 Å. By comparison, building brd totally *de novo* from the sequence with the three scripts described above gives a clearly different structure, with rmsd of 2.4 Å for Cα and 4.2 Å for all atoms to the real brd structure. This comparison highlights the sequence dependent irregularity of the helices.

Rhodopsin. A model of bovine rhodopsin was constructed for comparison to data by Oprian (Rao 1994). They report that a salt bridge between Lys296 in the seventh helix and Glu113 in the third helix is essential for regulated activity. Mutation of either Lys296 or Glu113 causes constitutive activity. The same effect is caused by addition of a negative charge through a mutation in the seventh helix, Ala292Glu, or in the second helix, Gly90Asp. These are assumed to interfere by competing for Lys296 with Glu113. The loss of regulation caused by Glu113Ala can be rescued by a simultaneous mutation Gly90Asp. The authors claim that Asp90 can substitute for Glu113 only if it is at Hbonding distance from Lys296. This information about close contacts in rhd was not used in building the model. Nevertheless, the salt bridge Lys296-Glu113 is found in the model. The distance Lys296-$N_\zeta H$ to Glu113-O_ε is 2.9 Å, which is slightly too long for a salt bridge. This suggests that the relative vertical positions of helices 3 and 7, directly adopted from brd, may need to be revised. Alternatively, the definition of helical segments in the sequence, and thus the relative position of Lys296 with respect to Glu113 might be inaccurate. No specific attention was paid to the definition of helix boundaries in rhd, but the helical sequences corresponding to those chosen for TRHR were used (see below). Lowering helix 7 by 1 Å, which amounts to the vertical distance between $C_\alpha(X_i)$ and $C_\alpha(X_{i+2})$ in a helix, brings

Lys296 to a good H-bonding distance of 1.9 Å from Glu113. This is approximately the accuracy in helix boundary definition.

The double mutant Glu113Ala/Gly90Asp was formed by residue replacement in CHARMm22, and the structure was optimized with distance constraints on the proposed salt bridge. The distance from Lys296-$N_\zeta H$ to Asp90 O_δ in the final structure is 1.6 Å, which corresponds to a good H-bond. Changes in the vertical position of helix 7 would not affect this bond. See Figure 5.5.

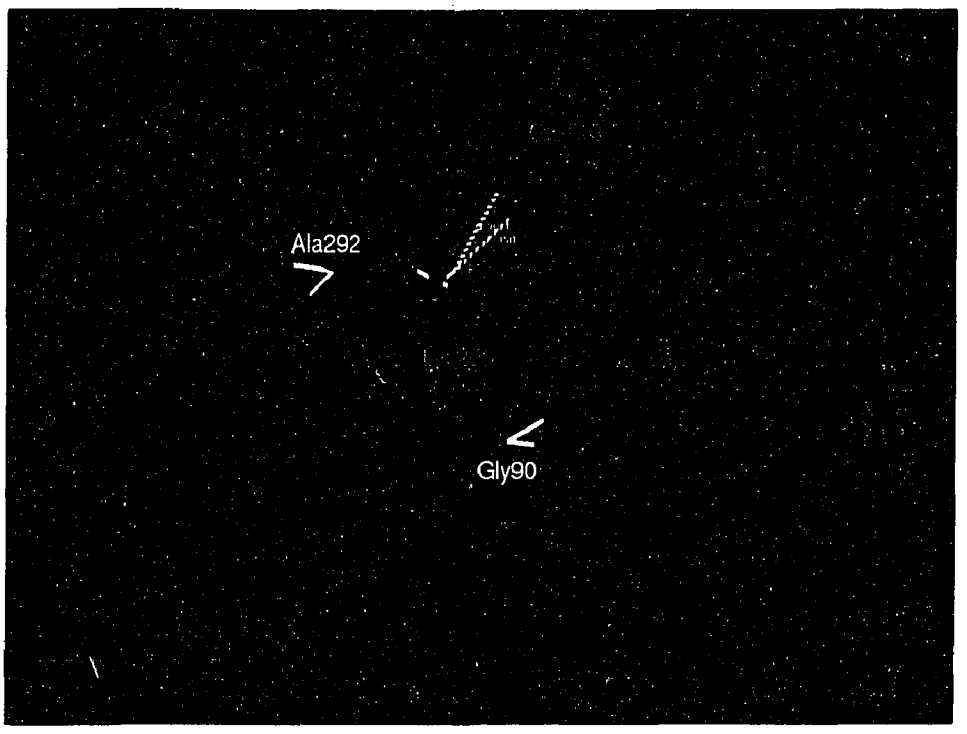


Figure 5.5. Proposed Hbond between Lys296 and Asp90 in the rhd model.

5.3.2 Construction of a TRH receptor model

A molecular model of TRH receptor (TRHR) was constructed according to the procedure described above. No structural data on TRHR is available, but there are inferences of the binding site from mutational studies. None of these data was used in the construction to form an unbiased starting structure for ligand docking. Helix boundaries were defined from an alignment of 39 peptide receptors, and are shown in Figure 2.4 on page 28. Helices formed from the chosen sequences were optimized, fitted to the GPCR template and oriented. The entire bundle was optimized in CHARMm by 3200 steps of adopted basis New-ton-Raphson energy minimization. Minimization did not change the positions of the helices, whether the surrounding continuum was represented by a constant dielectric of 1 or 4, or by a distance dependent dielectric. In all minimizations side chains moved to avoid close contacts and to form new interactions.

Figures 5.6 and Table 5.2 offer different views of the optimized receptor structure. Table 5.2 shows interhelical distances measured between 1/5, 1/2 and 4/5 heights of the helix axes, calculated with dials_and_ windows (Sklenar 1989). The interhelical angles, also in Table 5.2, are measured as dihedral angles between axis points of helices i and j ($C\text{-terminal}_i$-mid_i-mid_j-$C\text{-terminal}_j$) (Treutlein 1992). The angle between helices is 0°, if they are parallel; and 180°, if antiparallel.

The overall structure of the optimized model is similar to the template. Helices that are next to each other in the sequence are also close in space, as shown by the first off-diagonal elements in the distance matrices in Table II. In addition, helix 7 is close to the three first helices at all levels, and helix 3 comes close to helices 5, 6 and 7 at the bottom of the structure. The intracellular end of the helix bundle is more tightly packed than the extracellular part. The longest distance at every level is between helices 1 and 5.

a)

	h1	h2	h3	h4	h5	h6	h7
h1	0.0						
h2	10.7	0.0					
h3	17.9	9.6	0.0				
h4	28.5	20.9	11.3	0.0			
h5	29.8	25.5	17.6	13.2	0.0		
h6	21.1	21.0	17.1	19.8	12.0	0.0	
h7	11.1	13.0	13.2	20.9	19.9	11.0	0.0

b)

	h1	h2	h3	h4	h5	h6	h7
h1	0.0						
h2	12.0	0.0					
h3	17.9	10.1	0.0				
h4	27.8	19.2	10.6	0.0			
h5	29.1	23.1	13.7	13.4	0.0		
h6	20.8	19.8	13.7	20.4	12.0	0.0	
h7	11.4	12.1	10.0	19.6	18.2	10.2	0.0

c)

	h1	h2	h3	h4	h5	h6	h7
h1	0.0						
h2	12.4	0.0					
h3	18.8	12.2	0.0				
h4	27.0	19.2	10.6	0.0			
h5	30.0	24.1	12.9	15.4	0.0		
h6	23.2	21.1	12.7	20.9	10.8	0.0	
h7	10.6	12.7	12.0	22.1	20.5	12.6	0.0

d)

	h1	h2	h3	h4	h5	h6	h7
h1	0.0						
h2	-166.0	0.0					
h3	33.3	-162.9	0.0				
h4	-159.9	10.6	166.5	0.0			
h5	17.5	-167.5	-1.4	-171.0	0.0		
h6	-168.2	3.5	-175.5	-3.1	172.7	0.0	
h7	6.5	-174.6	24.2	-168.7	10.0	-176.4	0.0

Table 5.2. Interhelical distances (Å) a) at 1/5 height b) at 1/2 height c) at 4/5 height of the helices (h) and d) interhelical angles (°) in the TRHR model.

The values of the interhelical angles between neighboring helices (closer than 13.5 Å) are 3.1° - 21.8° for parallel helices, and -153.1° - 169.7° for antiparallel helices, which are in agreement with packing angles commonly seen in proteins (Chothia 1990). The only suspiciously small angle is between helices 5

and 4. Actually 4 and 5 are not strictly antiparallel to each other, because helix 5 is strongly kinked. The angle measured the extracellular halves of 4 and 5 is 10.4°. The value on the bottom is 1.9°, but it is less relevant, because there 4 and 5 are far apart from each other. In general the variation in the interhelical angles shows that the receptor is not simply a collection of parallel and anti-parallel helices.

The structure is not identical to the template, as a closer look at the distance matrices reveals. The template is made of straight lines, and the real helices are kinked. Helices 5, 6, and 7 are kinked at the prolines, the kink angles of axes are 166.2°, 164.1°, 165.4°, respectively. The helices without prolines are approximately straight, with kink angles h1 - 175.4°, h2 - 172.8°, h3 - 175.5°, h4 - 178.0°. Consequently, some distances increase from extracellular part to middle, and decrease again towards the intracellular side. See for example distances between helices 2 and 5, and 6 and 7 in Table 5.2. The tilt of the helices is similar to the template. By all these measures the global structure of the TRHR model is acceptable.

Interhelical interactions. The construction scheme did not impose any interactions between neighboring side chains, and thus the existence of interhelical H-bonds in the optimized structure may suggest potential structural roles. The conserved Asn43 in helix 1 is pointing towards helix 2, and forms a Hbond with Asp71 D71. The conserved serines Ser66 in helix 2 and Ser112 in helix 3, facing the interior, are Hbonding to the backbone CO of X_{i-4}. This is a direct consequence of the building procedure, that sets the value of the torsional angle χ_1 of serines to -60°, because it is the preferred rotamer in helices (McGregor 1987). In this conformation the Hbonding tendency of serines and threonines is satisfied by the backbone, and they can face the lipid as well as the interior of the protein. The possible interactions of Ser66 and Ser112 with

other residues have not been explored in the model. Further experimental results may suggest specific roles for these residues. Similarly, the interactions of Trp150 in helix 4 were not explored, and in the present model it makes no specific interactions. Pro203 affects the general form of helix 5 by kinking it, as desribed above. The CO of Phe199, four residues before the Pro, is left without a helical H-bond, and the sidechain of Arg283 from helix 6 binds to it. The conserved Trp279 stacks with Phe196 in helix 5. Asn316 in helix 7 interacts with the backbone CO of X_{i-4}. Only Trp279 of these conserved residues has been studied experimentally, and Trp279Ala mutant does not differ from the wild type receptor in binding or activation. The model does not help in explaining, why these residues are so highly conserved.

Charged residues. The orientation of charged residues in the model is logical. See Figure 5.6 on the next page. As expected, negatively charged residues, Asp71 and Asp195 face the inside of the helix bundle. Both have been mutated to alanine (Perlman 1992). The mutation Asp71Ala changes the affinity of the receptor to TRH only slightly, but the mutant receptor is inactive. In the model Asp71 is far away from the assumed binding pocket between helices 3 and 6. Asp71 forms an Hbond to the conserved Asn43 in the first helix. Disruption of this Hbond could change the overall structure of the receptor, and be the cause to the slight decrease in affinity. Mutation Asp195Ala does not affect binding significantly. In the model, however, D195 is part of an extensive Hbonding network, that includes residues involved in binding TRH (see below). The positively charged residues on the intracellular end of the helices, Arg52, Arg141, Lys143, Lys144, Lys266 and Lys326 face the lipid in the model, and could interact with the negatively charged lipid headgroups. The highly conserved Glu122 and Arg123 at the bottom of helix 3 both face inside the bundle, in agreement with possible functional roles.

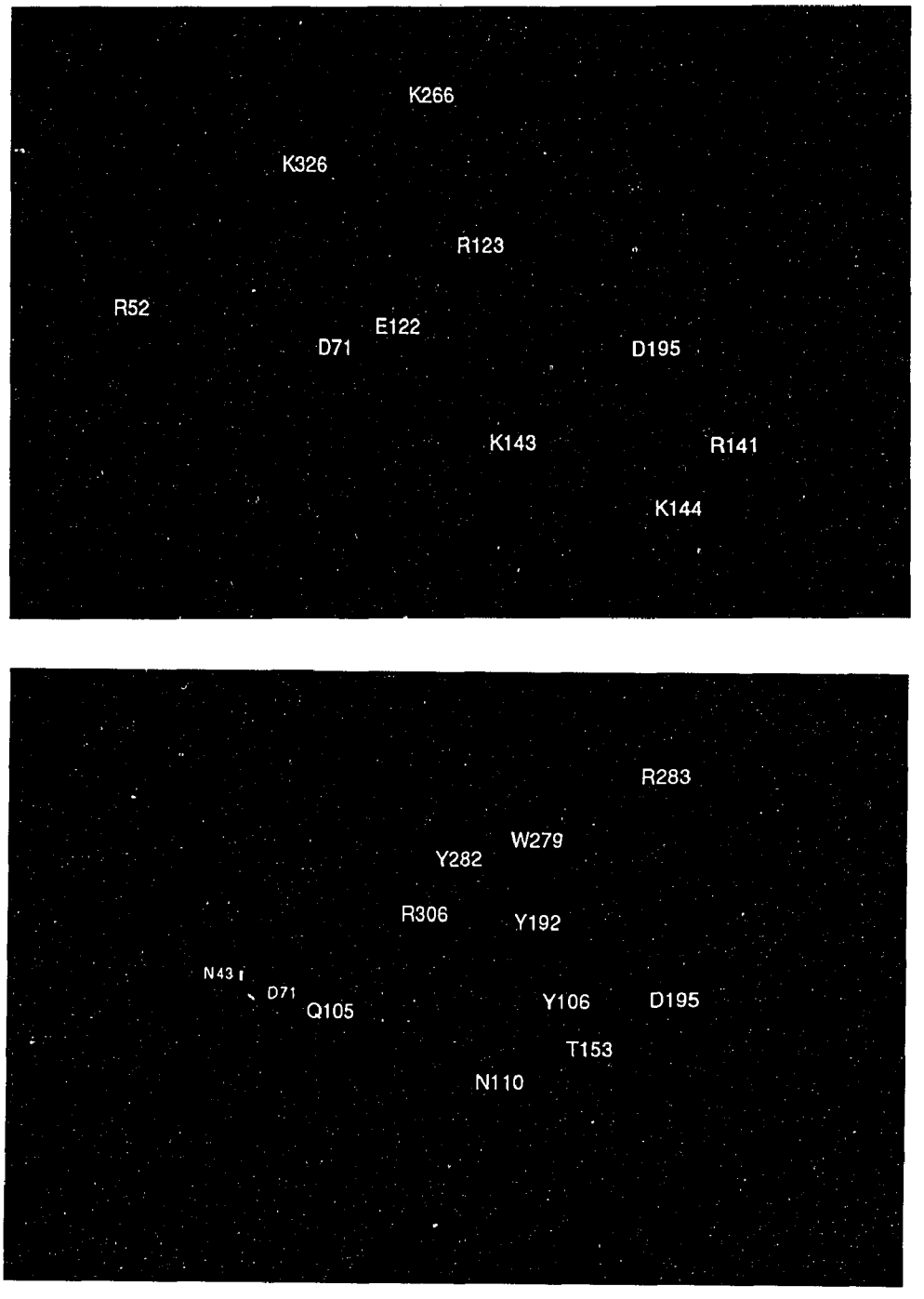


Figure 5.6. Extracellular views of the minimized helix bundle. Helix 1 is on the left. The upper picture shows the positions of the charged residues, except R283 and R306. The lower picture shows interhelical interactions in the binding site area.

Molecular dynamics simulations. The stability of the receptor model was tested by molecular simulations, which required special consideration. Besides the fact that the structure is built *de novo*, and thus might have packing defects, the system is not complete: the receptor consists of seven separate helices, without connecting loops. There is no hydrophobic medium packing to the outside of the helix bundle, and no water or phospholipid headgroups interacting with the charged residues at the ends of the helices. An unnatural system like this could unwind or the helices could drift apart. One could improve the description of the surroundings and add lipid and water layers around the receptor (D. Garmer, personal communication), but that would make the simulations several times more expensive, and the whole setup would need to be studied carefully. At this stage, the goal of running molecular dynamics is to test the general stability of the structure.

First trials to simulate the system with a standard protocol in CHARMm22 resulted in unstable structures. The backbone Cα, C and N were constrained with harmonic terms (25 kcal/mol) during heating from 0 to 300 K in 10 ps, and during 10 ps of equilibration. Once the constraints were released, helices 4 and 6 started unwinding within 0.5 ps. No specific interactions were found to cause the unwinding: all charged residues within 5 Å were checked. Change in the description of the surroundings (ε=4, ε=1 or ε=r) did not stabilize the simulations, neither did neutralization of the arginines and lysines that face lipid in the first intracellular turn of each helix (R52, R141, K153, K144, K266, K326). Obviously, the equilibration of the helix bundle was not sufficient.

Careful equilibration by stepwise heating and assignment of velocities produced stable dynamics in CHARMm22. The system was heated up gradually: 0 to 60 K in 3 ps, Gaussian dynamics for 7 ps at 60 K, similar heating and simulation from 60 to 180 K and from 180 to 300 K. The period of Gaussian dy-

namics had to be long enough, or the temperature dropped after it. The surroundings were represented by a continuum with distance dependent dielectric (ε=4 also worked), and the above mentioned Arg and Lys were neutralized. No constrains, specific force field terms or the presence of the ligand were necessary. Total length of the dynamics was 220 ps. Both energy and temperature were constant after the equilibration phase. See Figure 5.7.

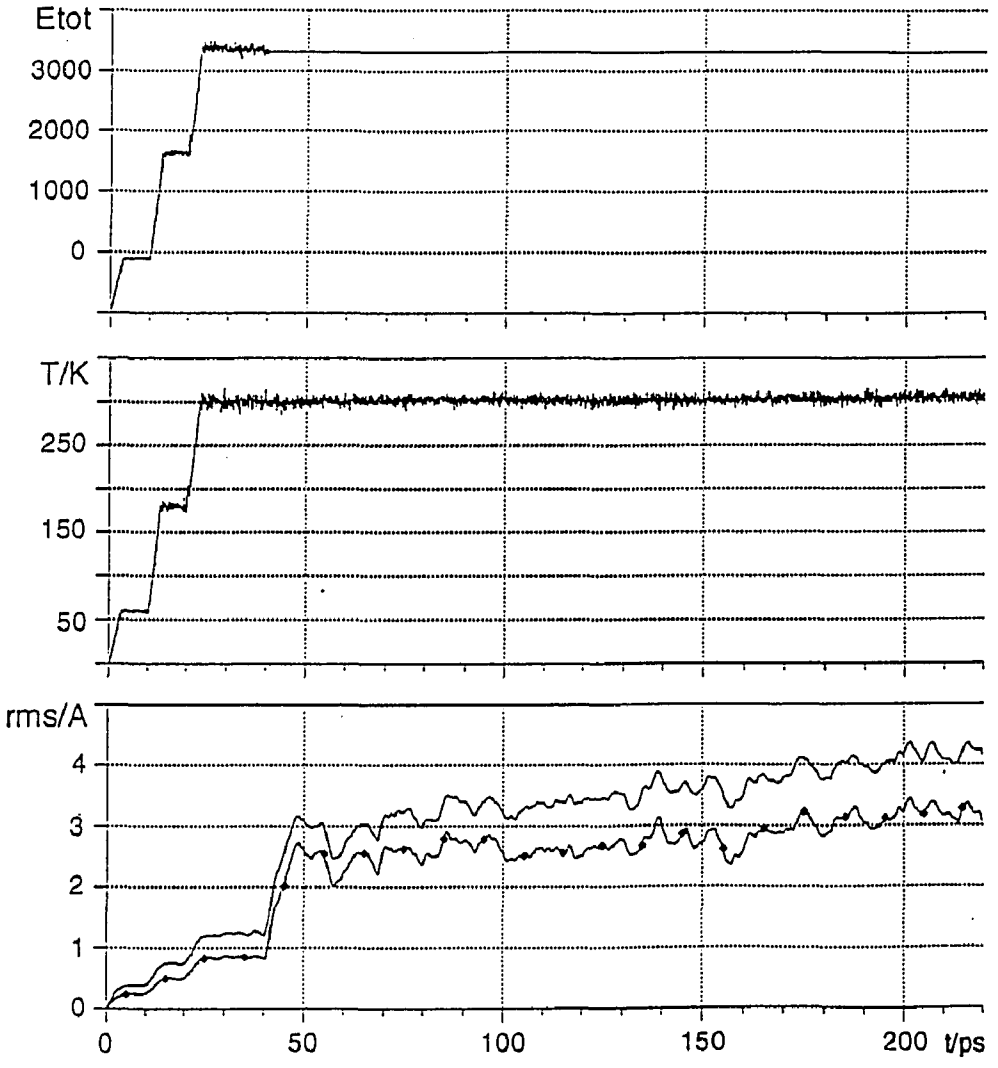


Figure 5.7. Total energy (kcal/mol), temperature (K) and rms deviation for all atoms and for Cαs (line marked with diamonds) as a function of time (ps).

Two average structures were constructed and minimized, one from the beginning of the free simulation (60-90 ps) and one from the end of the run (190-220 ps). The rms difference between the minimized initial structure and the dynamic end structure at 200 ps is 4 Å for all atoms, 3 Å for Cα's. See Figure 5.7 above. The two rms-curves follow each other, meaning that the changes occur mainly in packing, not in the structure of the helices. The integrity of the helices was confirmed by calculating the number of helical backbone Hbonds of the average structures and comparing them to the minimized starting structure. The average structures are more irregular than the starting structure, and with a strict Hbond definition (C=O ···· H <120° and O ···· H < 2.5 Å) they maintained ~ 65 % of the helical Hbonds of the original ideal structure. When the distance limit was relaxed (O ···· H < 2.8 Å) the majority, >90 % of the Hbonds of the starting structure, were present.

The helical axes of the average structures were compared to those of the starting structure to see possible changes in packing. The differences are not large, as can be expected from the small rms deviations. However, one can see a trend of the most exposed helices 1, 4, 5 to move closer to the rest of the bundle. The same effect is seen in the simulations of the ligand-receptor complex, described in the next chapter. See Figure 5.8 on the next page. For example, the 1-5 interhelical distances shorten by about 1.5 Å. The intracellular parts of helices 1, 2 and 6 bend. All the interhelical interactions described for the starting structure (See Figure 5.6) are conserved in dynamics, with the exception of the Hbond between Y192 and Y106. N110 is bound either to T153, or to D195. In short, the bundle of seven helices is stable in molecular dynamics simulations without any of the molecular surroundings, although their lack seems to affect the structure slightly.

Figure 5.8. Helix axes in the initial structure and after dynamics.

Comparison to ligand binding data. The most critical test of the receptor model is comparison to experimental data. Three residues in TRHR have been shown to interact directly with the ligand, Tyr106 in the third helix (Perlman 1994), Tyr282 in helix 6 and R306 in helix 7. There is also data on the interactions with N110. All these four residues face inside the receptor model, as shown in Figure 5.6. The ligand can be successfully docked between them, as is described in Chapter 6. The outstanding Arg283 in the sixth helix is essential for activation, and also important for binding. Arg283Ala mutation reduces the affinity by four orders of magnitude (Perlman 1995). In the model Arg283 H-bonds to the backbone COs of Phe199, left free because of Pro203, and to the backbone CO of Leu193.

The interior of the receptor between helices 3 through 7 is highly polar. Several hydrogen bonds between the helices can be formed by the side chains. The exact binding depends on the starting conformers and optimization conditions, but the residues that participate in interhelical interactions are always the same: Gln105, Tyr106, Asn110, Thr153, Asp195, Tyr192, Tyr282, Arg283 and Arg306. All the polar groups that are supposed to interact with the ligand, are occupied in Hbonding in the unliganded receptor. Introducing the ligand inside the pocket would break these Hbonds and change the whole internal structure of the receptor, as should happen in the transition to the activated state. According to the current model GPCRs exist in equilibrium between inactive state R and active state R* (De Lean 1980, Samama 1993), and agonists stabilize the active state, inverse agonists the inactive state. In this framework, some features of the unoccupied receptor described above are representative of the inactive state R. The unoccupied receptor itself cannot be related to any data, because physical sturdies are are present not possible on TRHR, and biological data only refers to an ocuupied receptor. Further information of the inactive state could be deduced from studies with a newly found inverse agonist to the TRHR (M. Gershengorn, unpublished results), and, when comparing results to studies with the agonist, of the activation mechanism of the system.

6 Simulation of the binding pocket

6.1 Introduction

Binding of TRH to TRHR was studied with a novel technique that combines stochastic dynamics and Monte Carlo steps (SD/MC) (Guarnieri 1994). These mixed mode calculations are advantangous over either Monte Carlo or molecular dynamics alone: First, a Monte Carlo simulation is not bound to local minima, but by random changes it searches the entire conformational space. This is extremely important in a *de novo* constructed model, that cannot be assumed perfect. Good conformational searching is necessary also for the ligand, which as a small peptide has multiple low-energy conformations, and should not be approximated as a rigid object. On the other hand, the additional dynamics steps relax bad contacts, which easily occur in a tightly packed environment.

The target of this study is the binding site. Possible new direct interactions with the ligand are looked for, and secondary interactions to the already identified key residues. A theoretical goal is accurate physical characterization of the binding interactions. This is the first time the mixed mode method has been used for a large complex, and its performance is of interest.

6.2 Methods

Construction of the starting structure. TRH was manually placed inside the helix bundle, guided by connections to Y106 and N110. Histidine was left uncharged because of the shown pH-dependence of binding (Perlman 1992), and the τ-tautomer was chosen. Side chains in the ligand and the receptor were rotated manually to bring R306 close to the terminal carboxamide and Y282 close to the imidazole ring. Figure 2.6 shows the target structure. The system has a total of 189 residues and 1678 atoms in an extended atom model.

Several initial structures were produced in the program MacroModel (Mohamadi, 1990), and the complexes were refined iteratively: a construct was formed, and flat-bottomed distance constraints were set on Hbonds between Y106-pGlu and R306-Pro-NH_2. The system was minimized first for 1500 steps with the distance constraints, and then for 1500 steps without them. In later cycles of iteration, the minimizations converged. Stronger constraints with looser ranges were tried first but they often deformed the peptide bonds. The structures were evaluated for contacts from ligand to Y106, N110, Y282 and R306 after free minimization. Y106 and R306 needed to form Hbonds to the ligand for the structures to be accepted, whereas for the less well defined N110 and Y282 a criterion of proximity to the ligand was sufficient. The planarity of the peptide bonds was also checked, and a deviation of more than 10° was not accepted. Total energy was not a good indicator for these acceptance criteria, but smoothness of the energy curve at the release of the constraints was. The process was continued until a complex that fulfilled all criteria simultaneously was achieved.

Mixed mode simulations. The structure built with manual docking and minimization was used as a starting structure for an extensive simulation of the binding pocket. The mixed mode stochastic dynamics/Monte Carlo method by Guarnieri (Guarnieri 1994) was used. The method is incorporated in MacroModel (Mohamadi, 1990) and uses the modified AMBER* force field (Weiner 1981; Weiner 1984; McDonald 1992). The method combines stochastic dynamics (SD) and Monte Carlo (MC) steps seamlessly into one simulation. This new kind of a simulation achieves both the good local sampling of dynamics and the barrier-crossing of Monte Carlo. In a MC step a given number of torsional angles in a preselected set is randomly changed, and the new configuration is evaluated by Metropolis criteria (Metropolis 1953). If the MC step is accepted, the dynamics step continues from the new position; if rejected, from

the previous one. The problem of combining two profoundly different methods, a deterministic and a stochastic one, to a continuos simulation has been solved by using a velocity Verlet algorithm (Swope 1982). In contrast to the normal Verlet algorithm (Verlet 1967) that advances the positions of atoms based on data at two time points,

$$\mathbf{r}(t + \delta t) = 2\,\mathbf{r}(t) - \mathbf{r}(t - \delta t) + \delta t^2\,\mathbf{a}(t) \quad (1)$$

r is position and **a** acceleration at time t

the velocity Verlet algorithm stores positions, velocities and accelerations at the same time point t, and calculates the next positions based on all three.

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t\,\mathbf{v}(t) + 1/2\,\delta t^2\,\mathbf{a}(t) \quad (2)$$

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + 1/2\,\delta t\,[\,\mathbf{a}(t) + \mathbf{a}(t + \delta t)\,] \quad (3)$$

r is position, **v** velocity and **a** acceleration at time t.

Accelerations are calculated in both cases as derivatives of the potential energy function at **r**. The Monte Carlo steps do not cause discontinuities in these dynamic equations, because no data from the (t-δt) time point -that is, before the MC step- is required. In a large system one has to pay special attention to the definition of the nonbonded cut-off distance, because one single MC steps can cause structurally large changes that bring several atoms close to each other suddenly. As a result of this, the kinetic energy of the system rises in an uncontrolled way, and the simulation stops. The mixed mode method has been shown to reproduce quantitatively binding free energies in a small molecule complex (Burger 1994).

Simulations procedure. The goal of the simulations is to explore the conformations available to the complex. Fifteen separate simulations were performed, each with a different seed number for the random number generator, to enhance the sampling of the conformational space, and to reduce the depend-

ence on the starting structure. This is not enough to reach numerically converged results in the large TRH-TRHR complex. As a comparison, pentane with two free dihedrals needed 10^2 ps for convergence (Guarnieri 1994), and this complex has 97 free dihedral angles. The simulation results do show recurring structural features, which is indicative of adequate sampling. Also, the simulations were started at elevated temperatures and then cooled to 310 K to enhance sampling.

To study the robustness of the starting complex, and to decide about the starting temperature, the receptor-ligand complex was heated to 1000 K in steps of 50 K, with 2 ps of stochastic dynamics at each temperature. The number of α-helical backbone Hbonds was monitored as a function of temperature. See Figure 6.1.

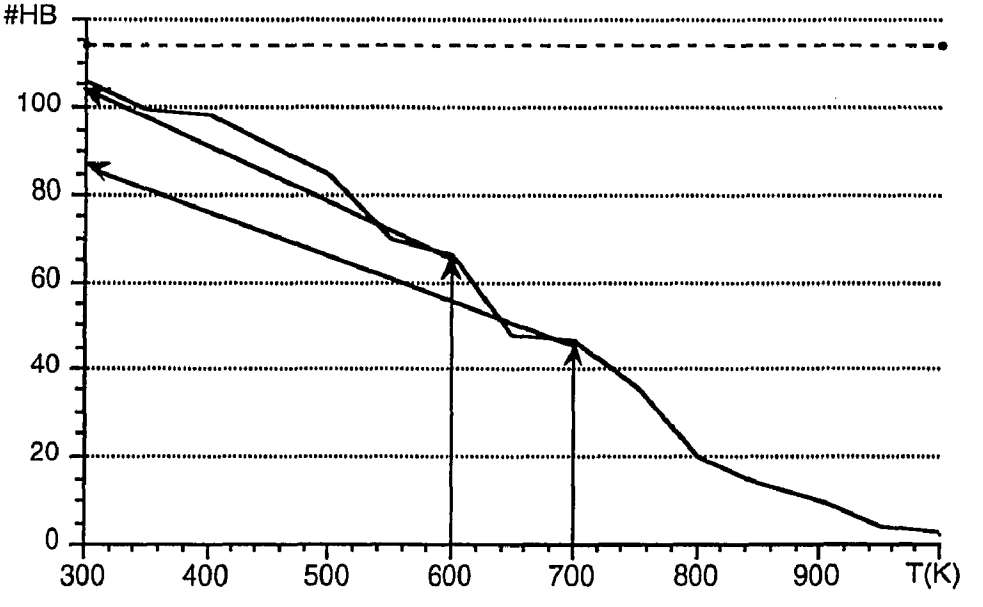


Figure 6.1. Temperature dependence of α-helical hydrogen bonds in the receptor model. The dashed line shows the number of hydrogen bonds in the minimized starting structure, the arrows show the different cooling trials discussed in the text.

At 1000 K practically all helicity was lost but the peptide chains stayed together. Largest Cα to Cα distance at 1000 K in the membrane plane was only 13 % larger than at 310 K. At 700 K the system looked helical and it had about half of the possible Hbonds. Annealing from 700 K to 310 K did not, however, rewind all the helices, but in each of five trials helix 1 had an extended stretch around the conserved GN residues (42-43 in TRHR). When cooling was started from 600 K, all peptide strands became helical again. All following simulations were started at 600K. The system was cooled in steps of 100 K with coupling constant of 0.4 ps, and each temperature was simulated for 30 ps. Analysis and discussion is from 310 K, if not otherwise mentioned.

The set of torsional angles to be varied by the Monte Carlo steps was chosen to include all non-peptide torsional angles of the ligand, and all side chain torsional angles of the residues that had at least one atom within 6 Å from any atom of the ligand. Forty-nine residues in the complex, with 97 free torsional angles, belonged to the active zone, as shown in Figure 6.2 on the next page. Five angles were selected randomly at each Monte Carlo step to be changed by a random value between -30° and +30°. The acceptance rate, which is the ratio of accepted steps to trial steps, was 17 % at 600 K and reduced to 10 % at 310 K. The step size and the acceptance ratio are related to each other: the probability of accepting small steps is larger than the probability of accepting large steps. The maximum step size of 30° corresponds to stepping from the minimum energy conformation to the top of the rotation barrier of a single bond. Larger moves, that would directly transfer the system from one minimum to another and thus enhance the sampling, were not tested in this case, because that approach seemed unlikely to succeed in a packed protein environment. Interestingly, the acceptance rate was only about 4 % in trial runs with a smaller active

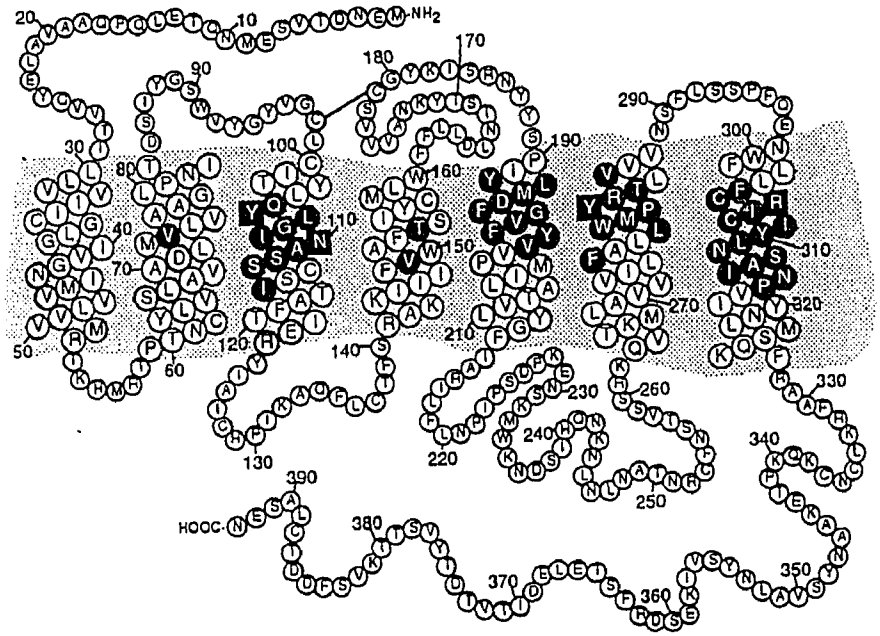


Figure 6.2. The active zone. Residues that are Monte Carlo active are shown black, and the four residues that guided the docking are in boxes.

zone. The small active zone is not spherical as the large one is, but some MC-active residues are surrounded by MC-inactive residues. Most probably MC steps are rejected more often because simultaneous change of several dihedral angles is not possible in a situation like this.

A weak constraining force was put between Y106 and the side chain CO of pGlu of the ligand. The force was zero when the Hbond is 2.0 ±0.5 Å and increased outside this range with a harmonic force constant of 50 kJ/Å^2. Distance monitors were set between Y106, N110, Y282 R306 and TRH, and angle monitors for all the torsional angles of the ligand. The time step in the simulations was 1.5 fs, and MC and SD steps were performed in 1:1 ratio.

The non-bonded cutoff distance was increased to 30 Å both for van der Waals and charge-charge interactions from the MacroModel default values of

7 Å and 12 Å, respectively. Thirty Ångströms is the upper limit for the largest distance between any two atoms in the ligand in an exended conformation. Without the increase, the simulations exit spontaneously because of overheating. This problem could in principle also be corrected by updating the list of non-bonded interactions every time a move is tried, and before it is accepted. This kind of updating scheme is not currently available in MacroModel. The present algorithm updates the list after a MC step is accepted. The surroundings of the complex were described by a distance dependent dielectric, $\varepsilon = r$, which is default for AMBER* parameters in MacroModel.

6.3 Results

Three different aspects are evaluated from the simulations: the methodological questions of the performance of the mixed mode method with a cooling scheme for this system, the evaluation of the structure in general, and the main question, ligand-residue interaction in the binding pocket. In all 15 simulations the energy decreased smoothly. See Figure 6.3 on the next page for an example of the evolution of energy and temperature in one simulation. The temperature fluctuates in the target range and settles to a new temperature within 5 ps of the transition. The energy follows the temperature, and it decreases slightly at each temperature, as is expected for a Metropolis simulation that selects preferentially low energy structures. In all the simulations the helix bundle stayed packed together, and no systematic unwinding was observed. In addition to overall stability, the demands of conformational searching were also met: the end results of the 15 mixed mode simulations differ structurally from each other, both regarding the contacts between the ligand and the receptor, and the exact form of the helices.

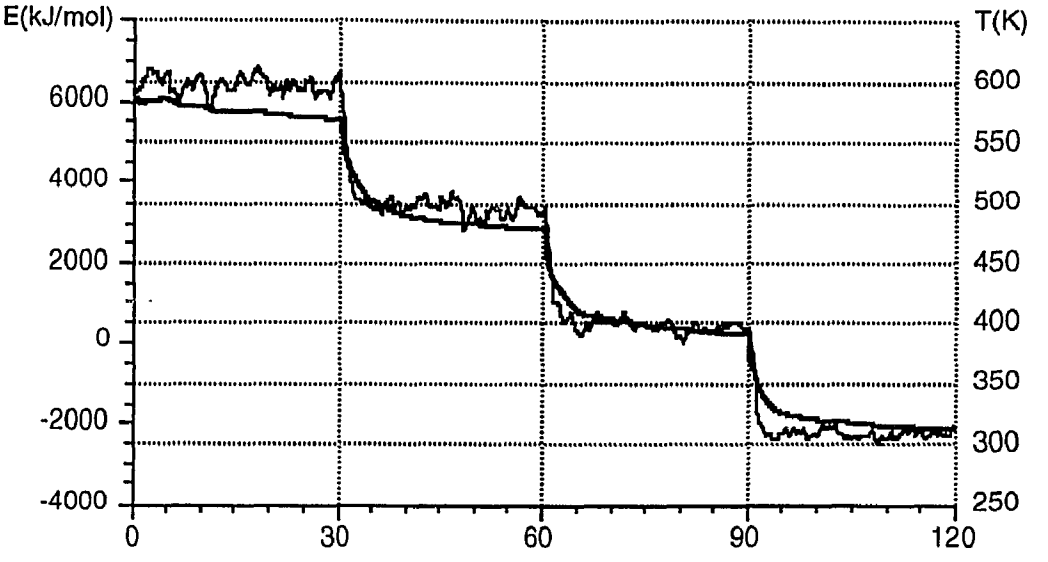


Figure 6.3. Total energy (thick line) and temperature (thin line) of the ligand-receptor complex as a function of time (ps) in one simulation.

The simulation protocol. The reason for choosing the mixed mode method and for applying heat to the system was to ensure good sampling of the conformational space in the complex. The performance of the procedure was evaluated from temperature dependence of fluctuations in ligand dihedral angles, and from the number of major changes in their values at different temperatures. There are six free dihedral angles in TRH. Their values are shown in Figure 6.4 on the next page. Panel a) shows the angular distributions combined from all 15 simulations at 310 K, panel b) shows the same distributions from one simulation, combined from all temperatures, and panel c) shows the values from the same simulation as b) but only at 310 K. Comparison of the three highlights the importance of the elevated temperatures, and of the multiple runs. Panels a) and b) cover the same angular space, although with different probablities. This

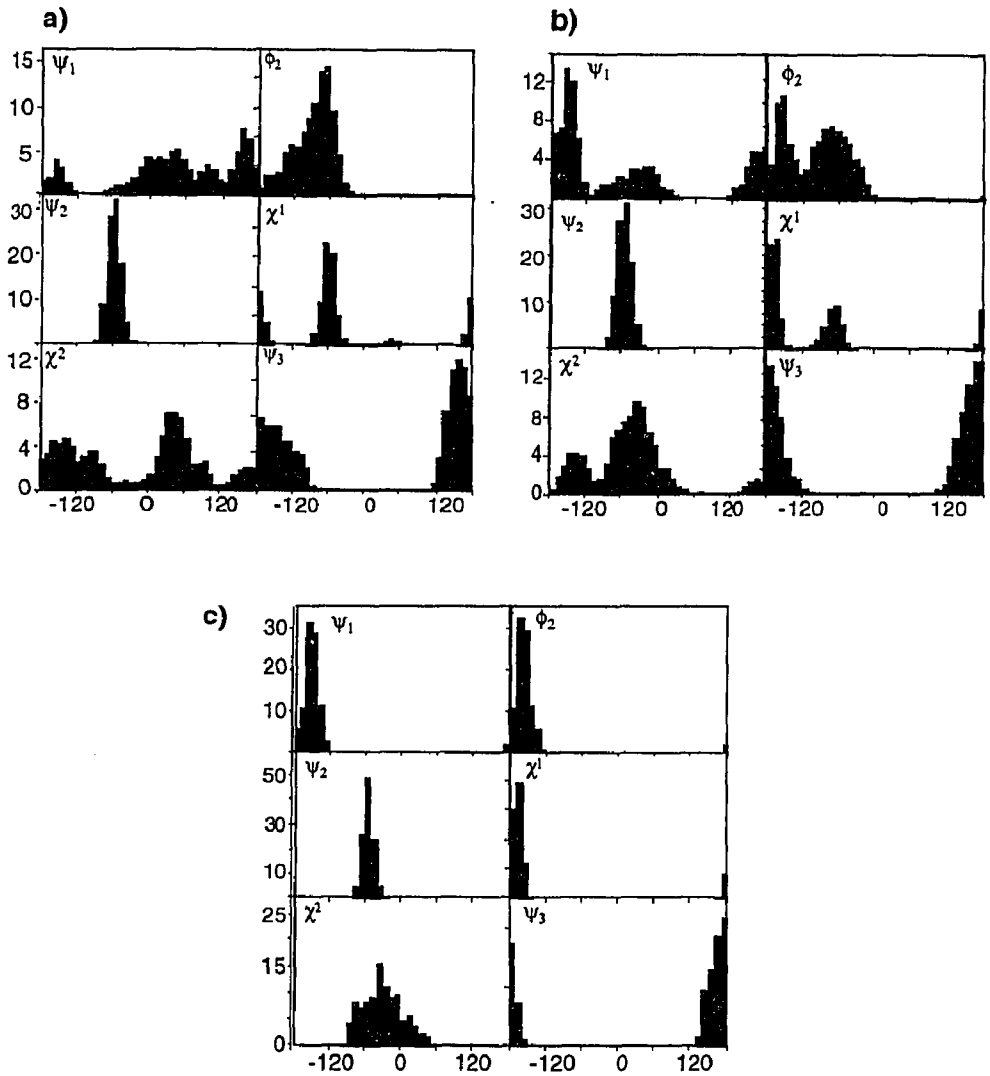


Figure 6.4. Torsional angle distributions of TRH inside the receptor model. The value of the angle in degrees is shown on the x-axis, and the fractional populationon the y-axis. The angles have been monitored every 1.5 fs, and for each 30 ps simulation at a given temperature there are 2000 values. Panel a) represents 30 000 structures combined from 15 simulations at 310 K, b) 8000 structures from one simulation at all four temperatures and c) 2000 structures from the same simulation as shown in previous panel but only at 310 K.

indicates that all the conformational space that can be sampled in hot temperature is also sampled in the combined runs at 310 K, although not with the same probabilities. Based on this comparison, 15 runs provides adequate sampling of the binding pocket. Comparing panels a) and c) shows the effect of the high temperature simulations: results of one run at 310 K do not come even close to the full, combined distributions. The extra freedom at high temperatures has produced sufficiently different starting structures for 310K runs, so that the results cover the entire available conformational space.

The side chain angles $\chi 1$ and $\chi 2$ show the trimodal and bimodal behavior that is expected. The only free angle of pGlu, $\psi 1$, is distributed all over the angular space, and the free angle of Pro, $\psi 3$, shows a broad distribution around 180° and a tiny population at 60°. Only the backbone torsions $\phi 2$ and $\psi 2$ of His do not show free rotation but are surprisingly confined to *gauche*$^-$.

The total number of large changes ($\geq$ 60°) in the dihedral angles in all the runs is 53, and the vast majority of them (42) occur at the high temperatures of 500 and 600 K. See Figure 6.5 on the next page for a representative plot of the evolution of torsional angles. The amount of fluctuations caused by the temperature alone is seen in the behavior of ϕ_1. This angle is inside the pGlu ring, and is not active in the MC steps. The other three angles, ψ_1, χ_1 and χ_2 are MC active, and show distinctly different rotameric states. For example, χ_1 moves from *gauche*$^-$ to *trans* at about 45 ps. Application of heat clearly enhances the conformational mobility of TRH, although still at 310 K there are large changes. A similar conclusion is reached from the study of fluctuations of the torsional angles, that decrease with cooling temperatures. For example, values for ϕ_2, ψ_3, χ_1 and their standard deviations are -73 ± 18, -78 ± 18, -87 ± 25 at 600 K, respectively, and -59 ± 10, -106 ± 14, -60 ± 7 at 310K. The effect is not always large but it is consistent, and totally independent of the direction of the change.

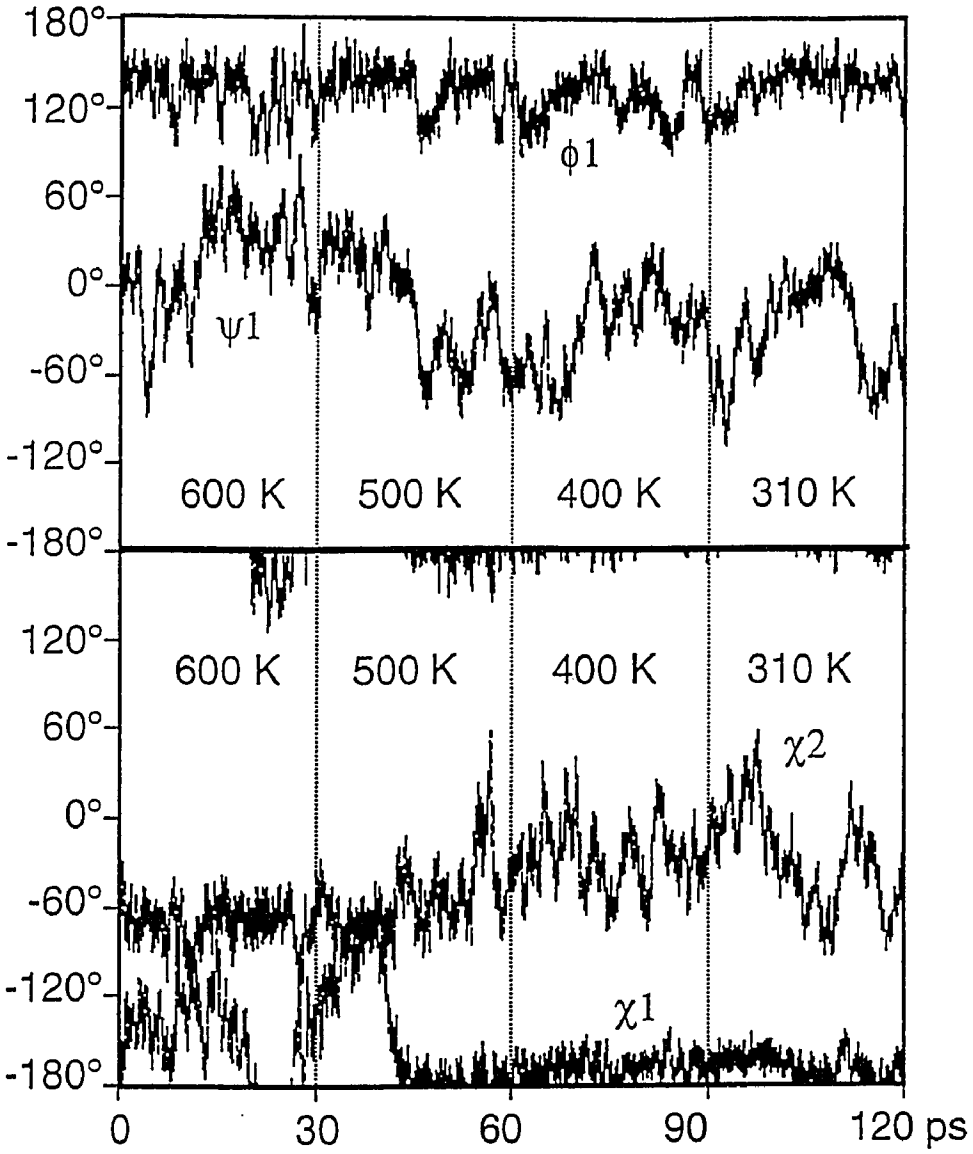


Figure 6.5. An example of the evolution of four torsional angles in the course of a simulation.

The general structure of the receptor. The integrity of the complex was evaluated by measuring the volume of the final structures. They were very similar, all within 1 % of the mean. All final structures had decreased in volume compared to the starting structure, on the average by 7 %. The same behavior

was seen in the molecular dynamics simulations of the receptor alone (See Chapter 5), and is most probably caused by the lack of a realistic environment. For structural evaluation of the helices, the number of helical Hbonds in the final structures was calculated at 310 K. On the average, helices had 87% of the Hbonds of the minimized starting structure.

Despite the conservation of the volume and the good helicity, overall changes are observed in the structure. The tops and bottoms of the helices fray or unwind non-systematically, which is not alarming for the study of the binding pocket. A different behaviour is shown by the helix ends that tilt towards the bundle and never away from it, as if to minimize the exposed area. This is also most probably caused by the non-realistic representation of the environment. The helix with most changes is helix 4, which is lying on the outside of the bundle and has strong positive charges at the intracellular end. Elimination of these charges, that presumably interact with the polar headgroups of the lipid would likely stabilize helix 4.

The binding pocket, on the other hand, is inside the protein, and does not deform greatly. The ligand moves relative to the helices (See Figure 6.10) but it stays inside the binding pocket in all runs, which is expected, because the packing around TRH is tight and there is always van der Waals attraction to keep it inside the receptor. Also the fact that there is no explicit solvent that could satisfy the polar groups of TRH, if not connected to the receptor, favors the complex. It is thus not possible to try to simulate expulsion of TRH from its binding site, for example by protonation of His, without a more detailed environment.

Binding pocket. For the analysis of the binding pocket the ligand is divided into four parts, so that each side chain from Cα on forms a separate group, and the backbone makes up a fourth group. The carboxy terminus is considered a part of proline. See Figure 6.6 below.

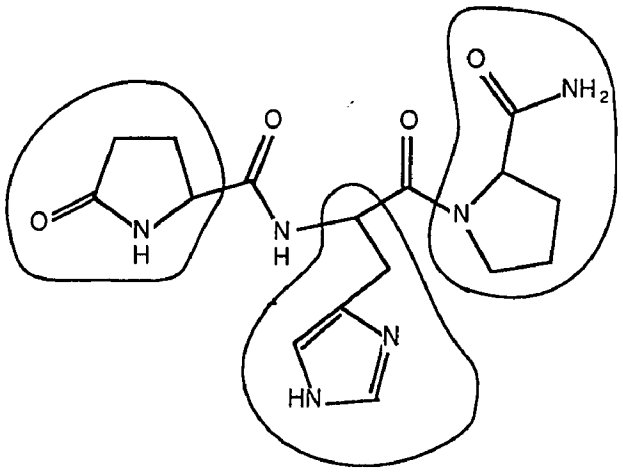
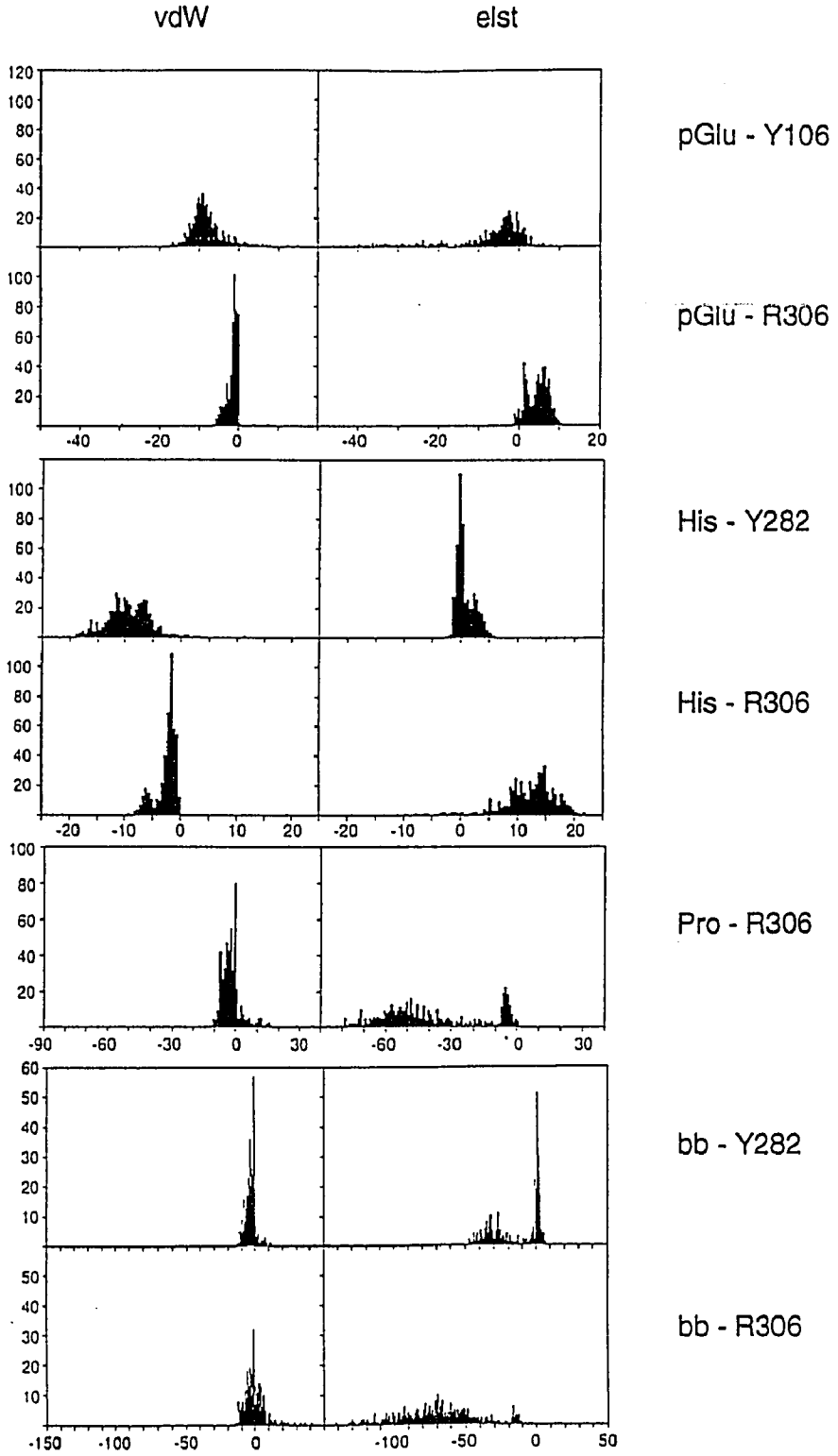


Figure 6.6. TRH divided into four fragments for analysis.

This is analogous to the experimental testing of ligand-receptor interactions, where each side chain can be modified separately but changes in the backbone are either impossible or hard to make. The nonbonded interactions are divided into van der Waals and electrostatic interactions. Figure 6.7 on the next page shows the distribution of the pairwise interaction energies between the groups in TRH and Y106, Y282 and R306 in all simulations at 310 K, and Table 6.1 on page 14 shows the average energy values separately for the fifteen simulations. The energies from Y106 to His, Pro or backbone are not included, because they are smaller than 0.5 kJ/mol.

Pyroglutamic acid and tyrosine 106. Interactions to pGlu in the simulations are straightforward: in all of the runs, pGlu is close to Y106, and none of the other residues of TRH has any significant interactions with Y106. However, the Y106-pGlu Hbond, suggested by experimental data and used to guide the docking, stays intact only in 4 out of 15 simulations. In the rest of the cases D195 from the upper part of helix 5 forms a strong ionic Hbond with Y106 and supplants the original Y106-pGlu interaction. Asp 195 is bound elsewhere in the runs that keep the Y106-pGlu Hbond. The interaction partners of D195 vary,

Figure 6.7. Combined distributions of pairwise interaction energies at 310 K.



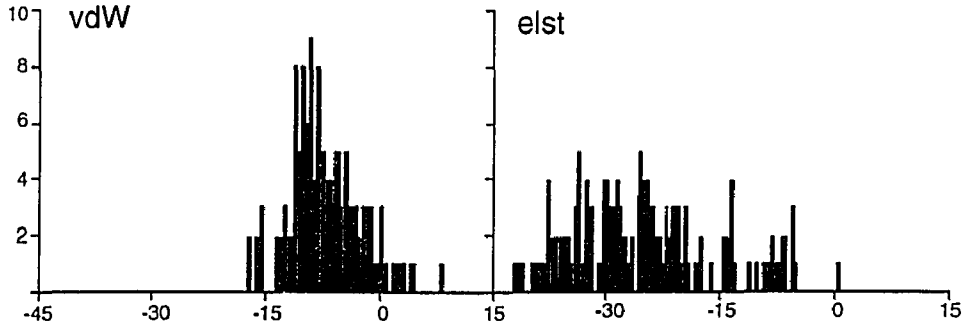
# of run	Y106 - pGlu vdW	elst	R306 - pGlu vdW	elst	Y282 - His vdW	elst	R306 - His vdW	elst
1	-7.8	-4.3	-0.8	1.7	-9.16	-0.0	-1.2	11.2
2	-4.4	-1.3	-1.6	5.1	-9.30	-0.1	-2.0	13.7
3	-6.9	-23.0	-0.5	1.7	-11.2	1.5	-1.7	9.4
4	-4.6	-1.9	-0.3	-0.6	-12.0	2.6	-0.7	5.8
5	-9.70	-7.6	-4.0	5.7	-6.25	-0.2	-2.9	18.1
6	-12.9	-1.9	-1.6	7.0	-11.3	-0.4	-2.2	14.3
7	-7.8	-3.0	-1.0	4.2	-12.2	1.7	-1.7	14.3
8	-7.9	-5.7	-1.0	5.3	-13.9	-0.8	-2.2	14.6
9	-1.3	-33.9	-0.2	1.2	-6.44	0.3	-6.7	0.4
10	-9.5	-3.9	-1.5	7.0	-9.06	-0.4	-2.9	13.0
11	-9.8	-11.2	-3.4	8.0	-7.22	3.8	-0.6	9.4
12	-9.8	-15.5	-1.6	4.3	-4.51	1.5	-1.7	12.3
13	-10.0	1.3	-2.9	5.7	-12.0	2.8	-5.4	9.9
14	-11.6	-4.7	-3.1	7.4	-10.1	-0.9	-1.9	16.8
15	-6.8	-4.9	-1.9	2.7	-9.4	-0.4	-5.7	16.8
ave	-8.1	-8.1	-1.7	4.4	-9.6	0.7	-2.6	12.0
stdD	3.0	9.44	1.2	2.6	2.6	0.4	1.9	4.6

# of run	R306 - Pro vdW	elst	Y282 - bbone vdW	elst	R306 - bbone vdW	elst
1	-1.9	-65.5	-4.5	-3.4	-1.0	-63.3
2	-4.3	-49.3	-1.5	0.7	4.6	-95.1
3	-1.6	-49.2	-1.1	-30.8	-2.2	-44.0
4	-0.6	-3.0	-3.0	-29.2	-1.3	-15.1
5	3.4	-64.7	-2.6	2.4	-5.3	-65.8
6	-0.6	-50.8	-8.5	-1.1	-5.6	-59.9
7	-0.0	-49.8	-4.8	-33.3	1.0	-66.1
8	-2.1	-6.0	-6.2	2.2	5.0	-111.9
9	-5.6	-27.2	-0.7	0.5	0.4	-70.2
10	-4.1	-43.1	-2.8	1.5	-1.6	-94.3
11	-0.8	-5.5	-3.8	-29.9	-1.3	-65.6
12	-5.4	-30.3	-7.2	-24.7	4.7	-96.5
13	-2.2	-55.9	-3.4	-33.2	-6.0	-53.6
14	-3.3	-32.9	-3.2	1.6	-1.6	-93.6
15	-1.0	-52.8	-0.8	-0.4	-8.0	-75.8
ave	-2.0	-39.1	-3.6	-11.8	-1.2	-71.4
stdD	2.3	20.8	2.3	15.7	4.0	24.5

Table 6.1. Average pairwise interaction energies at 310 K per simulation, and the average and standard deviation for all fifteen. Elst means electrostatic, vdW means van der Waals interaction. The first part shows interactions to pGlu and His, the second part to Pro and backbone.

and include W150, T153, backbone NHs of G197, V198 and F199, and both backbone and side chain NH of pGlu. When the side chain of D195 was manually pushed aside to interact with R283, all four subsequent test simulations had strong Y106-pGlu Hbonds. See Figure 6.8 for the Y106-pGlu pairwise interaction energies for the test runs. All the other energy distributions are similar to those shown in Figure 6.7. Previously the mean of pGlu-Y106 electrostatic interaction energies was -8.1 kJ/mol, and now it is -25.7 kJ/mol. The choice of R283 to bind D195 is not based on data, but as an ionic interaction it is the strongest possible to be formed, and sure to keep D195 bound. The test implies that D195 cannot be free and unbonded close to Y106 but it must have a good Hbonding partner, either some residue or a water molecule. There are no data on intra-receptor interactions in this area. Y106 and D195 are Hbonded to each other in the minimized unoccupied receptor model, as shown in Figure 5.6.

Figure 6.8. Interaction energies between pGlu and Y106 in the four test simulations at 310 K with D195 bound to R283.



There is also van der Waals attraction between pGlu and Y106. This arises most often from interactions between the sides of the rings.

Histidine of TRH. The experimental data on histidine of TRH is less clear to interpret, *e.g.* no Hbonds can be proposed, but there is strong indication of some kind of interaction between His and Y282 (M. Gershengorn, unpublished data). Accordingly, Y282 was placed close to His in the starting structure. See Figure 6.7 for the interaction energies to His. The hydroxyl group of Y282 forms an Hbond to the backbone CO of histidine in 6 simulations and this is discussed later in context of the backbone interactions. A more persistent but weak interaction is a van der Waals attraction between the rings of Y282 and His, present in all the simulations. The interaction energy is not large, only about 10 kJ/mol, but it is by far the largest attraction between the His side chain and the receptor. The relative positions of the two aromatic rings varies among the runs, from parallel stacking to angled overlap.

Proline of TRH. Proline CO is Hbonded to R306 in 12 out of 15 runs. See Figure 6.7 for the distribution of the pairwise interaction energies in individual structures at 310 K, and Table 6.1 for the average energies of the different simulations. In one of the three structures without R306-Pro interaction, R306 lies 6 Å above the ligand and does not have any contacts to TRH, but binds to Q105 and Y192. In the two other structures R306 is close to TRH and forms Hbonds to the backbone, but the terminal carboxamide is turned away and interacts with serines from helix3 and helix 7 (S112, S113, S313). In the structures with the R306-Pro interaction, the pattern of Hbonds varies both in the identity of the hydrogen involved and the number of Hbonds from one to a maximum of four. The average energies of simulations vary between 27 to 65 kJ/mol. The largest values in individual structures correspond to very short NH···O=C distances (shortest less than 1.5 Å), and systematically occur with repulsive van der Waals energies.

Backbone of TRH. Most of the attention in structure-activity studies on TRH has concentrated on the side chains. In a small peptide the backbone is expected to be mostly exposed, and can be involved in binding to the receptor.

The only free backbone torsion of pGlu is $\psi 1$. Its values cover practically 360°, with a small disallowed region around -100°. This agrees with the proposed free rotation of the $\psi 1$ angle (Feeney 1974) but poses specific questions about the role of interactions between pGlu and N110, as will be discussed later.

The backbone dihedral angles of His populate only one conformation, as mentioned earlier and shown in Figure 6.4. The distribution of ϕ_2 is wider than that of $\psi 2$, -89.2° ± 28.3° and -57.3° ± 13.1°, respectively. The backbone of the receptor bound TRH is thus nearly α-helical. This differs clearly from earlier experimental data (Donzel 1974; Vicar 1979) for free TRH showing it in the extended conformation. The idea that the ligand would bind only in a specific conformation is not new but there are no previous data, experimental or computational, about the bound conformation of TRH. The validity of the proposed binding conformation is naturally dependent on the accuracy of the 3D molecular model of the receptor.

The simulations show two recurring Hbonds partners for the backbone carbonyl groups: Y282 and R306. The hydroxyl group of Y282 binds to the CO of pGlu in six of the simulations with an average strenght of 30 kJ/mol. The average electrostatic interaction energy in the rest of the runs is 0.4 kJ/mol. The van der Waals attractions of the Hbonded and non-bonded cases are indistinguishable and small.

Hydrogen bonds between the backbone carbonyls of TRH and R306 are the most often occurring feature in the simulated structures. Only one simulation does not have a Hbond between the two. Instead, most simulations show two Hbonds between His C=O and R306, and some show the guanidino group in-

teracting with each of the backbone C=O groups of the ligand. See Figure 6.9. The energetic values of the interaction is 40-110 kJ/mol, which should be easy to detect, if analogs with modified backbone were available.

Three good structures. The strongest experimental data regarding Y106, Y282 and R306 were used as the final criteria for choosing structures to study more closely. Three simulations had these interactions simultaneously, and they were annealed by 20 ps steps of 20K down to 10 K to eliminate large fluctuations. The three structures are referred to as 1, 2, and 3, and two of them are shown in Figure 6.9. Structure 1 in Figure 6.9 a) has a 1.7 Å/ -30 ± 1.4 kJ/mol Hbond between Y106 and pGlu C=O. N110 is lying under the Y106-pGlu pair, and its terminal amide NH_2 makes an Hbond to Y106 OH. The distance from N110 side chain C=O to pGlu NH is 2.6 Å, which is long for an Hbond. The dipoles of N110 and pGlu are nearly perfectly matched. D195 interacts with the backbone NH of G197, V198 and F199, and does not interfere with the Y106-pGlu interaction. The histidine of TRH lies under Y282 and their van der Waals interaction energy is -13.7 ± 0.3 kJ/mol. The OH of Y282 makes an Hbond with the backbone C=O of pGlu, and also to Y192. The His side chain is surrounded by L274,W279, N312 and S313. The carboxy terminus of proline is Hbonded to R306, and R306 also contacts Y192. The hydrophobic part of the side chain of proline, $C_\beta C_\gamma C_\delta$, sits tightly packed between the side chains of I109, I116 and S113. Mutations of these residues to alanines should help binding of crowded analogs, for example MePro-TRH. Structure 1 differs from the two other cooled down structures regarding the interactions from the terminal NH_2. In structure 1 the carboxy terminus donates Hbonds to D71, S112 and Y310, while in structures 2 and 3 the Hbonds from NH_2 are with S313 and Y310. The latter two residues are being mutated.

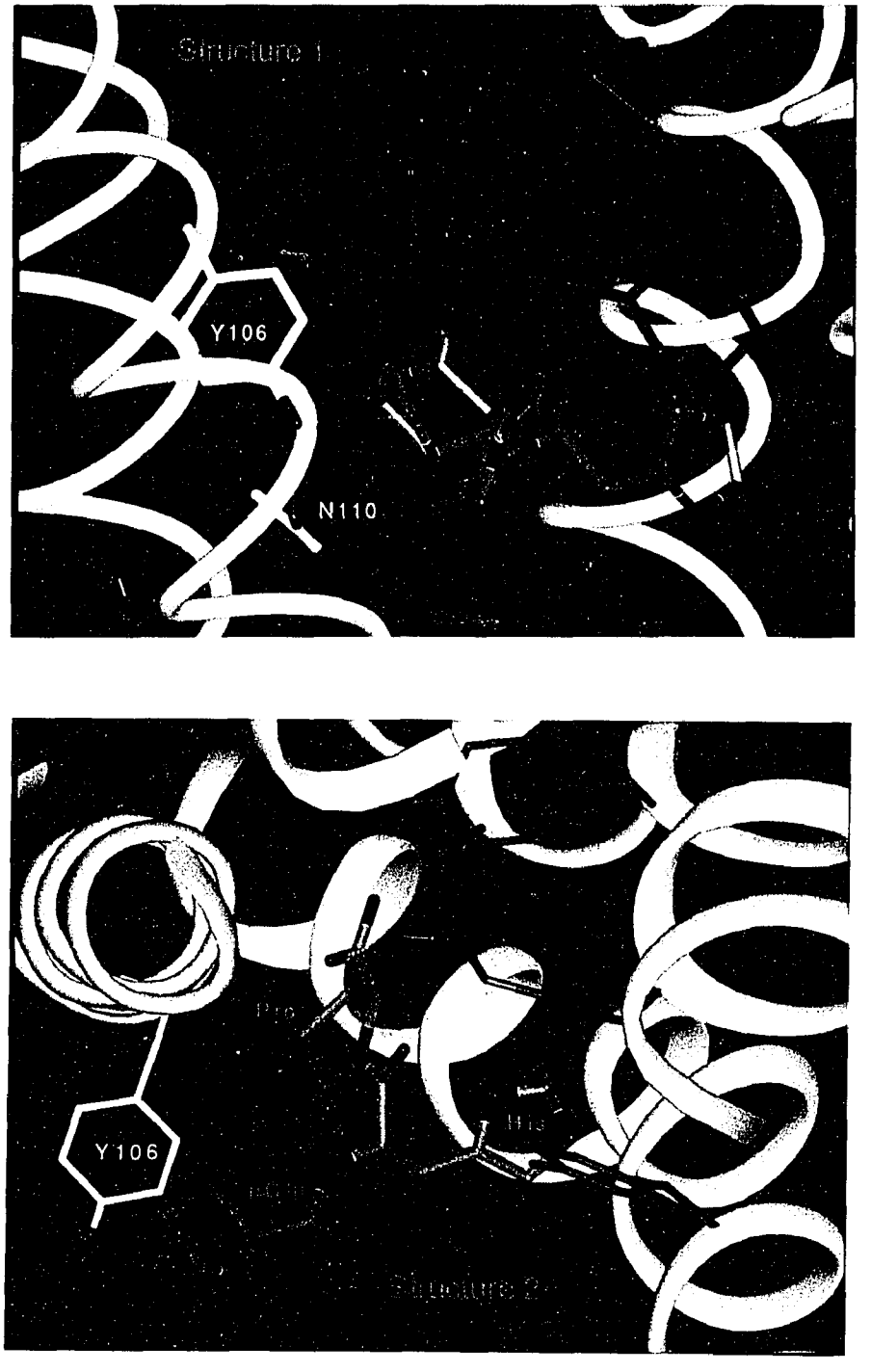


Figure 6.9. Binding site of structures 1 and 2. In structure 1, R306 forms three H-bonds with the peptide backbone. Y282 contacts the backbone in 1 but not in 2.

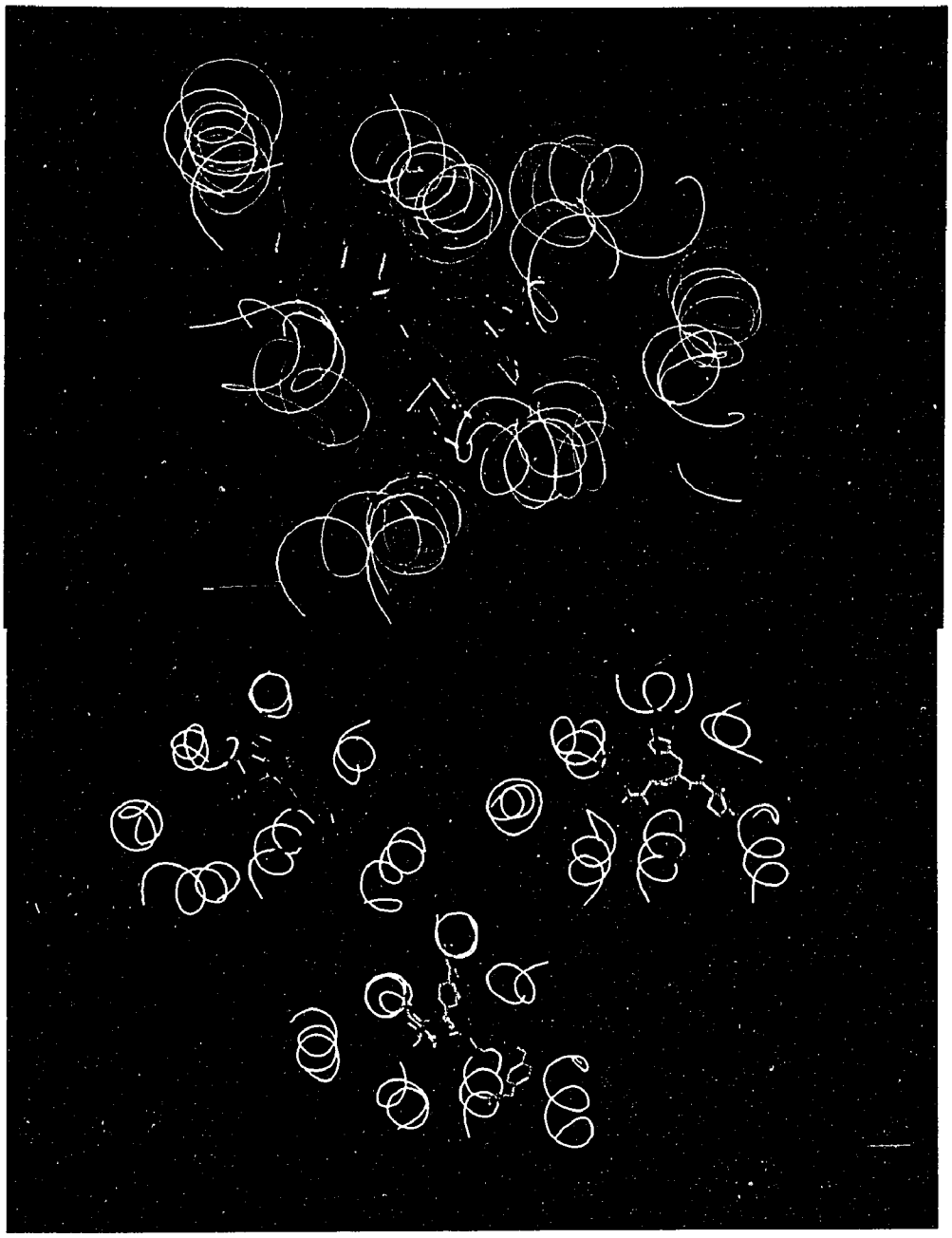


Figure 6.10. Top view of the three final complexes at 10 K. The structures are superimposed by the Cα-trace of the helices. Upper picture shows the structures together, with pGlu in blue, His in red, $ProNH_2$ in yellow, Y106 in orange, Y282 in green and R306 in purple. TRH and the contact residues are in different positions in each structure but the interactions are the same. Lower picture shows the structures separated. The colors serve only to identify structures from each other.

6.4 Discussion

The mixed mode simulations have proved to be a valuable tool in the study of large systems. The simulations produced several different structures of the TRH-TRHR complex with different ligand-receptor interactions, implying that the conformational search was effective. Methodologically, a clear improvement for the present calculations would be to increase the simulation time by cooling the system in smaller temperature steps. This might allow for a better rewinding of the helices and the simulations could be started at higher temperatures. A harder to reach improvement, although more critical, would be the inclusion of a more realistic model of the lipid-water environment. This would also make the addition of the loops to the helix bundle more relevant, that in turn should eliminate the fraying of helix ends, and damp their movements. At this stage of the accuracy, it is satisfying that among the results there are structures that comply with all available experimental data, and can thus be used as the starting structures for further mutational search of new possible contacts within the complex.

The best studied ligand-residue interaction between TRH and its receptor is the Y106-pGlu Hbond. Figure 6.11 on the next page, that shows the evolution of interactions between TRH and residues 106, 282 and 306 along one simulation, highlights the special character of this Hbond. Only pGlu has any significant interactions with Y106, none of the other parts of the ligand come close to it. This agrees both with physical and biological data. Several different NMR experiments have shown that pGlu is conformationally independent from His-ProNH$_2$: Substitutions made in pGlu do not affect the spectrum of His-ProNH$_2$, and vice versa (Deslauriers 1973; Donzel 1974); relaxation times for pGlu are faster than for the rest of the molecule (Deslauriers 1973; Feeney 1974), which is interpreted as fast rotation around $\psi 1$. Results from biological experiments

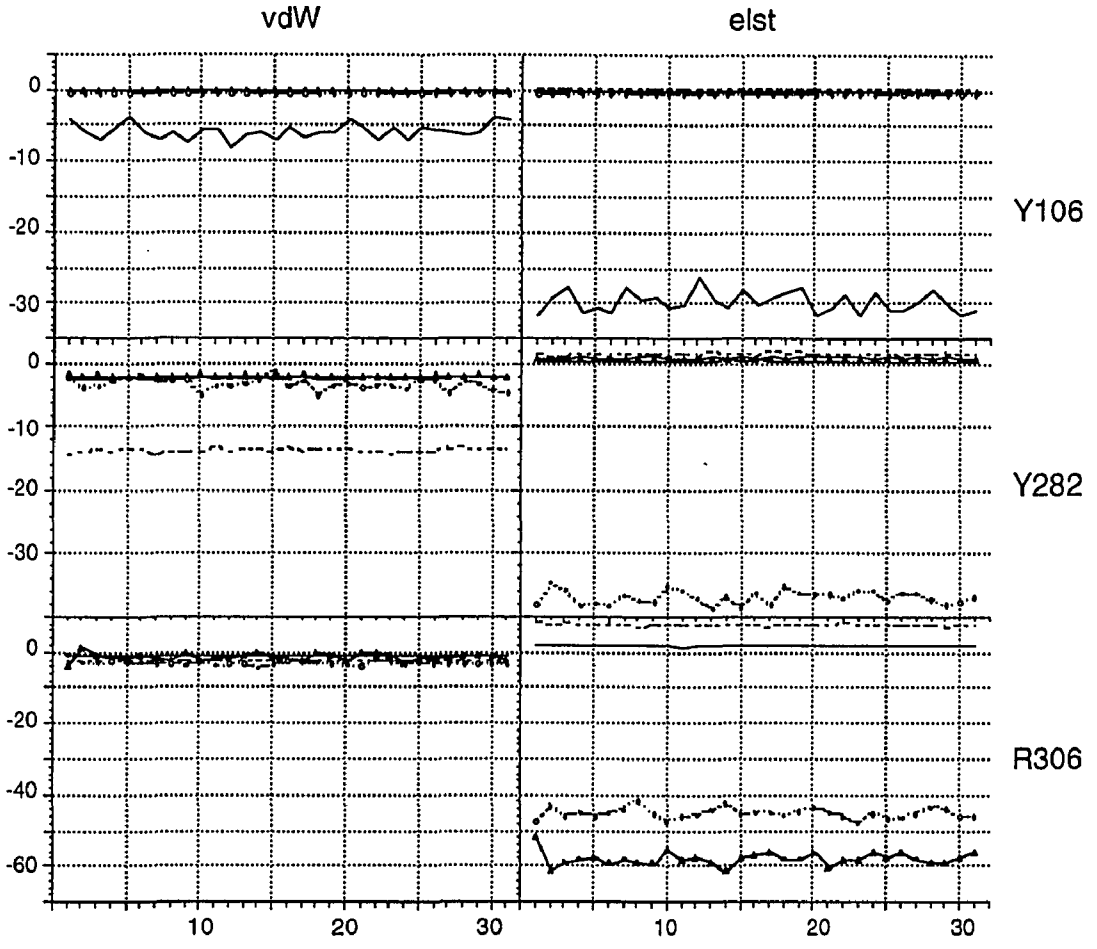
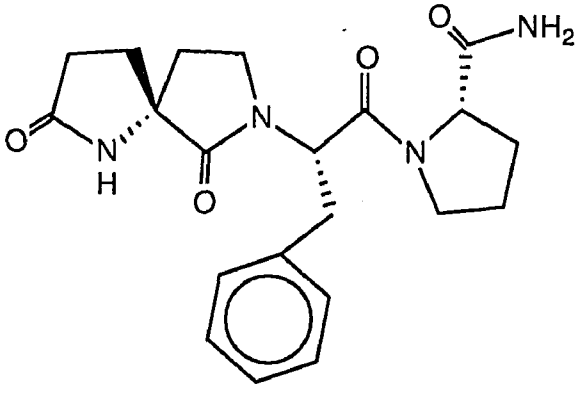


Figure 6.11. Nonbonded interactions between TRH and the three important residues in the receptor in one simulation at 10 K (structure 1). The y-axis shows energy in kJ/mol, and the x-axis time in ps. Left hand side shows the van der Waals interactions, and the right hand side the electrostatic interactions. Solid lines refer to pGlu, dashed lines to His, solid lines with black triangles to Pro, and dashed lines with open circles to the backbone. pGlu interacts only with Y106, and vice versa. Histidine shows vdW attraction to Y282 and electrostatic repulsion to R306. In this simulation Y282 Hbonds to the backbone. R306 forms Hbonds both with the backbone and ProNH_2. Interactions are similar at 310 K, but the fluctuations are much larger at room temperature.

also suggest that modifying pGlu-Y106 interaction induces strictly local changes in the receptor (Perlman 1994). The hypothesized Hbond between Y106 and pGlu CO was eliminated both by Y106F mutation of the receptor, and by $pGlu^1$ to Pro^1 change in the ligand. Both experiments produced the same effect, and more importantly, testing Pro1-TRH on Y106F mutant caused only a small additional increase in EC_{50}. Similar reciprocal studies on N110 produced smaller but still non-additive effect. It was deduced that N110 and Y106, just one helical turn apart from each other, would both form Hbonds to pGlu. This case is among the simulation results but it is not the dominating structure: most often N110 is attracted towards the intracellular part of helix 4. The coexistece of the two H-bonds can also be seen from the behavior of ψ_1 of pGlu. In these simulations ψ_1 covers 360 degrees. Only values around 160° are compatible with simultaneous interactions of pGlu to both Y106 and N110. A compound with constrained backbone at ψ_1 (see below) has been synthesized and tested for binding and activation on the TRH receptor (K. Moeller and M. Gershengorn, unpublished results). It behaves similarly to Phe^2-TRH, with K_d=290 nM and EC_{50} = 4400 nM. Both values correspond to 35 % of F^2-TRH affinity and activity. This indicates that the additional five membered ring does not eliminate any conformations or interactions needed for binding to the receptor. The spiro compound with the specified stereochemistry (see below) can exist only in a conformation



that has ψ_1 ~ 137°, which is compatible with the simultaneous formation of Hbonds to Y106 and N110. These data corroborate the existence of rather weak Hbond proposed by mutation experoments. Formation of two Hbonds from the same helix to neighboring functional groups in the ligand requires a very exactly defined organization of the binding pocket around pGlu.

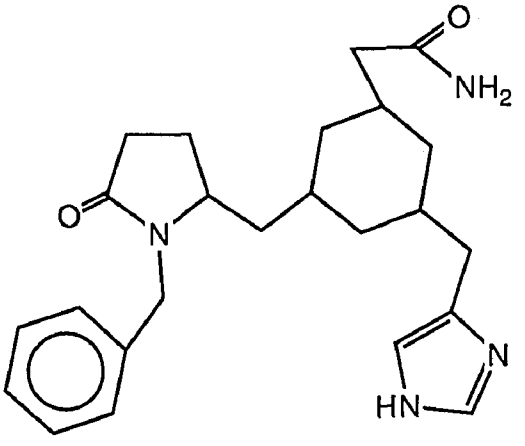
The important role of D195 in the simulations demands further experimental studies. If the aspartate indeed competes with the ligand for an Hbond to Y106, it could be mechanistically important in the release of TRH from the receptor. Experiments to test the effect of D195 mutations on activation are in progress.

A van der Waals attraction, that depends on the surface area and not on the identity of atoms, is not contradictory to the structure-activity data that portrays His as the only permissive position of TRH (Vale 1973; Szirtes 1984). Substitutions of His by a variety of other residues reduce activity, but no clear structural trend can be discerned. The situation is complicated by certain pairs of analogs, that have minimal structural differences,*e.g.* 1-MeHis and 3-MeHis, but differ by orders of magnitude in binding. A preliminary computational study suggests that the two analogs populate different conformations of ω_2 (unpublished results), which might cause the different affinity to the receptor.

Experimentally, the effect of Y282F mutation is an 8.3 fold reduction of affinity (M. Gershengorn, unpublished data). The interaction can not be tested by eliminating the backbone CO of the ligand, because an analog with reduced backbone is not available. An 8.3 fold reduction corresponds to loss of 5 kJ/mol in binding energy, much less than the 30 kJ/mol that is seen as the strength of the Hbond in the simulations. It should, however, be kept in mind that the numbers are not directly comparable. The number from the simulations measures directly the pairwise interaction energy between the two groups in question. The

experimental number includes everything that went into the formation of the complex: desolvation of the ligand, and breaking of any possible intrareceptor interactions of this residue in the absence of the ligand. In the unoccupied model Y282 is Hbonding to Y192. Thus, even if there is a large energetic gain in forming the Y282-backbone Hbond, a lot of energy also has to be invested into breaking the hydrogen bonding network of Y282 that existed without the ligand. In the case of Y282F, no such Hbonds exist, and the net difference of Y282 and F282 is reduced. There might also be kinetic consequences on binding from the fact that a Y282-backbone Hbond keeps the Y282 ring closed over His side chain.

In addition to the experimentally suggested Hbond between R306 and the carboxy terminus (Perlman 1995), the simulations also show several strong Hbonds between R306 and the backbone. If these Hbonds really exist, they do restrict the backbone mobility severely, and might be the most critical element in determining the conformational requirements of the bound ligand. Recently, a nonpeptide analog of TRH with reduced flexibility was described (Olson 1995). The structure is based on a cyclohexane. It has a modifed pGlu residue, and an imidazole side chain, and no backbone carbonyls (see below). Testing this



analog onTRHR and R306A mutant receptor should show the importance of the backbone interactions, if the benzyl group attached to pGlu does not interfere with binding.

Studying the ligand-receptor complex with a simulation method in contrast to only minimizing the docked structure offers the possibility to assess the flexibility of the binding pocket, and the mobility of the ligand. The simulations also make it possible to generalize the discussion on the binding site from residue-residue interactions into the physical nature and energetics of these interactions. The three residues of TRH and the three experimentally identified key residues of the receptor show three different kinds of interaction, characterized by their magnitude and directional requirements. The weakest and least ordered interaction is the van der Waals attraction between His and Y282. The magnitude of the attraction correlates with the shortest distance between the rings, but is not dependent on the exact identity of neighboring atoms. The attraction is small, 10 kJ/mol, but not easy to eliminate totally. Compared to this, the polar Hbond between Y106-pGlu is often broken, and totally dependent on the existence of Y106 hydroxyl group and pGlu carbonyl group. The simulations give average value of 30 kJ/mol for this interaction. The strongest interaction seen in the complex is the charged Hbond between R306 and Pro. It is nearly unbreakable and omnipresent, with the average attraction of 50 kJ/mol. The generalization of the ligand-receptor interaction from a set of residue-residue pairs into the nature of the functional groups involved is perhaps the most important result from the simulations.

The scenario of TRH binding to its receptor could be as follows: TRH binds initially to the extracellular loops. R306 forms strong, charged Hbonds to the backbone of TRH, which changes the conformation of the backbone from

extended to α-helical. The freely rotatable pGlu can now find its interaction partners Y106 and N110. Y282 sits over His, and may form a Hbond to the backbone carbonyl of pGlu. This corresponds to the two stage binding seen in kinetic experiments (Hinkle 1982). D195 is connected to something else, a water molecule or R283, as shown in the test simulations. The activation event, that requires R283, releases the interactions of D195, which then competes with pGlu for Hbonding to Y106. When the interactions between pGlu and the receptor are broken, TRH is released.

7 Conclusions

Structural hypotheses of receptor recognition and activation require a molecular model of the ligand-receptor complex. Only within an atomic structure can one test the feasibility of hypothesized interactions, both to rationalize experimental results and to project new possible contacts. As long as there are no experimental structures of the G-protein coupled receptors available, the only way to gain structural insights of GPCR function is to use molecular models. However, it is necessary to keep in mind that the models are based on very little physical data, and cannot be considered but good approximations.

The work on TRH receptor has produced a model that can include all the experimental data that is available at the moment. The model has proved to be sturdy enough to survive unconstrained molecular dynamics simulations and hot temperature Monte Carlo simulations. The receptor model should not be considered a final product but a reflection of one stage of knowledge. The several cycles of construction and docking that have preceded the present model are not described. However, it should be noted that even the first model of the complex, that consisted only of helix 3 and TRH in vacuum, was useful by suggesting the Y106-pGlu interaction. The recognition of the continuous need to improve models when new data become available was the reason to automate the model building process.

The model could be improved by adding the intracellular and extracellular loops to the helix bundle. This should stabilize the movement of the ends of helices. Another improvent would be to include a model of lipid membrane and solvent around the loops. A good description of the surroundings would eliminate many possible pitfalls, especially in future studies of the activation mechanism. The deepest uncertainty of the model concerns the relative vertical posi-

tions of the helices, and any data on direct helix-helix interaction that could clarify this aspect would be most welcome.

The mixed mode simulations on the TRH-TRHR complex explored the connections that had been suggested by previous work. A large set of torsional angles were kept free during the Monte Carlo steps, and the sampling of the conformational space of the ligand was very good. The interactions between R306-ProNH_2 and Y282-His were present in most of the simulations, and do not seem to be very sensitive to the exact structure of the binding site. The persistance of the R306-ProNH_2 interaction is caused both by the strength of the ionic Hbond and the flexibility of the R306 side chain. The Y282-His interaction is weaker, but as a van der Waals attraction it does not require exact relative positions of specific atoms, as long as the overlapping contact area does not change. The Hbond between pGlu and Y106 is isolated from the rest of the ligand but strongly affected by the position of the residues around Y106, especially D195. D195 competes with pGlu for the Hbond to Y106. Additional mutations of this residue are in progress. The simulations give an estimate for the mobility of the binding site, and describe the known interactions in physical terms. Interactions from the C-terminal NH to Y310 and S313 in helix 7 are proposed for testing.

The results about the conformation of TRH are intriguing. The real surprises were provided by conformational searches on the new constrained analogs (S)-XTRH and (R)-XTRH. The method of conformational memories (Guarnieri 1995) promises to be very useful in separating conformational effects of the ligand from changed ligand-receptor contacts. The constrained analogs offer a unique possibility to address the bioactive conformation of TRH directly, in contrast to indirect deductions from the crystal structures of analogs of varying activity (Eckle 1985). Because XTRHs are full agonists it can be assumed that they bind

and activate in the same way as TRH does. The compounds are so rigid that they can be assumed to have the same conformation in water and inside the receptor. TRH fits well to (S)-XTRH either in the conformation suggested by the studies in water (Chapter 4) or in the conformation that is observed in the simulations of the complex (Chapter 6). The two studies propose opposite conformations of the second peptide bond, but very similar overall forms for the ligand. The energy difference between the two possible binding conformations of TRH is large, 50 kJ/mol, but this is well within the range of interaction energies calculated in the simulations of the complex. Preliminary studies on the methylated analogs, τ-Me-TRH and π-Me-TRH would suggest *cis*-TRH as the binding conformation. The analogs differ in activity and seem to differ conformationally only in the size of *cis*-population. Obviously, the next set of simulations to do is a mixed mode study of *cis*-TRH inside the binding pocket, and a similar study of (S)-XTRH. See Figure 7.1.

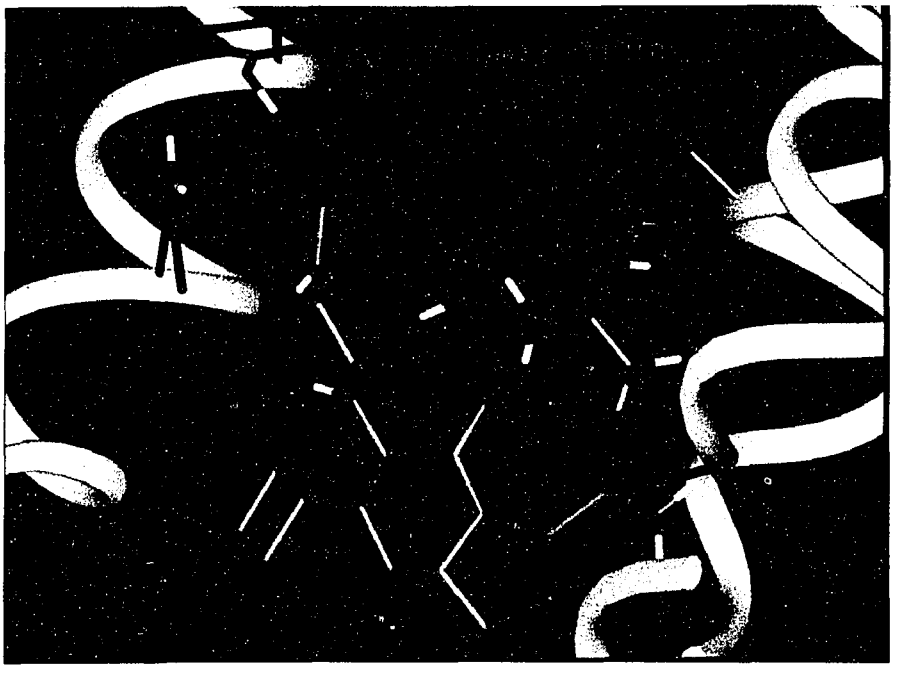


Figure 7.1. A possible structure of (S)-XTRH docked to the TRH receptor, in analogy to the studied TRH-TRHR complex.

The collaborative experimental-computational approach to study TRH and TRH receptor has been very fruitful. When the computational project started, not one single contact site between the ligand and the receptor was known. Three strong interaction sites have been identified, and several other residues are known to affect the binding significantly. The double functionality of pGlu has been elucidated, and the role of the carboxy terminus is known. There is also a reasonable hypothesis to follow for clarifying the complex role of His.

The computational studies on TRH and the TRH receptor have concentrated on the binding event, and opened many new questions for future studies. The emphasis of the results is in the conformational flexibility of amino acids and on the physical description of the observed residue-residue contacts in the binding site.

References

Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. and Weng, J. (1987). Protein Data Bank in Crystallographic Databases. Data Commission of the International Union of Crystallography. F. H. Allen, G. Bergerhoff and R. Sievers. Bonn/Cambridge/Chester: 107-132.

Anteunis, M. J. O., Borremans, F. A. M., Stewart, J. M. and London, R. E. (1981). "360-MHz NMR Conformational Analysis of Gly-Pro-X Peptides (X=Ala, Cha, Phe)." J. Am. Chem. Soc. **103**: 2187-2191.

Aragay, A. M., Katz, A. and Simon, M. I. (1992). "The $G\alpha_q$ and $G\alpha_{11}$ proteins couple the thyrotropin-relaesing hormone receptor to phospholipase C in GH_3 rat pituitary cells." J. Biol. Chem. **267**: 24983-24988.

Ashworth, R., Yu, R., Nelson, E. J., Dermer, S., Gershengorn, M. C. and Hinkle, P. M. (1995). "Visualization of the thyrotropin-releasing hormone receptor and its ligand during endocytosis and recycling." Proc. Natl. Acad. Sci. U.S.A. **92**: 512-516.

Balbes, L. M., Mascarella, S. W. and Boyd, D. B. (1994). A Perspective of Modern Methods in Computer-Aided Drug Design. Reviews in Computational Chemistry. Ed: K. B. Lipkowitz and D. B. Boyd. New York, VCH Publishers. **V:** 337-379.

Baldwin, J. M. (1993). "The probable arrangement of the helices in G protein-coupled receptors." EMBO J. **12**: 1693-1703.

Baldwin, J. M. (1994). "Structure and function of receptors coupled to G proteins." Curr. Opinion Cell Biol. **6**: 180-190.

Ballesteros, J. A. and Weinstein, H. (1992). "Analysis and refinement of criteria for predicting the structure and relative orientations of transmembranal helical domains." Biophys. J. **62**: 107-109.

Ballesteros, J. A. and Weinstein, H. (1995). Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors. Receptor Molecular Biology. Ed: S. C. Sealfon, Academic Press. **25:** 366-428.

Belle, J., Montagut, M. and Bellocq, A.-M. (1972). "Analyse conformationelle de l'hormone hypothalamique TRF de liberation de la thyrostimuline." C. R. Acad Sc. Paris, Series C **275**: 471-474.

Bellocq, A.-M., Castensson, S. and Sievertsson, H. (1977). "Conformational analysis of an analogue of thyroliberin (T.R.F.)." Biochem. Biophys. Res. Comm. **74**: 577-583.

Bellocq, A. M. and Dubien, M. (1976). "Proton nuclear magnetic resonance studies of N_{ε}-Me-His^2 thyroliberin. An hyperactive analogue of thyroliberin (thyrotropin releasing factor)." Biochem. Biophys. Acta **420**: 1-7.

Blagdon, D. E., Rivier, J. and Goodman, M. (1973). "Proposed Tertiary Structure for the Hypothalamic Thyrotropin-Releasing Factor." Pro. Natl. Acad. Sci. USA **70**: 1166-1168.

Bohm, H.-J. (1992). "The Computer Program LUDI: A New Method for the De Novo Design of Enzyme Inhibitors." J. Comp. Aided Mol. Design **6**: 61.

Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C. and Kraut, J. (1982). "Crystal Structures of *Eschericia coli* and *Lactobacillus casei* Dihydrofolate Reductase Refined at 1.7 A Resolution." J. Biol. Chem. **257**: 13650-13662.

Bond, R. A, Leff, P., Johnson, T. D., Milano, C. A., Rockman, H. A., McMinn, T. R., Apparsundaram, S., Hyek, M. F., Kenakin, T. P., Allen, L. F. and Lefkowitz, R. J. (1995). "Physiological effects of incverse agonists in transgenic mice with myocardial overexpression of the β_2-adrenoceptor." Nature **374**: 272-276.

Breneman, C. M. and Wiberg, K. B. (1990). "Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials. The Need for High Sampling Density in Formamide Conformational Analysis." J. Comp. Chem. **11**: 361-373.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. and Karplus, M. (1983). "CHARMM: A Program for Macromolecular Energy, Minimization and Dynamics Calculations." J. Comp. Chem. **4**: 187-217.

Brünger, A. (1988). "Crystallographic Refinement by Simulated Annealing. Application to a 2.8 Å Resolution Structure of Aspartate Aminotransferase." J. Mol. Biol. **203**: 803-816.

Burger, M. T., Armstrong, A., Guarnier, F., McDonald, D. Q. and Still, W. C. (1994). "Free Energy Calculations in Molecular Design: Predictions by Theory and Reality by Experiment with Enantioselective Podand Ionophores." J. Am. Chem. Soc. **116**: 3593-3594.

Burt, A. K., Hutchins, C. W. and Greer, J. (1991). "Predicting receptor-ligand interactions." Curr. Opin. Str. Biol. **1**: 213-218.

Caflisch, A., Niederer, P. and Anliker, M. (1992). "Monte Carlo Docking of Oligopeptides to Proteins." Proteins **13**: 223-230.

Cherfils, J., Duquerroy, S. and Janin, J. (1991). "Protein-Protein Recognition Analyzed by Docking Simulation." Proteins **11**: 271-280.

Cherfils, J. and Janin, J. (1993). "Protein docking algorithms: simulating molecular recognition." Curr. Opin. Str. Biol. **3**: 265-269.

Chothia, C. and Finkelstein, A. V. (1990). "The classification and origins of protein folding patterns." Annu. Rev. Biochem. **59**: 1007-10039.

Chothia, C., Levitt, M. and Richardson, D. (1981). "Helix to Helix Packing in Proteins." J. Mol. Biol. **145**: 215-250.

Choudhary, M. S., Sachs, N., Uluer, A., Glennon, R. A., Westkaemper, R. B. and Roth, B. L. (1995). "Differential Ergoline and Ergopeptide Binding to 5-Hydroxytryptamine$_{2A}$ Receptors: Ergolines Require an Aromatic Residue at Position 340 for High Affinity Binding." Mol. Pharm. **47**: 450-457.

Cockle, S. M., Aitken, A., Beg, F. and Smyth, D. G. (1989). "A Novel Peptide, Pyroglutamylglutamylproline Amide, in the Rabbit Prostate Complex, Structurally Related to Thyrotrophin-releasing Hormone." J. Biol. Chem. **264**: 7788-7791.

Coy, D. H., Hirotsu, Y., Redding, T. W., Coy, E. J. and Schally, A. V. (1975). "Synthesis and Biological Properties of the 2-L-*b*-(Pyrazolyl-1)alanine Analogs of Luteinizing Hormone-Releasing Hormone and Thyrotropin-Releasing Hormone." J. Med. Chem. **18**: 948-949.

Cramer, C., J. and Truhlar, D., G. (1991). "General Parametrized SCF Model for Free Energies of Solvation in Aqueous Solution." J. Am. Chem. Soc. **113**: 8305-8311.

Cramer, C. J. and Truhlar, D. G. (1994). "Quantum Chemical Conformational Analysis of 1,2-Ethanediol: Correlation and Solvation Effects on the Tendency to Form Internal Hydrogen Bonds in the Gas Phase and Aqueous Solvation." J. Am. Chem. Soc.

Crippen, G. M. (1982). "Conformational Analysis by Energy Embedding." J. Comp. Chem. **3**: 471-476.

Crippen, G. M. (1987). "Why Energy Embedding Works?" J. Phys. Chem. **91**: 6341-6343.

Cronet, P., Sander, C. and Vriend, G. (1993). "Modeling of transmembrane seven helix bundles." Prot. Eng. **6**: 59-64.

Dahl, S. G., Edvardsen, Ø. and Sylte, I. (1991). "Molecular dynamics of dopamine at the D_2 receptor." Proc. Natl. Acad. Sci. USA **88**: 8111-8115.

Deisenhofer, J. and Michel, H. (1989). "The photosythetic reaction center from the purple bacterium Rhodopseudomonas viridis." Science **245**: 1463.

DesJarlais, R. L., Sheridan, R. P., Seibel, G. L., Dixon, J. S., Kuntz, I. D. and Venkataraghavan, R. (1988). "Using Shape Complementarity as an Initial Screen in Designing Ligands for a Receptor Binding Site of Known Tree-Dimensional Structure." J. Med. Chem. **31**: 722-729.

Deslauriers, R., Garrigou-Lagrange, C., Bellocq, A.-M. and Smith, I. C. P. (1973). "Carbon-13 nuclear magnetic resonance studies on thyrotropin-releasing factor and related peptides." FEBS Lett. **31**: 59-66.

Deslauriers, R., Smith, I. P. and Walter, R. (1974). "Conformational Flexibility of the Neurophyseal Hormones Oxytocin and Lysine-vasopressin. A Carbon-13 Spin-Lattice Relaxation Study of Backbone and Side Chains." J. Am. Chem. Soc. **96:7**: 2289-2291.

De Lean, A., Stadel, J.M. and Lefkowitz, R. L. (1980). "A Ternary Complex Model Explains the Agonist-specific Binding Properties of the Adenymate Cyclase-coupled β-Adrenergic Receptor." J. Biol. Chem. **255**: 7108-7117.

Di Nola, A., Roccatano, D. and Berendsen, H. J. C. (1994). "Molecular Dynamics Simulations of the Docking of Substrates to Proteins." Proteins **19**: 174-182.

Donnelly, D., Johnson, M. S., Blundell, T. L. and Saunders, J. (1989). "An analysis of the periodicity of conserved residues in sequence alignments of G-protein coupled receptors." FEBS Lett. **251**: 109-116.

Donnolly, D., Overington, J. P. and Blundell, T. L. (1994). "The prediction and orientation of α-helices from sequence alignments: the combined use of environment-dependent substitution tables, Fourier transform methods and helix capping rules." Prot. Eng. **7**: 645-653.

Donzel, B., Goodman, M., Rivier, J., Ling, N. and Vale, W. (1975). "Synthesis and cnformations of hypothalamic hormone releasing factors: two TRF-analogues containing backbone N-methyl groups." Nature **256**: 750-751.

Donzel, B., Rivier, J. and Goodman, M. (1974). "Conformational Studies on the Hypothalamic Thyrotropin Releasing Factor and Related Compounds by ^{1}H Nuclear Magnetic Resonance Spectroscopy." Biopolymers **13**: 2631-2647.

Eckle, E. and Stezowski, J. J. (1985). "Conformational Properties of Central Nervous System Active Thyrotropin Releasing Hormone Analogues: Probing Structure-Activity Relationships at the Molecular Level." J. Med. Chem. **28**: 125-137.

Eisen, M. B., Wiley, D. C., Karplus, M. and Hubbard, R. E. (1994). "HOOK: A Program for Finding Novel Molecular Architectures That Satisfy the Chemical and Steric Requirements of a Macromolecule Binding Site." Proteins **19**: 199-221.

Falck-Pedersen, E., Heinflick, M., Alvira, M., Nussenzveig, D. R. and Gershengorn, M. C. (1994). "Expression of thyrotropin-releasing hormone receptors by adenovirus-mediated gene transfer reveals that thyrotropin-releasing hormone desensitization is cell specific." Mol. Pharm. **45**: 684-689.

Feeney, J., Bedford, G. R. and Wessels, P. L. (1974). "^{1}H Nuclear magnetic resonance studies of thyrotropin releasing factor (TRF)." FEBS Lett. **42**: 347-351.

Findlay, J. and Eliopoulos, E. (1990). "Three-dimensional modelling of G protein-linked receptors." TiPS **11**: 492-499.

Fischer, G., Bang, H. and Mech, C. (1984). "Nachweis einer Enzymkatalyse für die cis-trans-Isomerisierung der Peptidbindung in prolinhaltigen Peptiden." Biomed. biochim. Acta **43**: 1101-1111.

Fischer, S., Dunbrack, R. L. J. and Karplus, M. (1994). "Cis-Trans Imide Isomerization of the Proline Dipeptide." J. Am. Chem. Soc. **116**: 11931-11937.

Frisch, M. J., Trucks, G. W., Head-Gordon, M., Gill, P. M. W., Wong, M. W., Foresman, J. B., Johnson, B. G., Schlegel, H. B., Robb, M. A., Replogle, E. S., Gomperts, R., Andres, J. L., Raghavachari, K., Binkley, J. S., Gonzalez, C., Martin, R. L., Fox, D. J., Defrees, D. J., Baker, J., Stewart, J. J. P. and Pople, J. A. (1992). Gaussian 92. Pittsburgh, Gaussian Inc.

Galardy, R. E., Alger, J. R. and Liakopoulou-Kyriakides, M. (1982). "s-*Cis* and s-*trans* isomerism in acylproline analogs." Int. J. Peptide Protein Res. **19**: 123-132.

Garduño-Juarez, R., Morales, L. and Perez-Neri, F. (1990). "Global minimum energy conformations of thyrortopin releasing hormone analogs by simulated annealing." J. Mol. Struc. (Theochem) **208**: 279-300.

Goodford, P. J. (1985). "A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules." J. Med. Chem. **28**: 849-857.

Goodsell, D. S. and Olson, A. J. (1990). "Automated Docking of Substrates to Proteins by Simulated Annealing." Proteins **8**: 195-202.

Goren, H. J. and Bauce, L. G. (1977). "Forces and Structural Limitations of Binding of Thyrotrophin-Releasing Factor to the Thyrortopin-Releasing Receptor: the Pyroglutamic Acid Moiety." Mol. Pharmacol. **13**: 606-614.

Grant, G., Ling, N., Rivier, J. and Vale, W. (1972). "Orientation Restrictions of the Peptide Hormone, Thyrotropin-Releasing Factor, Due to Intramolecular Hydrogen Bonding." Biochemistry **11**: 3070-3073.

Grathwohl, C. and Wütrich, K. (1976). "Nmr Studies of the Molecular Conformations in the Linear Oligopeptides H-(L-Ala)$_n$-L-Pro-OH." **15**: 2043-2057.

Grathwohl, C. and Wütrich, K. (1976). "The X-Pro Peptide Bond as an Nmr Probe for Conformational Studies of Flexible Linear Peptides." Biopolymers **15**: 2025-2041.

Guarnieri, F. and Still, C. (1994). "A Rapidly Convergent Simulation Method: Mixed Monte Carlo/Stochastic Dynamics." J. Comp. Chem. **15**: 1302-1310.

Guarnieri, F. and Wilson, S. R. (1995). "Conformational Memories and a Simulated Annealing Program That Learns: Application to LTB_4." J. Comp. Chem. **16**: 648-653.

Guillemin, R., Burgus, R. and Vale, W. (1971). The Hypothalamic Hypophysiotropic Thyrotropin-Releasing Factor. Vitamins and Hormones. Ed: R. S. Harris, E. Diczfalusy, P. L. Munson and J. Glover. **29:** 1-39.

van Gunsteren, W. F., King, P. M. and Mark, A. E. (1994). "Fundamentals of drug design from a biophysical viewpoint." Quat. Rev. Biophys. **27**: 435-481.

Haar, W., Fermandjian, S., Vicar, J., Bláha, K. and Fromageot, P. (1975). "^{13}C-nuclear magnetic resonance study of [85 % ^{13}C-enriched proline]thyrotropin releasing factor: ^{13}C-^{13}C vicinal coupling constants and conformation of the proline residue." Proc. Nat. Acad. Sci. USA **72**: 4948-4952.

Hart, T. N. and Read, R. J. (1992). "A Multiple-Start Monte Carlo Docking Method." Proteins **13**: 206-222.

Head-Gordon, T. and Stillinger, F. H. (1993). "Predicting polypeptide and protein structures from amino acid sequence: Antlion method applied to melittin." Biopolymers **33**: 293-303.

Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E. and Downing, K. H. (1990). "Model for the Structure of Bacteriorhodopsin Based on High-resolution Electron Cryo-microscopy." J. Mol. Biol. **213**: 899-929.

Hibert, M., Trumpp-Kallmayer, S., Bruinvels, A. and Hoflack, J. (1991). "Three-Dimensional Models of Neurotransmitter G-Binding Protein-Coupled Receptors." Mol. Pharm. **40**: 8-15.

Higashijima, T., Tasumi, M. and Miyazawa, T. (1977). "^1H Nuclear Magnetic Resonance Studies of N-Acetyl-L-Proline N-Methylamide. Molecular Conformations, Hydrogen bondings, and Thermodynamic Quantities in Various Solvents." Biopolymers **16**: 1259-1270.

Hinkle, P. M. and Kinsella, P. A. (1982). "Rapid Temperature-dependent Transformation of the Thyrotropin-releasing Hormone-Receptor Complex in Rat Pituitary Tumor Cells." J. Biol. Chem. **257**: 5462-5470.

Hodes, Z. I., Némethy, G. and Scheraga, H. A. (1979). "Influence of Hydration on the Conformational Stability and Formation of Bends in Terminally Blocked Dipeptides." Biopolymers **18**: 1611-1634.

Hsieh, K.-P. and Martin, T. F. J. (1992). "Thyrotropin-releasing hormone and gonadotropin-releasing hormone receptors activate phospholipase C by coupling to the guanosine triphosphate-binding proteins G_q and G_{11}." Mol. Endocrinol. **6**: 1673-1681.

Hu, C. H., Shen, M. and Schaefer, H. F. I. (1993). "Glycine Conformational Analysis." J. Am. Chem. Soc. **115**: 2923-2929.

Ijzerman, A. P., van Galen, P. J. M. and Jacobson, K. A. (1992). "Molecular modeling of adenosine receptors." Drug Des. Discov. **9**: 49.

Jackson, I. M. D. (1982). "Thyrotropin-releasing hormone." New Engl. J. Med. **306**: 146-155.

Jorgensen, W. L. and Gao, J. (1988). "Cis-Trans Energy Difference for the Peptide Bond in the Gas Phase and in Aqueous Solution." J. Am. Chem. Soc. **110**: 4212-4216.

Kamiya, K., Takamoto, M., Wada, Y., Fujino, M. and Nishikawa, M. (1980). "Molecular Conformation of Thyrotropin-releasing Hormone from the X-ray Structural Analysis of its Tartrate." J.C.S.Chem. Comm.: 438-439.

Khan, Z., Aitken, A., del Rio Garcia, J. and Smyth, D. G. (1992). "Isolation and Identification of Two Neutral Thyrotropin Releasing Hormone-like Peptides, Pyroglutamylphenylalanineproline Amide and Pyroglutamylglutamineproline Amide, from Human Seminal Fluid." J. Biol. Chem. **267**: 7464-7469.

Kirkpatrick, S., Gelatt, C. D. J. and Vecchi, M. P. (1983). "Optimization by Simulated Annealing." Science **220**: 671-680.

Kontoyianni, M. and Lybrand, T. P. (1993). "Three-dimensional modles for integral membrane proteins: Possibilities and pitfalls." Persp. Drug Disc. Design **1**: 291-300.

Konvicka, K., Ballesteros, J., A. and Weinstein, H. (1995). A Structural and Functional Role for the Conserved NP/DP Motif in Helix 7 of GPCR Inferred from a Model of the GnRH Receptor. Biophysical Society Meeting, San Fransisco.

Kostrowicki, J. and Scheraga, H. A. (1992). "Application of the Diffusion Equation Method for Global Optimization to Oligopeptides." J. Phys. Chem. **96**: 7442-73449.

Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langeridge, R. and Ferrin, T. E. (1982). "A Geometric Approach to a Macromolecule-Ligand Interactions." J. Mol. Biol. **161**: 269-288.

Kyte, J. and Doolittle, R. F. (1982). "A SImple Method for Displaying the Hydropathic Character of a Protein." J. Mol. Biol. **157**: 105-132.

Leach, A. R. (1994). "Ligand Docking to Proteins with Discrete Side-chain Flexibility." J. Mol. Biol. **235**: 345-356.

Lee, C. (1994). "Predicting Protein Mutant Energetics by Self-consistent Ensemble Optimizations." J. Mol. Biol. **236**: 918-939.

Liakopoulou-Kyriakides, M. and Galardy, R. E. (1979). "s-Cis, and s-Trans Isomerism of the His-Pro Peptide Bond in Angiotensin and Thyroliberin analogues." Biochemistry **18**: 1952-1957.

Lipinsky, D. and Gershengorn, M. C. (1992). "$G\alpha_{11}$ and $G\alpha_q$ guanine nucleotide regulatory proteins differentially modulate the response to thyrotropin-releasing hprmone in *Xenopus* oocutes." FEBS Lett. **307**: 237-240.

MacArthur, M. W. and Thornton, J. M. (1991). "Influence of Proline Residues on Protein Conformation." J. Mol. Bio. **218**: 397-412.

Madison, V. and Kopple, K. D. (1980). "Solvent-Dependent Conformational Distributions of Some Dipeptides." J. Am. Chem. Soc. **102**: 4855-4863.

MaloneyHuss, K. and Lybrand, T. P. (1992). "A three-dimensional structure for the β_2 adrenergic receptor protein based on computer modeling studies." J. Mol. Biol. **225**: 859-871.

Mapelli, C., Elrod, L. F., Switzer, F. L. and Stammer, C. H. (1989). "Conformational Properties of 2,3-Methanopyroglutamic Acid in Peptides: NMR and X-Ray Diffraction Studies." Biopolymers **28**: 123-128.

Mapelli, C., van Hallbeck, H. and Stammer, C. H. (1990). "Synthesis and Conformational Studies by ^{1}H- and ^{13}C-NMR Spectroscopy of a Novel, Sterically Constrained Analogue of Thyrotropin-Releasing Hormone." Biopolymers **29**: 407-422.

Marshall, G. R. (1982). Conformational Studies of Peptide hormones as a Basis for Analog Design. Chemical Regulation of Biological Mechanisms Proceedings of the 1st Medicinal Chemistry Symposium, Royal Society of Chemistry, London

Matsuzaki, T. and Iitaka, Y. (1971). "The Crystal Structure of Acetyl-L-proline-N-methylamide." Acta Cryst. **B27**: 507-516.

Matus-Leibovitch, N., Nussenzveig, D. R., Gershengorn, M. C. and Oron, Y. (1995). "Truncation of thyrotropin-releasing hormone carboxyl tail causes constitutive activity and leads to impaired responsiveness in *Xenopus* oocutes and AtT20 cells." J. Biol. Chem. **270**: 1041-1047.

McDonald, Q. D. and Still, W. C. (1992). "AMBER* Torsional Parameters for the Peptide Backbone." Tetrahedron Lett. **33**: 7743-7746.

McGregor, M. J., Islam, S. A. and Sternberg, M. J. E. (1987). "Analysis of the Relationship Between Side-chain Conformation and Secondary Structure in Grobular Proteins." J. Mol. Biol. **198**: 295-310.

Metcalf, G. and Jackson, I. M. D., Eds. (1989). Thyrotropin-Releasing Hormone. Biomedical Significance. Annals of the New York Academy of Sciences. New York, The New York Academy of Sciences.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953). "Equation of State Calculations by Fast Computing Machines." J. Chem. Phys. **21**: 1087-1092.

Miranker, A. and Karplus, M. (1991). "Functionality Maps of Binding sites: A Multiple Copy Simultaneous Search Method." Proteins **11**: 29-34.

Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T. and Still, W. C. (1990). "MacroModel - An Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics." J. Comp. Chem. **11**: 440.

Montagut, M., Lemanceau, B. and Bellocq, A.-M. (1974). "Conformational Analysis of Thyrotropin Releasing Factor by Proton Magnetic Resonance Spectroscopy." Biopolymers **13**: 2615-2629.

Morgan, B. A., Bower, J. D., Dettmer, P. W., Metcalf, G. and Schafer, D.J., (1979). "Novel TRH analogs with increased neuropharmacological activity." Proceedings of the Sixth American Peptide Symposium, Peptides, Pierce Chemical Company, Rockford, Illinois.

Nicholls, A., Sharp, K. A. and Honig, B. (1991). "Protein Folding and Association: Insights From the Interfacial and Thermodynamic Properties of Hydrocarbons." Proteins **11**: 281-296.

Nordvall, G. and Hacksell, U. (1993). "Binding-Site Modeling of the Muscarinic m1 Receptor: A Combination of Homology-Based and Indirect Approaches." J. Med. Chem. **36**: 967-976.

Nussenzveig, D. R., Heinflick, M. and Gershengorn, M. C. (1993). "Agonist-stimulated internalization of the thyrotropin-releasing hormone receptor is dependent on two domains in the receptor carboxyl terminus." J. Biol. Chem. **268**: 2389-2392.

Nutt, R. F., Hirschmann, R. and Veber, D. F. (1979). "Synthesis of <Aad-His-(4R, 5R)-5-methyl-Tzl-NH_2, a TRH analog with high CNS activity." Proceedings of the Sixth American Peptide Symposium, Pierce Chemical Company, Rockford, Illinois.

O'Dowd, B. F., Hnatowich, M., Caron, M. G., Lefkowitz, R. J. and Bouvier, M. (1989). "Palmitoylation of the Human β_2-Adrenergic Receptor." J. Biol. Chem. **264**: 7564-7569.

Olson, G. L., Bolin, D. R., Bonner, M. P., Bös, M., Cook, C. M., Fry, D. C., Graves, B. J., Hatada, M., Hill, D. E., Kahn, M., Madison, V. S., Rusiecki, V. K., Sabaru, R., Sepinwall, J., Vincent, G. P. and Voss, M. E. (1993). "Concepts and Progress in the Development of Peptide Mimetics." J. Med. Chem. **36**: 3039-3049.

Olson, G. L., Cheung, H.-C., Chiang, E., Madison, V. S., Sepinwall, J., Vincent, G. P., Winokur, A., and Gary, K. A. (1995). "Peptide Mimetics of Thyrotropin-Releasing Hormone Based on a Cyclohexane Framework: Design, synthesis, and Cognition-Enhancin g Properties." J. Med. Chem. **38**: 2866-2879.

Paulssen, E. J., Paulssen, R. H., Gautvik, K. M. and Gordeladze, J. O. (1992). "The thyroliberin receptor interacts directly with a stimulatory guanine nucleotide-binding proteins in the activation of adenylyl cyclase in GH3 rat pituitary tumour cells." Eur. J. Biochem. **204**: 413-418.

de la Peña, P., del Camino, D., Pardo, L. A., Dominguez, P. and Barros, F. (1995). "G_s couples thyrotropin-releasing hormone receptors expressed in *Xenopus* oocytes to phospholipase C." J. Biol. Chem. **270**: 3554-3559.

Perlman, J. H., Laakkonen, L., Osman, R. and Gershengorn, M. C. (1994b). "A Model of Thyrotropin-releasing Hormone (TRH) Receptor Binding Pocket." J. Biol. Chem. **269**: 23383-23386.

Perlman, J. H., Laakkonen, L., Osman, R. and Gershengorn, M. C. (1995). "Distinct roles for arginines in transmembrane helices 6 and 7 of the thyrotropin-releasing hormone receptor." Mol. Pharm.

Perlman, J. H., Laakkonen, L., Thaw, C. N., Osman, R. and Gershengorn, M. C. (1993). "Analysis of the role of transmembrane helix three of the thyrotropin-releasing hormone (TRH) receptor in directly binding TRH." Trans. Assoc. Am. Phys. **CVI**: 162-167.

Perlman, J. H., Nussenzveig, D. R., Osman, R. and Gershengorn, M. C. (1992). "Thyrotropin-releasing hormone binding to the mouse pituitary receptors does not involve ionic interactions. A model for neutral peptide binding to G protein-coupled receptors." J. Biol. Chem. **267**: 24413-24417.

Perlman, J. H., Thaw, C. N., Laakkonen, L., Bowers, C. Y., Osman, R. and Gershengorn, M. C. (1994a). "Hydrogen Bonding Interaction of the Thyrotropin-releasing Hormone (TRH) with Transmembrane Tyrosine 106 of the TRHR Receptor." J. Biol. Chem. **269**: 1610-1613.

Petitt, B. M., Matsunaga, T., Al-Obeidi, F., Gehrig, C., Hruby, V. J. and Karplus, M. (1991). "Dynamical search for bis-penicillamine enkephalin conformations." Biophys. J. **60**: 1540-1544.

Piela, L., Kostrowicki, J. and Scheraga, H. A. (1989). "The Multiple-Minima Problem in the Conformational Analysis of Molecules. Deformation of the Potential Energy Hypersurface by the Diffusion Equation Method." J. Phys. Chem. **93**: 3339-3346.

Prasad, C. (1984). Thyrotropin-Releasing Hormone. Handbook of Neurochemistry. Ed: A. Lajtha. New York, Plenum. **8:** 175-195.

Presta, L. G. and Rose, G. D. (1988). "Helix Signals in Proteins." Science **240**: 1632-1641.

Probst, W. C., Snyder, L. A., Schuster, D. I., Brosius, J. and Sealfon, S. C. (1992). "Sequence Alignment of the G-Protein Coupled Receptor Superfamily." DNA Cell Biol. **11:1**: 1-20.

Quick, M. W., Simon, M. I., Davidson, N., Lester, H. A. and Aragay, A. M. (1994). "Differential coupling of G protein α subunits to seven-helix receptors expressed in *Xenopus* oocutes." J. Biol. Chem. **269**: 30164-30172.

Radzicka, A., Pedersen, L. and Wolfenden, R. (1988). "Influences of Solvent Water on Protein Folding: Free Energies of Solvation of Cis and Trans Peptides Are Nearly Identical." Biochemistry **27**: 4538-4541.

Rao, V. R., Cohen, G. B. and Oprian, D. D. (1994). "Rhodopsin mutation G90D and a molecular mechanism for congenital night blindness." Nature **367**: 639-641.

Reynolds, W. F., Peat, I. R., M.H., F. and Lyerla, J. R. J. (1973). "Determination of the Tautomeric Form of the Imidazole Ring of L-Histidine in Basic Solution by Carbon-13 Magnetic Resonance Spectroscopy." J. Am. Chem. Soc. **95**: 328-331.

Robson, B. and Osguthorpe, D. J. (1979). "Refined Models for Computer Simulation of Protein Folding. Applications to the Study of Conserved Secondary Structure and Flexible Hinge Points During the Folding of Pancreatid Trypsin Inhibitor." J. Mol. Biol. **132**: 19-51.

Roitberg, A. and Elber, R. (1991). "Modeling side chains in peptides and proteins: Application of the locally enhanced sampling and the simulated annealing methods to find minimum energy conformations." J. Chem. Phys. **95**: 9277-9287.

Rydel, T. J., Tulinsky, R., Bode, W. and Huber, R. (1991). "Refined Structure of the Hirudin-Thrombin Complex." J. Mol. Biol. **221**: 583-601.

Samama, P., Cotecchia, S., Costa, T. and Lefkowitz, R. J., (1993). "A Mutation-induced Activated State of the β_2-Adrenergic Receptor." J. Biol. Chem. **268**: 4625-4636.

Sankararamakrishnan, R. and Vishveshwara, S. (1990). "Conformational Studies on Peptides with Proline in the Right-Handed α-Helical Region." Biopolymers **30**: 287-298.

Saunders, M., Houk, K. N., Wu, Y.-D., Still, W. C., Lipton, M., Chang, G. and Guida, W. C. (1990). "Conformation of Cycloheptadecene. A comparison of Methods for Conformational Searching." J. Am. Chem. Soc. **112**: 1419-1427.

van Schaik, R. C., Berendsen, H. J. C., Torda, A. E. and van Gunsteren W. F. (1993). J. Mol. Biol. **234**: 751.

Schambye, H. T., Hjorth, S. A., Weinstock, J. and Schwartz, T. W. (1995). "Interaction between the Nonpeptide Angiotensin Antagonist SKF-108,566 and Histidine 256 (HisVI:16) of the Angiotensin Type 1 Receptor." Mol. Pharm. **47**: 425-431.

Schertler, G. F. X., Villa, C. and Henderson, R. (1993). "Projection structure of rhodopsin." Nature **362**: 770-772.

Schmidt, J. M., Brüschweiler, R., Ernst, R. R., Dunbrack, R. L. J., Joseph, D. and Karplus, M. (1993). "Molecular Dynamics Simulation of the Proline Conformational Equilibrieum and Dynamics in Antamanide Using the CHARMM Force Field." J. Am. Chem. Soc. **115**: 8747-8756.

Shenkin, P. S. and McDonald, D. Q. (1994). "Cluster Analysis of Molecular Conformations." J. Comp. Chem. **15**: 899.

Sievertsson, H., Castensson, S. and Bowers, C. Y. (1974). "On the conformation of thyrotropin releasing hormone." FEBS Lett. **42**: 340-342.

Sklenar, H., Etchebest, C. and Lavery, R. (1989). "Describing Protein Structure: A General Algorithm Yielding Complete Helicoidal Parameters and a Unique Overall Axis." Proteins **6**: 46-60.

Smolyar, A. and Osman, R. (1993). "Role of Threonine 342 in Helix 7 of the 5-Hydroxytryptamine Type 1D Receptor in Ligand Binding: An Indirect Mechanism for Receptor Selectivity." Mol. Pharm. **44**: 882-885.

Stamnes, M. A. and Zuker, C. S. (1990). "Peptidyl-prolyl *cis-trans* isomerases, cyclophilin, FK506-binding protein, and ninaA: four of a kind." Curr. Opin. Cell Biol. **2**: 1104-1107.

Still, W. C., Tempczyk, A., C., H. R. and Hendrickson, T. (1990). "Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics." J. Am. Chem. Soc. **112**: 6127-6129.

Still, W. C., Tempczyk, A., Hawley, R. C. and Hendrickson, T. (1990). "Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics." J. Am. Chem. Soc. **112**: 6127-6129.

Strader, C. D., Sigal, I. S., Register, R. B., Candelore, M. R., Rands, E. and Dixon, R. A. F. (1987). "Identification of residues required for ligand binding to the β-adrenergicreceptor." Proc. Natl. Acad. Sci. USA **84**: 4384-4388.

Straub, R. E., Frech, G. C., Joho, R. H. and Gershengorn, M. C. (1991). "Expression cloning of cDNA encoding the mouse pituitary thyrotropin-releasing hormone receptor." Proc. Natl. Acad. Sci. USA **87**.

Suryanarayana, S. and Kobilka, B. K. (1993). "Amino acid substitutions at position 312 in the seventh hydrophobic domain of the β_2-adrenergic receptor modify ligand binding specificity." Mol. Pharm. **44**: 111-114.

Swaminathan, S., Ravishanker, G., Beveridge, D. L., Lavery, R., Etchebest, C. and Sklenar, H. (1990). "Conformational and Helicoidal Analysis of the Molecular Dynamics of Proteins: "Curves". Dials and Windows for a 50 psec Dynamic Trajectory of BPTI." Proteins **8**: 179-193.

Swope, W. C., Andersen, H. C., Berens, P. H., Wilson, K. R. (1982). "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters." J. Chem. Phys. **76**: 637-649.

Szirtes, T., Kisfaludy, L., Palosi, E. and Szporny, L. (1984). "Synthesis of Thyrotropin-Releasing Hormone Analogues. 1. Complete Dissociation of Central Ner-

vous System Effects from Thyrotropin-Releasing Activity." J. Med. Chem. **27**: 741-745.

Teeter, M. M., Froimowitz, M., Stec, B. and DuRand, C. J. (1994). "Homology Modeling of the Dopamine D_2 Receptor and ItsTesting by docking of Agonists and Tricyclic Antagonists." J. Med. Chem. **37**: 2874-2888.

Thomas, W. A. and Kevin Williams, M. (1972). "^{13}C nuclear Magnetic Resonance Spectroscopy and cis/trans Isomerism in Dipeptides containing Proline." J.C.S. Chem. Comm.: 994.

Torjesen, P. A., Bjøro, T., Østberg, B. C. and Haug, E. (1988). "Thyrotropin-releasing hormone-stimulated inositol trisphosphate formation is liable to thyrotropin-releasing hormone-indiced desensitization by a calcium-mediated mechanism." Mol. Cell. Endocrinol. **56**: 107-114.

Treutlein, H. R., Lemmon, M. A., Engelman, D. M. and Brünger, A. T. (1992). "The Glycophorin A Transmembrane Domain Dimer: Sequence-Specific Propensity for a Right-Handed Supercoil of Helices." Biochemistry **31**: 12726-12733.

Trumpp-Kallmeyer, S., Bruinvels, A., Hoflack, J. and Hibert, M. (1991). "Recognition site mapping and receptor modelling: Application to 5-HT receptor." Neurochem. Int. **19**: 397-406.

Tsuboi, M., Shimanouchi, T. and Mizushima, S.-i. (1959). "Near Infrared Spectra of Compounds with Two Peptide Bonds and the Configuration of a Polypeptide Chain. VII. On the Extended Forms of Polypeptide Chains." J. Am. Chem. Soc. **81**: 1406-1411.

Tuffery, P., Etchebest, C., Popot, J.-L. and Lavery, R. (1994). "Prediction of the Positioning of the Seven Transmembrane α-helices of Bacteriorhodopsin. A Molecular Simulation Study." J. Mol. Biol. **236**: 1105-1122.

Unger, V. M. and Schertler, G. F. X. (1995). "Low-Resolution Structure of Bovine Rhodopsin Determined by Electron Cryo-Microscopy." Biophys. J. **68**: 1776-1786.

Unkefer, C. J., Walker, R. D. and London, R. E. (1983). "Hydrogen-1 and carbon-13 nuclear magnetis resonance conformational studies of the His-Pro peptide

bond: conformational behaviour of TRH." Int. J. Peptide Protein Res. **22**: 582-589.

Vale, W., Grant, G. and Guillemin, R. (1973). Chemistry of the Hypothalamic Releasing Factors - Studies on Structure-Function Relationship. Frontiers in Neuroendocrinology. Ed: W.F. Ganong and L. Martini. **11:** 375-402.

Vásques, M., Némethy, G. and Scheraga, H. A. (1994). "Conformational Energy Calculations on Polypeptides and Proteins." Chem. Rev. **94**: 2183-2239.

Vicar, J., Abillon, E., Toma, F., Pirou, F., Litner, K., Bláha, K., Fromageot, P. and Fermandjian, S. (1979). "The two conformations of TRH in solution." FEBS Lett. **97**: 275-278.

Ward, D. J., Finn, P. W., Griffiths, E. C. and Robson, B. (1987). "Comparative conformation-activity relationships for hormonally- and centrally-acting TRH." Int. J. Peptide Protein Res. **30**: 263-274.

Weiner, P. K. and Kollman, P. A. (1981). "AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions." J. Comp. Chem. **2**: 287-303.

Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G. A., Profeta, S. J. and Weiner, P. (1984). "A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins." J. Am. Chem. Soc. **106**: 765-784.

Wess, J., Maggio, R., Palmer, J. R. and Vogel, Z. (1992). "Role of conserved threonine and tyrosine residues in acetylcholine binding and muscarinic receptor activation: a study with m3 muscarinic receptor point mutaions." J. Biol. Chem. **267**: 19313-19319.

Wong, M. W., Frisch, M. J. and Wiberg, K. B. (1991a). "Solvent Effects. 1. The Mediation of Electrostatic Effects by Solvents." J. Am. Chem. Soc. **113**: 4776-4782.

Wong, M. W. and Wiberg, K. B. (1991b). "Hartree-Fock second derivatives and electric field properties in a solvent reaction field: Theory and applications." J. Chem. Phys. **95**: 8991-8998.

Wong, M. W., Wiberg, K. B. and Frisch, M. J. (1992a). "Solvent Effects. 2. Medium Effect of the Structure, Energy, Charge Density, and Vibrational Frequencies of Sulfamic Acid." J. Am. Chem. Soc. **114**: 523-529.

Wong, M. W., Wiberg, K. B. and Frisch, M. J. (1992b). "Solvent Effects. 3. Tautomeric Equilibria of Formamide and 2-Pyridone in the Gas Phase and Solution. An ab Initio SCRF Study." J. Am. Chem. Soc. **114**: 1645-1652.

Wütrich, K., Tun-Kyi, A. and Schwyzer, R. (1972). "Manifestation in the ^{13}C-NMR spectra of two different molecular conformations of a cyclic pentapeptide." FEBS Lett. **25**: 104-108.

Wütrich, K., von Freyberg, B., Weber, C., Wider, G., Traber, R., Widmer, H. and Braun, W. (1991). "Receptor-Induced Conformation Change of the Immunosuppressant Cyclosporin A." Science **254**: 953-954.

Yang, J. and Tashjian, A. H. J. (1993). "Regulation of endogenous thyrotropin-releasing hormone (TRH) receptor messanger RNA by TRH in GH_4C_1 cells." Mol. Endocrinol. **7**: 753-758.

Yaron, A. and Naider, F. (1993). "Proline-Dependant Structural and Biological Properties of Peptides and Proteins." Crit. Rew. Biochem. Mol. Bio. **28**: 31-81.
Yue, S.-Y. (1990). "Distance-constrained molecular docking by simulated annealing." Prot. Eng. **4**: 177-184.

Zhang, D. and Weinstein, H. (1993a). "Ligand selectivity and the molecular properties of the 5-HT2 receptor: computational simulations reveal a major role for the transmembrane helix 7." Med. Chem. Res. **3**: 357-369.

Zhang, D. and Weinstein, H. (1993b). "Signal Transduction by the 5-HT2 Receptor: A Mechanistic Hypothesis from Molecular Dynamics Simulations of the Three-Dimensional Model of the Receptor Complexed to Ligands." J. Med. Chem. **36**: 934-938.

Zhou, W., Flanagan, C., Ballesteros, J. A., Konvicka, K., Davidson, J. S., Weinstein, H., Millar, R. P. and Sealfon, S., C. (1994). "A Reciprocal Mutation Reveals Helix 2 and Helix 7 Proximity in the Gonadotropin-Releasing Hormone Receptor." Mol. Pharm. **45**: 165-170.

Zimmerman, S. S. and Scheraga, H. A. (1977). "Influence of Local Interactions on Protein Structure. I. Conformational Energy Studies of N-Acetyl-N'-Methyl-amides of Pro-X and X-Pro Dipeptides." Biopolymers **16**: 811-843.