Static and Dynamic Roles of Extracellular Loops in G-Protein-Coupled Receptors: A Mechanism for Sequential Binding of Thyrotropin-Releasing Hormone to Its Receptor

Anny-Odile Colson,* Jeffrey H. Perlman,# Alex Smolyar,* Marvin C. Gershengorn,# and Roman Osman*
*Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York, New York 10029, and #Division of Molecular Medicine, Department of Medicine, Cornell University Medical College and the New York Hospital, New York, New York 10021 USA

ABSTRACT Small ligands generally bind within the seven transmembrane-spanning helices of G-protein-coupled receptors, but their access to the binding pocket through the closely packed loops has not been elucidated. In this work, a model of the extracellular loops of the thyrotropin-releasing hormone (TRH) receptor (TRHR) was constructed, and molecular dynamics simulations and quasi-harmonic analysis have been performed to study the static and dynamic roles of the extracellular domain. The static analysis based on curvature and electrostatic potential on the surface of TRHR suggests the formation of an initial recognition site between TRH and the surface of its receptor. These results are supported by experimental evidence. A quasi-harmonic analysis of the vibrations of the extracellular loops suggest that the low-frequency motions of the loops will aid the ligand to access its transmembrane binding pocket. We suggest that all small ligands may bind sequentially to the transmembrane pocket by first interacting with the surface binding site and then may be guided into the transmembrane binding pocket by fluctuations in the extracellular loops.

INTRODUCTION

The thyrotropin releasing hormone receptor (TRHR) is a member of a large family of transmembrane proteins (GPCR) for which an interaction with an intracellular G protein is a critical part of the signal transduction pathway mediated by the receptor. It is thought that all GPCRs have a common tertiary structure composed of seven transmembrane helices. The topology of these membrane-bound proteins is defined by an extracellular amino terminus and an intracellular carboxy terminus. Consequently, the helices are connected by three intracellular and three extracellular loops. As is the case for most integral membrane proteins, and especially for GPCRs, the three-dimensional structure of TRHR is not known. To gain understanding about the relationship between structure and function of TRHR, we developed a combined theoretical/experimental approach. The structural models of the receptor are constructed and studied by theoretical simulation methods (Laakkonen et al., 1996), and the structural/functional inferences are tested by experimental molecular biological approaches (Gershengorn and Osman, 1996; Perlman et al., 1994a,b, 1995a, 1996).

In a previous work (Laakkonen et al., 1996), we presented a de novo model for the transmembrane domain of the receptor that was constructed based on a generic template for GPCRs developed from an analysis of homologous GPCR sequences (Baldwin, 1993). The template consisted of helical axes, which provided guidelines for positioning, tilting, and orienting the helices in a transmembrane bundle. Guided by mutagenesis and simulation studies (Laakkonen et al., 1996; Perlman et al., 1994a,b, 1995a, 1996), we suggested that the binding pocket of TRH (p-Glu-His-ProNH_2) consists at least of four residues that lie within the upper half of the transmembrane core. The four residues that were previously identified as part of the TRH binding pocket are Tyr106 and Asn110 in helix 3, Tyr282 in helix 6, and Arg306 in helix 7, and they form specific interactions with TRH. Thus, Tyr106 forms a H-bond to the C═O of p-Glu, and the dipole of Asn110 is antiparallel to the CO-NH dipole of p-Glu, giving rise to a stabilizing dipole-dipole interaction. Tyr282 in helix 6 forms a stacking or hydrophobic interaction with the imidazole of the His residue of TRH, and Arg306 in helix 7 forms an ionic H-bond with the C═O of the prolineamide group and the backbone of TRH. Novel mixed-mode Monte Carlo/stochastic dynamics simulations were used to test the validity of the complex between the hormone and the receptor and to investigate the nature of the hormone-receptor interactions (Laakkonen et al., 1996). An extracellular view of the space-filling model of the occupied receptor described in the previous work (Laakkonen et al., 1996) shows that the ligand is fully buried. This raises the question of the mechanism of TRH access to the transmembrane binding pocket, and especially the role of the extracellular loops in occluding or helping the access of the ligand into the binding pocket.

Recent experimental work supports the hypothesis that the extracellular domain of the TRHR plays a role in ligand access to the binding pocket. For example, mutations of various residues in the extracellular loops reduced apparent binding of TRH by a significant amount (Han and Tashjian, 1995a,b). In particular, mutation of N289 to Ala or Asp in

Received for publication 21 April 1997 and in final form 17 November 1997.

Address reprint requests to Dr. Roman Osman, Mount Sinai School of Medicine, Department of Physiology and Biophysics, Box 1218, 1 Gustave L. Levy Place, New York, NY 10029. Tel.: 212-241-5609; Fax: 212-860-3369; E-mail: osman@inka.mssm.edu.

0006-3495/98/03/1087/14 $2.00

extracellular loop 3 reduced the affinity to TRH by approximately 10- or 100-fold, respectively. The fact that affinity of the N289D mutant to the Pro[1]TRH derivative increased (>10-fold) whereas that of the N289A mutant did not change compared with the wild-type receptor led the authors to conclude that the interaction between p-Glu and N289 is specific (Han and Tashjian, 1995b). Our observations that the binding pocket is entirely positioned in the transmembrane domain (see above) and the results presented by Han and Tashjian suggest that TRH initially interacts with residues in the extracellular loops (e.g., Asn289) and subsequently moves into the transmembrane binding pocket. Recent experiments to determine the rate constants for hormone binding (Perlman et al., 1998) supported by simulations (Colson et al., 1997) are in agreement with this suggestion. Such a sequential binding model rationalizes the experimental findings and highlights the importance of the extracellular domain in molecular models of the TRH receptor, and possibly in GPCRs in general.

In an effort to elucidate the mechanisms by which TRH accesses its transmembrane binding pocket and the potential involvement of the extracellular domains in such mechanisms, we have constructed a model that encompasses the seven transmembrane helices and the three extracellular loops of TRHR. Despite the rapidly increasing number of molecular models of GPCRs and a systematic approach to the construction of the transmembrane domain (Ballesteros and Weinstein, 1995), only few models have included the intra- and extracellular loops (Dahl et al., 1991; Findlay and Eliopoulos, 1990; Kyle et al., 1994; MaloneyHuss and Lybrand, 1992; Sylte et al., 1996). None of these works addressed critically the difficulties involved in the construction of the loops in the absence of structural guidelines. Consequently, the constructed models of the loops were not described in detail, nor was their relevance to the binding process critically evaluated.

We present results that demonstrate the validity of the construction model. Results from annealings and long (1-ns) molecular dynamics (MD) simulations show the possible formation of a surface binding pocket, the properties of which are complementary to TRH. Finally, a quasi-harmonic analysis of trajectories from long simulations suggests the possible importance of anti-correlated conformational motions of the TRHR extracellular loops in hormone binding.

MATERIALS AND METHODS

Construction of the extracellular loops

All calculations were performed using the CHARMM program version 23 (Brooks et al., 1983b). The calculations were carried out with the all-atom parameter set par_all22_prot. The SHAKE algorithm was employed to fix all bonds to hydrogen atoms, and the environment was represented by a distance-dependent dielectric function.

The model receptor studied in this work consists of the seven transmembrane helices for which the structure was obtained from previous work performed in our laboratory (Laakkonen et al., 1996) and the three extracellular loops of which the amino acid sequence and two-dimensional topology was described previously (Straub et al., 1990). No consideration was made of the amino-terminal domain, because previous experimental results have shown that most of this domain is not essential for agonist binding (Han and Tashjian, 1995a). As the ultimate goal of this work is to study the conformation and mobility of the extracellular loops in relation to the ligand's access to the transmembrane binding pocket, the seven helix bundle remained frozen in the geometry obtained previously (Laakkonen et al., 1996). Simulations of the relaxed system are ongoing in our laboratory.

The three extracellular loops shown in Fig. 1 were constructed as four fragments: D85-L99 (EC1), F161-C179 (EC2A), G180-S189 (EC2B), and N289-E298 (EC3), so that the disulfide bond between the conserved cysteines of loops 1 and 2, i.e., C98 and C179, was maintained throughout all simulations. This bridge has been shown to be necessary in maintaining TRHR in a high-affinity conformation (Perlman et al., 1995b). The loops were then attached at one end to their respective target helix via a *trans* peptide bond. The backbone ϕ and ψ torsion angles of the loops were manually rotated so that the free end of each loop came into proximity with the other respective target helix. Adopted Basis Newton Raphson (ABNR) minimization was performed on the initial structure to establish a proper peptide bond between the carboxy end of the loop and the amino terminus of the helix. Dihedral constraints were applied to the ω angles of the loop-helix junctions to maintain the peptide bond in a *trans* orientation. Once appropriate peptide bond lengths were obtained at these junctions, the unconstrained loops were minimized for 1000 steps.

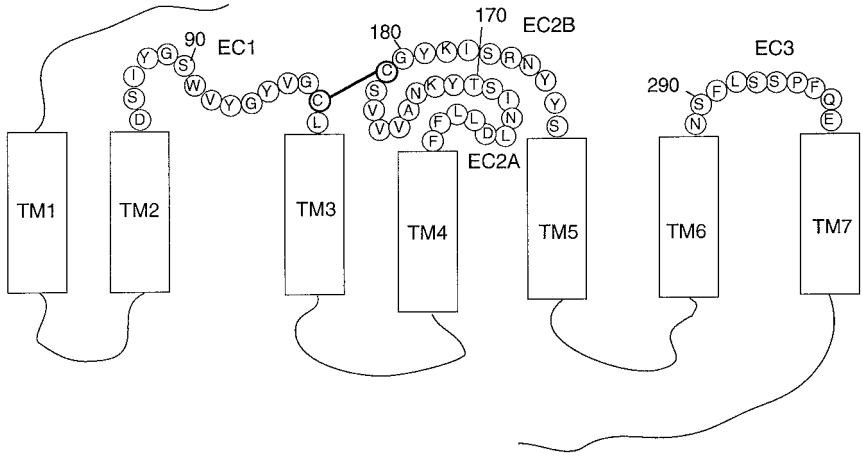

FIGURE 1 Schematic description of the extracellular domains of TRHR.

Simulated annealings and molecular dynamics simulations

The minimized structure was heated to 1500 K in 14 ps followed by 7 ps of constant temperature MD simulation at 1500 K. Fourteen structures were extracted from the trajectory at 1500 K by sampling every 0.5 ps. Each structure was cooled down to 300 K in 60 ps and subsequently subjected to 100 ps of constant-temperature MD simulation at 300 K.

Fourteen energy-minimized averaged structures were obtained over the stabilized portion of each trajectory and clustered according to their pairwise root mean square deviations into conformational families employing the program Xcluster (Shenkin and McDonald, 1994). One of the structures from the major family obtained in the clustering procedure was employed as the starting structure for a 1-ns MD simulation. The 1-ns simulation was performed at 300 K applying the SHAKE algorithm to all hydrogen atoms. The step size was 1.0 fs, and the coordinates were recorded every 0.1 ps. The nonbonded lists were generated using an atom-based cutoff of 13.0 Å and updated every 25 steps.

Quasi-harmonic analysis

Normal mode analysis of the root mean square positional fluctuations of atoms in proteins have been found to correlate well with experimentally observed properties (Brooks and Karplus, 1983a, Go et al., 1983). The inherent limitations of normal mode analysis of a nonharmonic force field can be overcome by the use of quasi-harmonic models (Teeter and Case, 1990) in which effective vibrational modes are computed such that the second moments of the amplitude distribution matches those found in a MD simulation using the complete anharmonic force field. The resultant quasi-normal modes are a quadratic approximation to the potential of mean force instead of the potential energy. In the present work, the quasi-harmonic analysis was performed with CHARMM on a 1-ns trajectory of TRHR in which the transmembrane helices were held at fixed positions. In an effort to reduce the size of the matrices to be calculated in the quasi-harmonic analysis, the energy-minimized average structure of this simulation was represented by the coordinates of the center of mass of each residue. This structure served as the reference to orient the reduced trajectory of the coordinates of the centers of mass. The quasi-harmonic analysis was subsequently performed and resulted in 162 effective vibrational modes that represent the fluctuations of 54 residues of the extracellular loops.

Excursion distances for each center of mass were determined from the trajectory projected on the first four quasi-normal modes using Quanta (Molecular Simulations, University of York, York, UK). They represent the maximal distances traveled by each center of mass within the same effective vibrational mode.

Materials

[^{3}H]Methyl (Me)TRH was obtained from DuPont (Wilmington, DE). *myo*-[^{3}H]Inositol was obtained from Amersham Corp. (Arlington Heights, IL). TRH was from Calbiochem (La Jolla, CA) and MeTRH from Sigma Chemical Co. (St. Louis, MO). Restriction endonucleases were from New England Biolabs (Beverly, MA). The cloning vector pBluescript was from Stratagene (La Jolla, CA), and the expression vector pCDM8 from Invitrogen (San Diego, CA). Dulbecco's modified Eagle's medium and fetal calf serum were from Collaborative Research (Bedford, MA).

Cell culture and transfection

COS-1 cells were maintained and transfected as described previously (Straub et al., 1990). In brief, cells were seeded 1 or 2 days before transfection at 0.7×10^6 to 1.5×10^6 cells/100-mm dish. Cells were transfected using the DEAE-dextran method as described and maintained in Dulbecco's modified Eagle's medium with 10% fetal calf serum for 1 day at which time cells were harvested and seeded into 12-well plates at 10^5 cells/well in Dulbecco's modified Eagle's medium with 5% fetal calf serum.

Mutagenesis

The full-length mouse TRHR cDNA in pBluescript (pBSmTRHR) (Straub et al., 1990) or pCDM8 (pCDM8mTRHR) (Gershengorn and Thaw, 1991) was used for mutation. The polymerase chain reaction was used to generate fragments containing the R185A mutation, which was subcloned into pBSmTRHR. A fragment derived from digestion with *Xho*I and *Not*I was then subcloned into pCDM8mTRHR. The mutation was confirmed by the dideoxy chain termination method.

Receptor binding studies

One day after reseeding into 12-well plates, binding experiments were carried out in buffer with cells in monolayer for 1 h at 37°C as described previously (Perlman et al., 1992).

Inositol phosphate formation

One day after transfection, cells in monolayer in 12-well plates were labeled with 1 μCi of *myo*-[^{3}H]inositol/ml. Stimulation of inositol phosphate formation was measured 1 day later for 1 h at 37°C by methods previously described (Perlman et al., 1994a).

RESULTS AND DISCUSSION

Simulated annealings

Selection of relevant structures

Fourteen simulated annealings were conducted as described above. At the end of the cooling period, a 100-ps MD simulation at 300 K was conducted on each of the structures. The calculated root mean square deviations (RMSD) from the structure at the beginning of the 300 K simulation along each of the 14 100-ps trajectories reveal that the structures have stabilized after 20 to 40 ps and suggest that they fluctuate around an average. We calculated the average RMSDs and the standard deviations in the final 60–80 ps with respect to the 100-ps structure to demonstrate that the deviation and the fluctuations are small. The RMSDs of each of the 14 simulations were (in Å) 1.2 ± 0.1, 1.1 ± 0.1, 1.3 ± 0.3, 1.4 ± 0.5, 1.8 ± 0.3, 1.8 ± 0.3, 1.4 ± 0.3, 1.5 ± 0.2, 1.0 ± 0.3, 1.7 ± 0.2, 1.5 ± 0.3, 1.4 ± 0.1, 1.7 ± 0.3, and 1.7 ± 0.5. Consequently, the last 60- to 80-ps interval was used to obtain the energy-minimized average structures for each simulation. To examine whether the structures fall in the allowed range of conformational space, the ϕ, ψ, and $\chi 1$ angles were measured in the 14 energy-minimized average structures for all residues in the extracellular loops. The Ramachandran plot is presented in Fig. 2, and an analysis performed with PROCHECK (Laskowski et al., 1993; Morris et al., 1992) reveals that 86% of the observed values are clustered in the sterically allowed regions. The histograms of the distributions of the $\chi 1$ angles presented in Fig. 3 show that the majority of the residues in the loops take conformations characteristic of non-α/β structures (McGregor et al., 1987). This analysis confirms that the

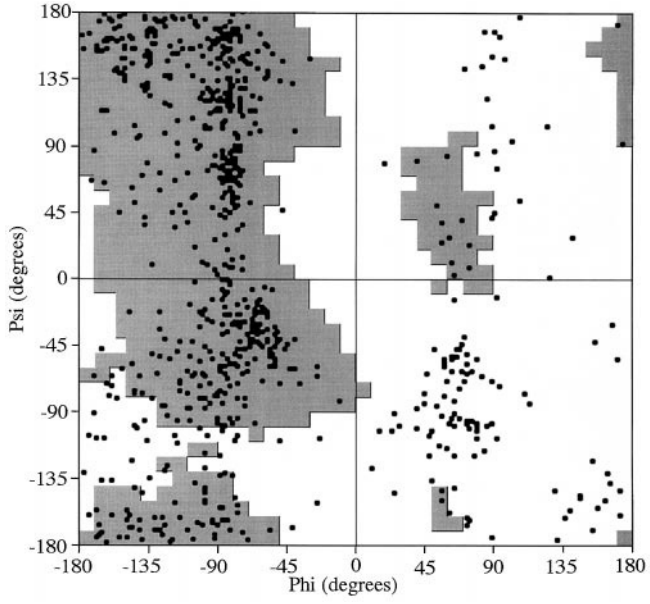

FIGURE 2 Ramachandran plot showing combinations of the conformational angles ϕ and ψ of residues in the extracellular loops of the 14 energy-minimized average structures obtained in separate annealings. The shaded area corresponds to the most favored and additional allowed regions defined in PROCHECK (Laskowski et al., 1993; Morris et al., 1992).

simulated annealing protocol followed by a constant temperature simulation generates structures that are in a conformationally allowed region and that belong to proteins without a well defined secondary structure.

Clustering

The 14 minimized averaged structures were subsequently clustered according to their pairwise RMSDs for Cα values into conformational families employing the program Xcluster (Shenkin and McDonald, 1994). At a level of 1.89-Å resolution, the cluster analysis reveals the existence of one major and six minor families. The major family is composed of seven members, suggesting that 50% of the structures have a common fold of the extracellular loops. Fig. 4 shows the general fold of the extracellular loops in two members of the major family (top) and in two members of the minor families (bottom). Clearly, among members of the major family, EC1 is always positioned over helices 2 and 3 and packs against EC2A. EC2A consistently starts with a short loop at its amino terminus and moves toward helix 1 or 7 to maintain the disulfide bridge at the carboxy terminus of EC1. EC2B (Gly180 to Ser189) lies below EC2A and thus closer to the helical bundle. It spans the area between helices 3, 5, 6, and 7. Consequently, it is positioned above the transmembrane binding pocket. The backbone of EC3 is U-shaped, but the side chains form close contacts with those of EC2B, which occlude access to the transmembrane binding pocket. Members of the various minor families differ from those of the major family mostly by EC1, which often lies outside helices 1 and 2 (when viewed from the extracellular side). Consequently, EC2A moves toward helices 1 and 3 (as opposed to 1 and 7 in the major family), which results in a less convoluted EC2B. Differences in folding patterns of the extracellular loops clearly emerge across the various families as exemplified in Fig. 4, although the conserved patterns described above remain within clustered groups. The conservation of folding patterns among 50% of the structures obtained in this work suggest that, despite the short annealing and simulation time, the simulated annealing protocol employed here is quite satisfactory.

As mentioned above, examination of the structures from the major family reveals that the loops pack against each other, generating an apparently impenetrable structure for the ligand to access its transmembrane binding pocket. EC1 tightly packs against EC2A, whereas EC2B packs against both EC2A and EC3, hence keeping EC2A from interfacing with EC3. As the transmembrane binding pocket is positioned underneath EC2B and EC3, the interface between EC2B and EC3 becomes the major obstruction for the ligand to access the transmembrane domain from the outside. The absence of a clear path into the transmembrane binding pocket raises the question of the molecular steps between the dissociated ligand from the receptor and its final bound state in the transmembrane domain. Two possibilities can be suggested. In one, a critical step is the recognition of specific elements on the surface of the receptor by the ligand. Such an interaction between the ligand and the surface binding pocket could induce a conformational change in the loops that may open a path into the transmembrane region. Ligand binding to surface pockets is well known in examples of peptide-antibody complexes (Garcia et al., 1992; Stanfield and Wilson, 1995). Furthermore, the interaction between a nonamer peptide and the antibody recognition domain, Fab 17/9, was shown to induce dramatic changes in the position of the loops and domains in the antibody (Rini et al., 1992). Thus, an initial complex between TRH and a surface binding site could change the interaction between the residues that obstruct access to the transmembrane region and trigger the opening of a pathway to access the binding pocket.

An alternative mechanism could involve a spontaneous unfolding motion of the loops that would result in an intermittent opening of a pathway to the transmembrane domain. A recent study performed on the access of small ligands into the core of lysozyme (Feher et al., 1996) suggests that relatively large-scale fluctuations of backbone atoms are required for entry of the ligands into the core of the protein. Such a mechanism could play a role in TRH access to its transmembrane binding pocket.

The surface binding mechanism depends on the presence of a recognition site with properties complementary to those of TRH. The large-scale fluctuation mechanism depends on the vibrational properties of the loops and in particular on an anti-correlated motion of EC2B and EC3 that may open a pathway to the transmembrane binding pocket. Below we present an analysis of the recognition elements on the surface of the receptor. To address these questions, we have

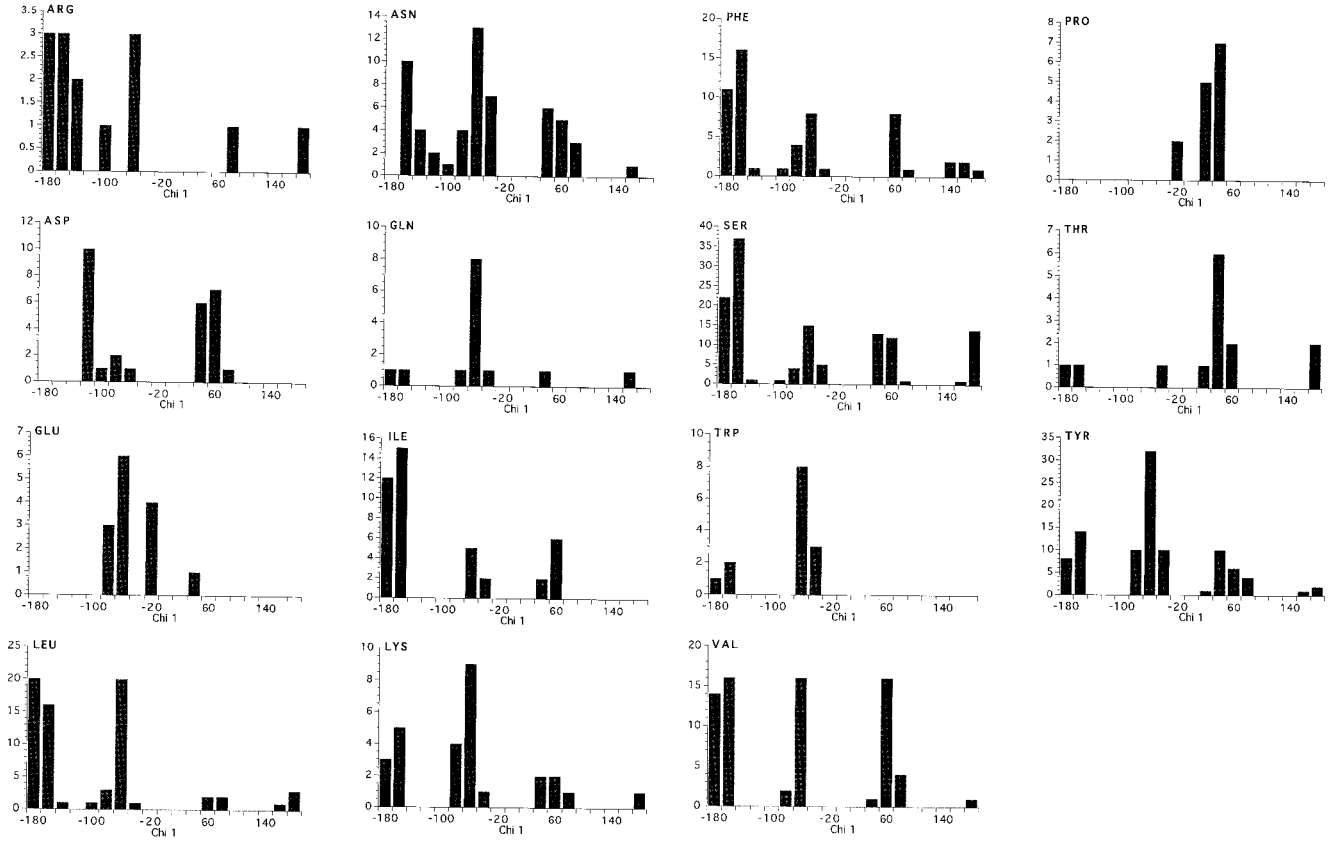

FIGURE 3 Histograms of $\chi 1$ distribution for residues in the extracellular loops of the 14 energy-minimized average structures obtained in separate annealings. The bars represent 20° intervals.

performed a 1-ns simulation of the extracellular loops anchored to the frozen transmembrane helices. The starting structure employed for this simulation was one of the members that belongs to the major family obtained from the clustering analysis described above. The structure derived from this 1-ns simulation performed at 300 K was an energy-minimized average structure obtained over the last 500 ps of the trajectory. The calculated average RMSD from the structure obtained at 1 ns during the last 500 ps was 1.14 ± 0.14 Å. The small value of the average RMSD and in particular the small standard deviation reflect the fact that the structure has reached a stable conformation over the last 500 ps of the trajectory. The trajectory was also used to perform a quasi-harmonic analysis of the vibrational properties of the loops.

Recognition elements

Characterization of surface binding properties could be divided into two elements. One is the presence of cavities that present a sterically confined space into which the ligand can fit. The other are the electrostatic properties of such a cavity, which show complementarity to the charge distribution in the ligand. These two elements could be associated, respectively, with the entropic driving force due to release of bound water in the cavity and due to the enthalpic driving force generated by specific polar interactions between the ligand and the cavity. To examine these properties we have constructed the molecular surface of this average structure employing the program GRASP (Nicholls, 1992). The surface curvature is presented in Fig. 5 *A*. The surface has several cavities, one of which is positioned in the interface between loop 2 (EC2B) and loop 3 (EC3). The cavity defined by Tyr181, Lys182, Arg185, Asn186, Tyr187, Asn289, Ser290, and Phe296 (Fig. 5 *B*) is consistently observed on the surface of the average structures of members of the large clustered family (not shown here) and persists in snapshots of the structures extracted from long MD trajectories. This suggests the formation of a surface binding site that can become a putative entry point into the transmembrane binding pocket.

To probe the significance of the putative surface binding pocket, we mutated three residues that define the boundaries of the pocket. R185A, N289A, and Y181F mutant TRH receptors were constructed and tested. Results with N289A and Y181F TRHRs are described elsewhere (Perlman et al., in press, 1998). The affinity of N289A was 10-fold lower than that of native TRHR. The affinity of Y181F was too low to be determined in binding experiments and was estimated from the potency of TRH in activating the hydrolysis of phosphatidyl inositol (EC50s). We have shown previously (Perlman et al., 1994a,b, 1996) that the relative EC50

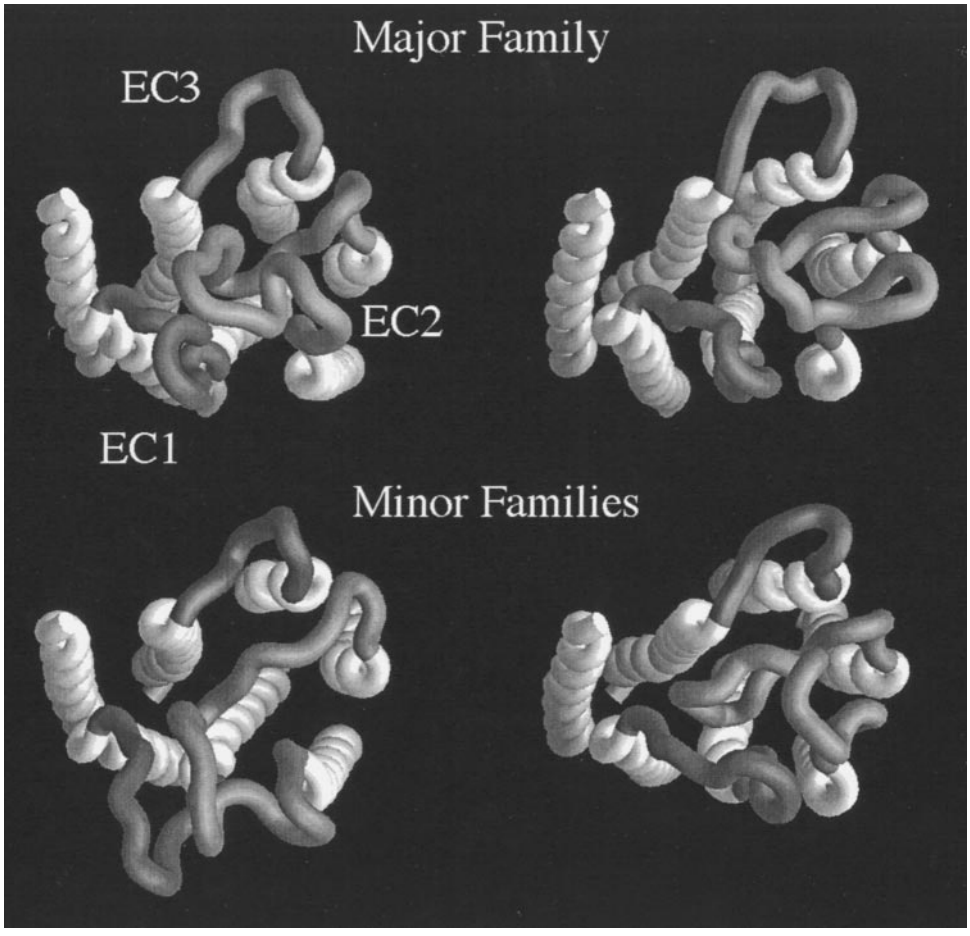

FIGURE 4 Folding patterns of the extracellular loops among two members of the major family (see text for details) and members of two individual minor families. Conserved patterns remain within clustered groups, but significant variations can be observed across the various families.

of a mutant receptor compared with the wild-type receptor is a good representation of the change in relative affinities. Compared with wild-type TRHR, the EC50 of TRH for Y181F was 3300-fold higher. The affinity of R185A TRHR was also too low to be determined in binding experiments and was estimated through activation experiments as described in Materials and Methods. The EC50 of TRH for R185A TRHR was 2200 nM (1700–2800 nM, 95% confidence interval; $n = 4$) compared with 0.59 nM for wild-type TRHR (Perlman et al., 1997), indicating an affinity decrease of 3700-fold. These results indicate that mutations of Tyr181, Arg185, and Asn289 lower the affinity of TRH to the mutant receptors and suggest that these residues are important for binding TRH. They are consistent with the idea that Tyr181, Arg185, and Asn289 are part of the proposed surface binding pocket.

Some of the residues of the cavity defined above have been mutated previously by Han and Tashjian (1995a,b). Mutant receptors Y181F, N289A, and S290A were found to exhibit a significant loss of binding compared with native TRHR when assayed with a single dose of *N*-τ-methylhistidine-TRH. The binding of mutant N186A was the same as the native receptor. This further supports the suggestion that the surface binding site is defined by residues in EC2B and EC3.

To further explore the properties of this cavity, we have calculated the electrostatic potential on the surface defined by the extracellular loops (Fig. 5 *C*). The map shows that the center of the putative surface binding pocket has a positive electrostatic potential surrounded by areas of negative potential. The positive potential is generated primarily by Lys182 (EC2B) whereas the surrounding negative potential is produced by Glu298 (EC3) and polar residues in EC3 (e.g., Ser290 and Ser293). It is difficult to present the charge distribution on TRH, which could demonstrate its complementarity to the electrostatic potential of the putative binding pocket. Examination of the TRH structure shown in Fig. 5 *D* reveals that the oxygens of three of the four carbonyls point in the same direction. Such an arrangement of negatively charged groups could be attracted by the positive potential generated by Lys182. The structure also shows that the positively charged groups (N—H) are distributed on the periphery of the molecule and could be attracted by the negative potential generated by the polar groups around the pocket. These properties could be illustrated by the electrostatic potential on the molecular surface of TRH as shown in Fig. 5 *D*. The negative potential defined by the backbone carbonyls could interact with the positive charge of Lys182 (EC2B) and the positive potential defined by the histidine, the pyroglutamyl, and the terminal ProNH2 could be inter-

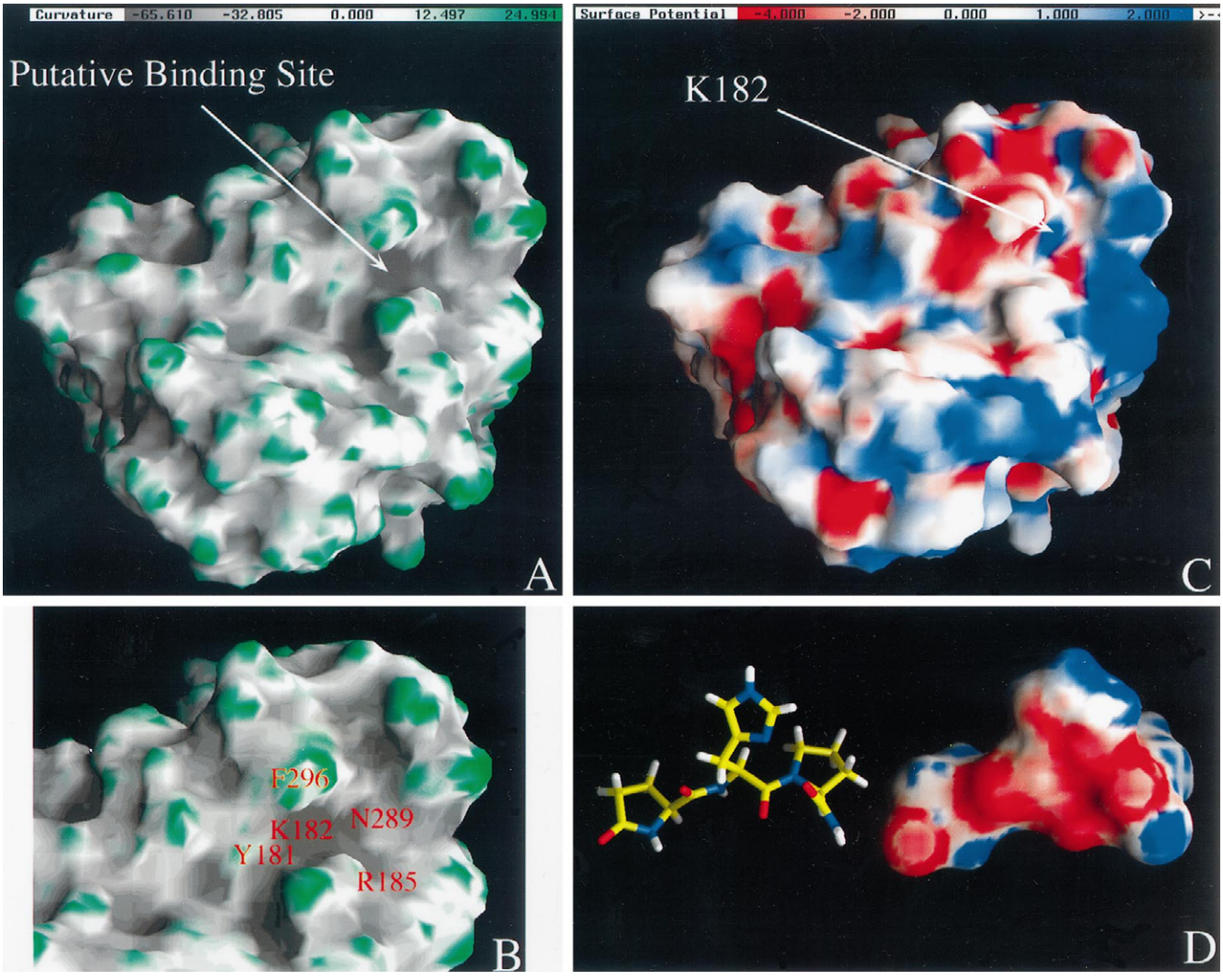

FIGURE 5 (*A*) Molecular surface curvature of the extracellular domains of TRHR. Negative curvature (gray) corresponds to concave areas, and positive curvature (green) to convex areas. The most significant crevice, which could suggest the formation of a putative entry point on the molecular surface, is positioned at the interface between loop 2 and loop 3. (*B*) Selected residues that define the putative surface binding site. See text for the complete list of residues that define this site. (*C*) Electrostatic potential of the extracellular domains of TRHR mapped onto the molecular surface. Blue areas correspond to positive potential and red areas to negative potential. At the bottom of the putative binding site, Lys182 generates a positive potential surrounded by areas of negative potentials. (*D*) The structure of TRH and the electrostatic potential on its molecular surface. The molecule can be superimposed on the putative binding pocket by a 180° rotation around the vertical axis of the figure.

acting with the negative charges of Ser290, Ser293, and Glu298 (EC3). Our finding that the affinity of E298A TRHR is threefold lower than wild-type TRHR is consistent with a role for Glu298 in binding (Perlman et al., 1998).

It is interesting to examine the putative binding pocket at the atomic level. The intention of this analysis is to provide a molecular basis for the properties of the pocket and will of course depend on the validity of the structure of the loops. Nevertheless, the consistency between the static and dynamic properties of the pocket (see below) provides a point of interest. A detailed picture is shown in Fig. 6. Lys182 (EC2B) lines the bottom of the cavity by forming H-bonds with the side chains of Ser290 (EC3) and Glu298 (EC3), which is itself H-bonded to the OH group of Tyr181 (EC2B). This intricate hydrogen bonding network between EC3 and EC2B results in the tight packing of the loops described above. To probe the extent of permanency of the hydrogen-bonding network, we have examined the fraction of time the interactions between Lys182, Glu298, and Ser290 (see Fig. 6) are interrupted in the course of a 1-ns simulation. The Tyr181-Glu298 interaction, monitored by the (Tyr181)-O_ζ-H···O_ϵ(Glu298) distance, was maintained at $\leq$1.8 Å 99% of the time. The (Lys182)N_ζH···O_ϵ(Glu298), and (Lys182)N_ζH···O_γ(Ser290) were maintained 88% and 45% of the time, respectively. Thus, Glu298 in this model of the extracellular loops is responsible for maintaining the rigid scaffold of the putative binding pocket observed from the snapshots of the molecular surface as a function of time. To test the effect of the continuum dielectric representation of the solvent, we conducted simulations with a distance-

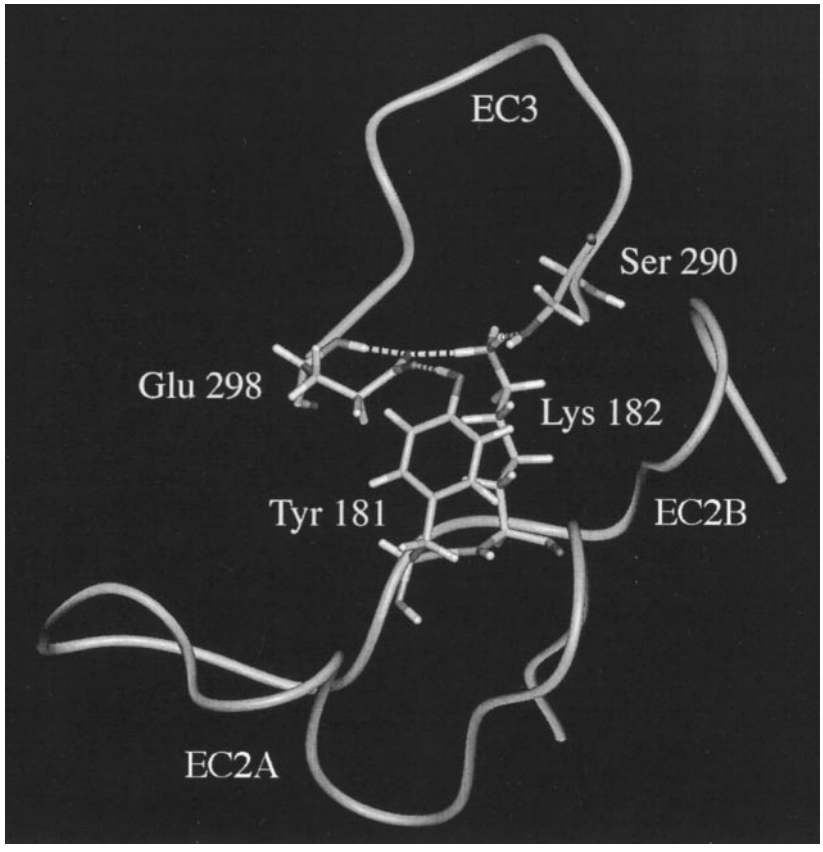

FIGURE 6 Atomic representation of selected residues in the putative surface binding pocket. Lys182 (EC2B) lines the bottom of the pocket and forms H-bonds with the side chains of Ser290 (EC3) and Glu298 (EC3). Glu298 is H-bonded to the hydroxyl group of Tyr181 (EC2B).

dependent dielectric function that has a fourfold stronger dependence on distance. We have constructed a molecular surface on the minimized average structure from this simulation. The surface curvature and surface electrostatic potential were nearly identical to those obtained from simulations with $\epsilon = r$. We conclude that the distance-dependent continuum dielectric function did not significantly bias the simulations in favor of polar residue packing. However, additional studies with an explicit representation of the solvent will be needed to address the significance of the approximate representation of the solvent.

The presence of a surface cavity with size and electrostatic properties complementary to those of the ligand is consistent with the suggestion that an initial recognition of TRH occurs on the surface of the receptor formed by the extracellular loops. This suggestion is supported by the observation that residues that define portions of the cavity (e.g., Tyr181, Arg185, and Asn289) also affect binding affinity. The corollary of this suggestion is that TRH binds to the surface before accessing the transmembrane binding pocket.

Dynamic analysis

Quasi-harmonic analysis

The static representation described heretofore cannot address the question of the involvement of loop motion in guiding the ligand to its binding pocket in the transmembrane domain. To address this problem, we chose to use quasi-harmonic analysis to describe loop motions and the correlated displacements between different loops. Such an approach has been taken previously in studies of deoxymyoglobin (Seno and Go, 1990a,b), lysozyme (Amadei et al., 1993; Horiuchi and Go, 1991), thermolysin (van Aalten et al., 1995a), and retinol binding protein (van Aalten et al., 1995b) employing normal mode analysis, quasi-harmonic analysis, and essential dynamics. They revealed that low-frequency motions in the essential subspace of the protein are related to its functional behavior. For instance, Amadei et al. (1993) show that the essential dynamics of lysozyme can be related to the functional behavior of the protein, such as opening and closing of the active site, whereas Horiuchi and Go (1991) extracted the motions that represented the hinge bending of the two lobes forming the active site of the cleft of lysozyme. van Aalten et al. (1995b) also used essential dynamics to study differences in dynamics between the holo and apo forms of the cellular retinol-binding protein. Their results reveal inhibition of essential motions upon ligand binding and show large correlated motions of retinol with regions of the protein, pointing to a possible retinol entry/exit site.

In the present work, a quasi-harmonic analysis was performed on a trajectory from a 1-ns MD simulation on the structure shown in Fig. 6, which belongs to the major family described above. As the transmembrane binding pocket in TRHR is located directly underneath EC2B and EC3, and

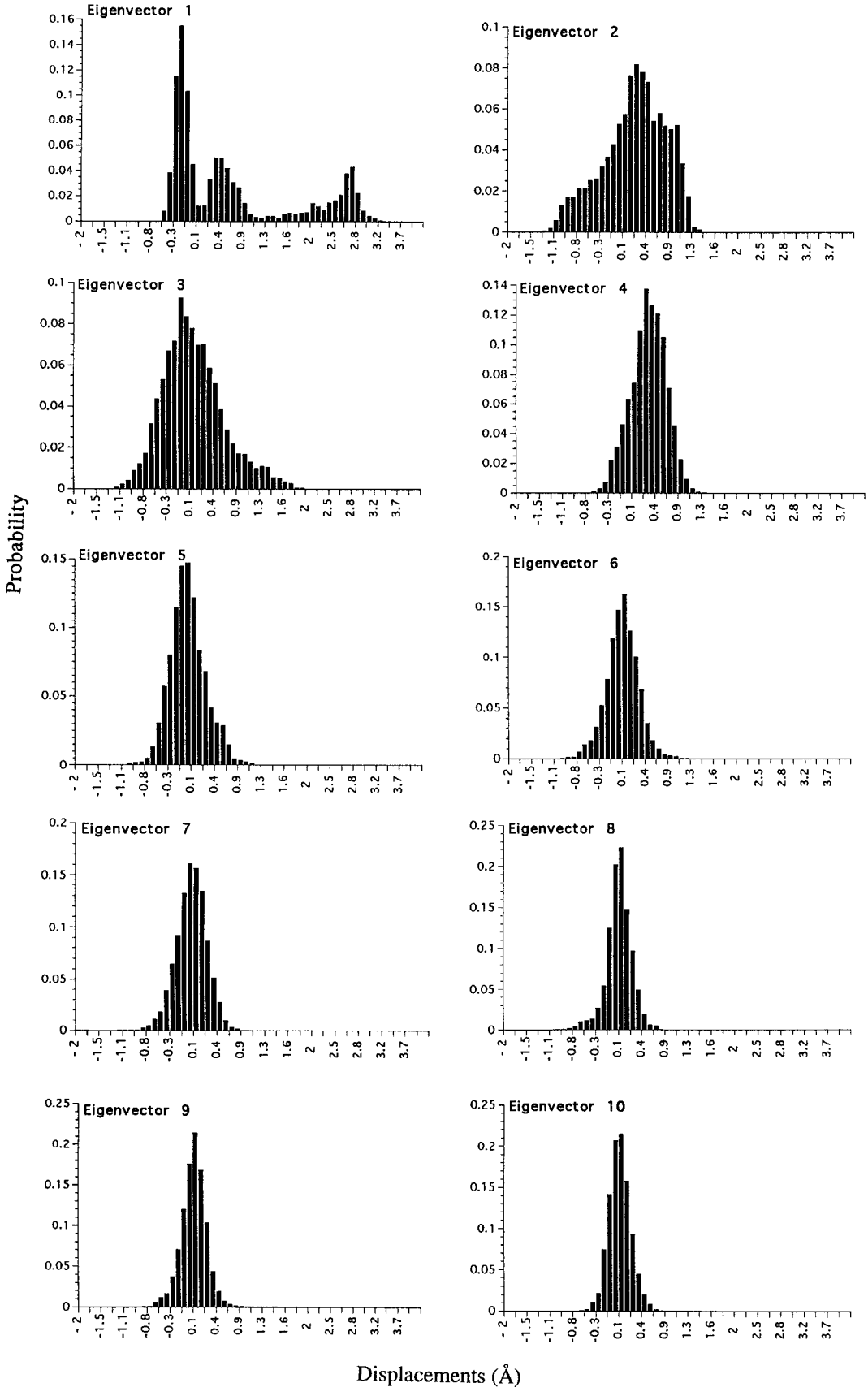

FIGURE 7 Sampling distribution probabilities for displacements in the first 10 eigenvectors. Each vertical bar represents the probability of finding the system in a displacement interval of 0.1 Å as a function of displacement.

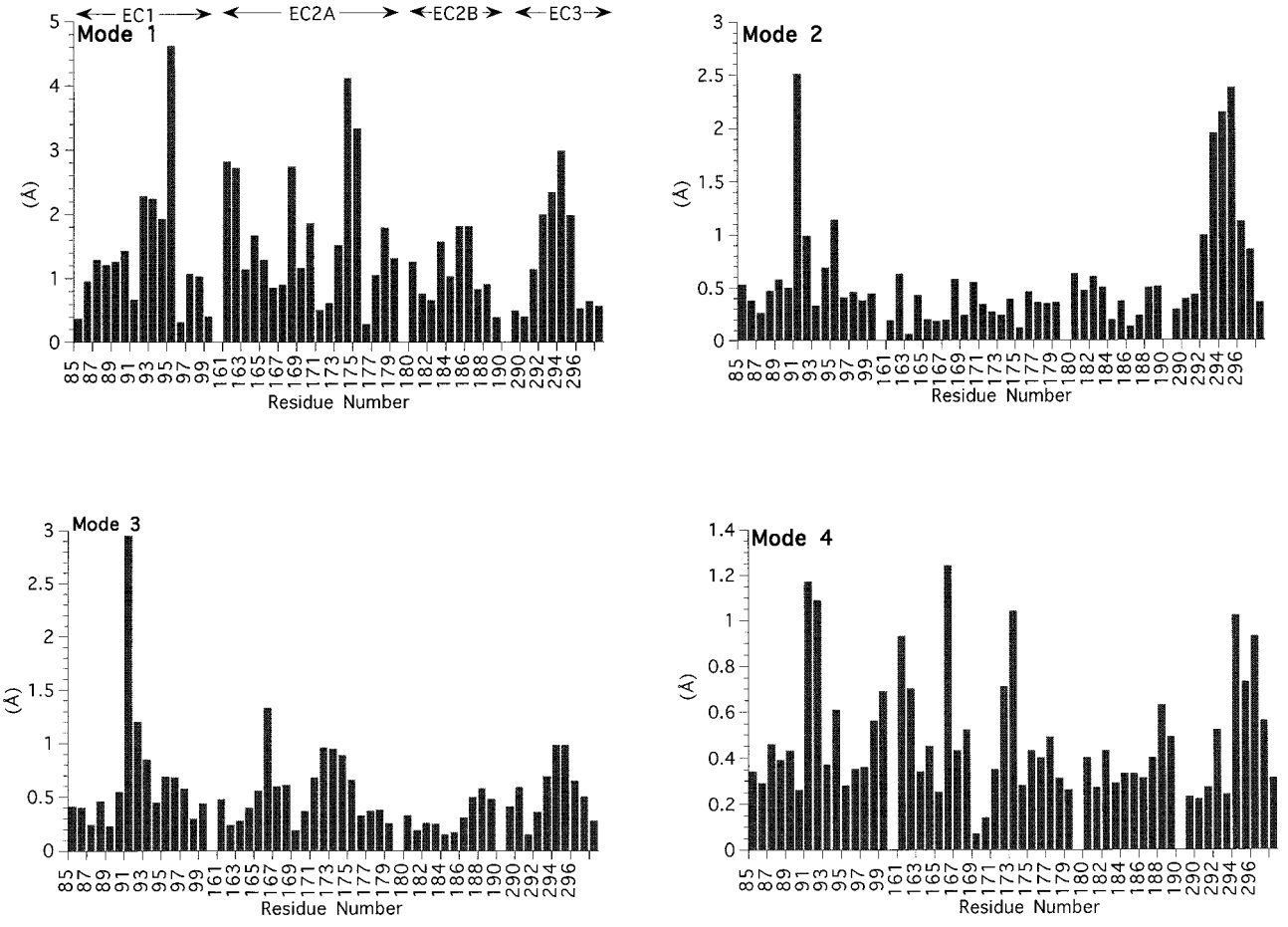

FIGURE 8 Excursion distances of the center of mass of each extracellular residue in the first four modes of vibration. Note the changes in the vertical scale in the different modes.

the putative surface binding site is defined by residues from these loops, we focus our discussion on the motions of this extracellular domain.

The amount of motion associated with the calculated eigenvectors can be represented by the cumulative positional fluctuation defined as follows (Amadei et al., 1993):

$$\Phi_n = \frac{\Sigma_{i=1}^{n} \lambda_i}{\Sigma_{i=1}^{N} \lambda_i},$$

where λ_i is the eigenvalue, n is the cumulative index of the normal modes, and N is the total number of vibrational modes. The cumulative positional fluctuation, Φ_n, describes the fraction of the total variance from the average position accounted by the first n vibrational modes. Our results show that 83% of the total variance due to motion of the extracellular domain is described by the first 10 eigenvectors of a total of 162. These modes carry little vibrational energy but are responsible for large displacements. In contrast, the higher vibrational modes, although they carry most of the zero-point vibrational energy due to their higher frequencies, contribute very little to the total motion.

Projection of the trajectory onto the individual eigenvectors allows for a more detailed description of the motion accounted for by each eigenvector. The sampling distributions for the displacements of motions from the average along the first 10 eigenvectors in the 1-ns trajectory are presented in Fig. 7. They indicate that non-Gaussian distributions are found among the first few eigenvectors. Thus, the motions described by these vibrational modes are highly anharmonic. Anharmonic effects are of importance in larger-scale collective motions in proteins (Ichiye and Karplus, 1991) because they are associated with low-energy vibrations and they increase the overall amplitude of fluctuation of the protein over that predicted by a harmonic vibration (Teeter and Case, 1990). Thus, we focus on the low-frequency modes to sample large-scale motions of possible biological importance in the TRHR.

Distance analysis

To identify the residues that contribute to the large-scale motions in the low-frequency modes, we have calculated the excursion distances of the center of mass of each residue of the extracellular loops in the first four vibrational modes. They were obtained from the projected trajectories of each of the first four modes and represent the distances traveled

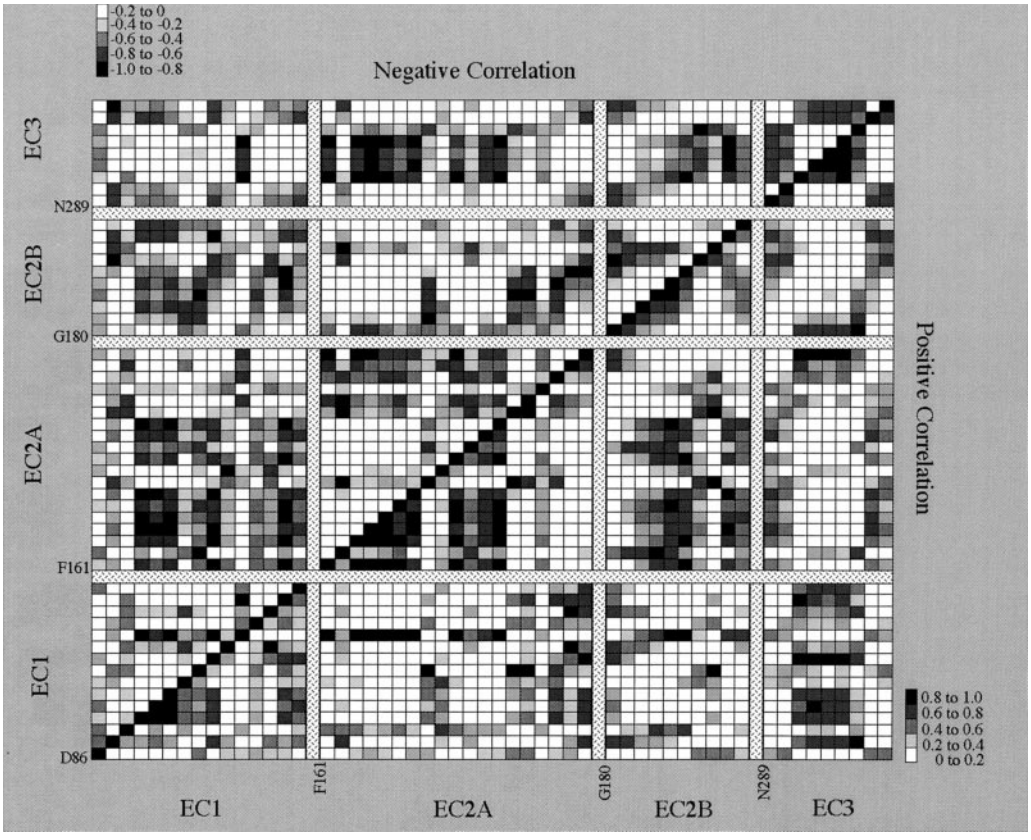

FIGURE 9 Correlation between residues in the extracellular loops in the 1-ns trajectory projected along the first eigenvector. The upper left triangle represents negative correlations, and the lower right triangle represents positive correlations. The degree of correlation is color coded from white to black, with white representing no correlation and black representing very strong correlation.

by each center of mass within the mode and are shown in Fig. 8. In the first mode, fluctuations in excess of 2.5 Å are observed, whereas, consistent with the decrease in the fractional contribution to the total displacement, the individual contributions from the following modes significantly decrease as the frequency increases. As expected, residues identified as being part of the static recognition elements (i.e., Tyr181, Lys182, and Glu298) undergo very little motion along any of the eigenvectors because they participate in an extensive hydrogen bond network (see above). On the other hand, the stretch of amino acids I183–N186 in EC2B and F291–P295 in EC3 show large excursions exceeding 1 Å.

Excursion analysis alone cannot provide information about the directionality of these motions. To assess whether the motions are correlated, we present in Fig. 9 a correlation matrix calculated for the 1-ns trajectory projected along the first eigenvector. The lower triangle presents positive correlation, whereas the negative correlation is presented in the upper triangle. Significant positive correlation in the first mode is observed between EC1, the carboxy end of EC2A, and the central portion of EC3. There is also a significant positive correlation between the amino portion of EC2A and EC2B. As positive correlation indicates motions in the same direction, these results suggest that a set of interactions between EC1 and the carboxy end of EC2A, as well as between the amino end of EC2A and EC2B are constraining their motion to one unit. This is consistent with the close packing observed in the analysis of the static averaged structure (see above). Negative correlations indicate motions in opposite directions. Among the most pronounced are the negative correlation between the amino end of EC2A and, respectively, with EC1 on one side and EC3 on the other. However, as these portions are not in contact, they represent large-scale motions that may not be relevant to the interaction of the ligand with the surface binding site. The only negative correlation of adjacent loops is between EC2B and EC3. It is clear from this analysis that residues in EC2B, especially Ile183 through Tyr188, and in EC3, Phe291 through Pro295, have anti-correlated motions. Such movements could result in an opening and closing of a putative entry channel into the receptor, and therefore identification of the moving residues could suggest a mechanism for ligand entry. In addition, the amplitude of this motion is sufficiently large to observe large fluctuations in interresidue distances. To illustrate the motion, atomic displacement vectors in the first vibrational mode are presented in Figure 10 *A*. The apparent mobility of the loops is reflected in the displacement vectors, which clearly outline an anti-corre-

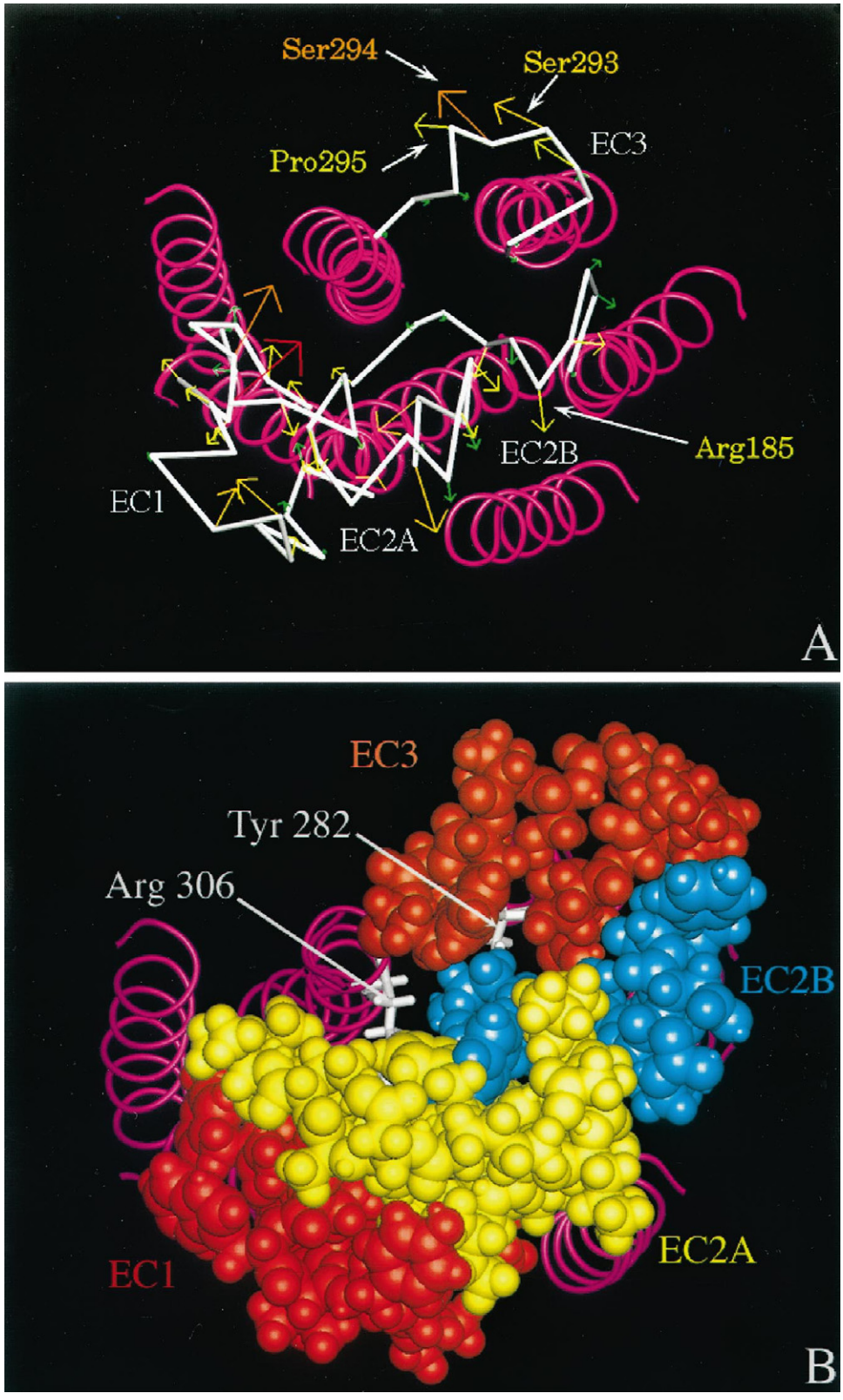

FIGURE 10 (*A*) Displacement vectors in the first mode represented by color-coded arrows. The vectors were multiplied by a factor of 50 to ensure that all the arrows are visible. The arrows give a good representation of the anti-correlated motion between EC2B (e.g., Arg185) and EC3 (e.g., Ser293, Ser294, and Pro295). (*B*) Space-filling representation of the minimized average structure of the 1-ns simulation over the last 500 ps. The transmembrane binding pocket represented by Tyr282 and Arg306 (white) is clearly visible when the side chains of Y181 and Phe296 are rotated (see text).

lated motion between EC3 and EC2B, more specifically between the motions of the segments Leu292-Pro295 and Ile183-Asn186. Analysis of the projected trajectory onto the first quasi-normal mode trajectory reveals that the distance between the center of mass of Tyr188 and Ser294 oscillates between 10.13 and 13.84 Å, whereas that between Ser184 and Ser294 oscillates between 11.05 and 14.28 Å. Large motions with low frequencies suggest "soft" degrees of freedom in the extracellular loops. This suggest that even weak interactions with the ligand can trigger changes in

local conformation that will result in a possible opening of the channel.

Connection between static and dynamic definition of the putative entry point

The residues in the putative binding pocket can be classified into two different categories. One set of residues, Tyr181, Lys182, Ser290, and Glu298, has low mobility because of the hydrogen bond interactions among these residues and thus can be considered as a static component responsible for TRH recognition. Another set of residues, Ile183-Asn186 in EC2B and Leu292-Pro295 in EC3, are engaged in large anti-correlated positional fluctuations. We suggest that the interaction of TRH with the surface binding pocket contributes to the disruption of the interactions between the static residues, introducing a possibility of enhancing the anti-correlated vibrational motions of the loops. Such a mechanism will ensure the specificity of ligand recognition defined by the static residues as well as provide a way of inducing conformational changes in the loops to open a pathway into the transmembrane pocket. To examine whether conformational changes in the extracellular residues can give rise to such a pathway, we have manually rotated the side chains of residues at the rim (Tyr181 (χ_1 by 100°) and Phe296 (χ_1 by 100°)) and at the bottom (Lys182) (χ_1 by 168° and χ_2 by −170°) of the surface binding pocket. As shown in Figure 10 *B*, residues in the transmembrane binding pocket (i.e., Tyr282 and Arg306) become exposed when such rotations of the surface residues are induced. Additional studies of the hormone-receptor complex in which TRH is bound to the surface of TRHR will have to be conducted to elucidate the conformational changes that can be induced by such an interaction.

CONCLUSION

In the present work, we have focused our attention on the extracellular domain of the TRH receptor and showed an apparent conserved pattern in the folding of the loops through the use of simulated annealing and cluster analysis. Among the most important conclusions from this study is that a static representation of the extracellular loops does not present an apparent access pathway into the transmembrane binding pocket but suggests the formation of a putative entry point on the surface of the receptor with steric and electrostatic properties complementary to those of TRH. A dynamic representation supported by quasi-harmonic analysis shows an anti-correlated motion between EC2B and EC3, which could result in the formation of a path for TRH to access its binding pocket. Although this path appears to be occluded by Lys182, which hydrogen bonds to residues in EC3, an interaction with TRH could induce conformational changes in the surface binding pocket to expose an entrance to the transmembrane binding site.

Our results therefore suggest that both recognition of specific elements on the receptor surface and loop motion are necessary for the ligand to access its transmembrane binding pocket. These results point to a multi-step mechanism for binding of TRH to its receptor. TRH initially interacts with residues within the extracellular loops and then moves into the transmembrane binding pocket, aided by the motion of the loops. We suggest that sequential binding interactions similar to these between TRH and its receptor may occur between small ligands and all GPCRs and that small ligands may be guided into transmembrane binding pockets by extracellular loops.

We thank the National Energy Research Supercomputer Center at Lawrence Livermore National Laboratory for providing computational support for this work.

This work was supported by National Institutes of Health Research Fellowship Award DK 09647 (to A.-O. Colson), National Institutes of Health Physician Scientist Award DK 02101 (to J. H. Perlman), and grant DK 43036 (to M. C. Gershengorn and R. Osman).

REFERENCES

Amadei, A., B. M. Linssen, and H. J. C. Berendsen. 1993. Essential dynamics of proteins. *Proteins.* 17:412–425.

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

Ballesteros, J. A., and H. Weinstein. 1995. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G-protein-coupled receptors. *In* Receptor Molecular Biology. Academic Press, New York. 366–428.

Brooks, B., and M. Karplus. 1983a. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA.* 80:6571–6575.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983b. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Colson, A. O., J. H. Perlman, R. Jain, L. A. Cohen, R. Osman, and M. C. Gershengorn. 1997. Sequential binding of TRH to the TRH receptor: a model for small ligand binding to G-protein coupled receptors. *Biophys. J.* 72:A122.

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

Feher, V. A., E. P. Baldwin, and F. W. Dahlquist. 1996. Access of ligands to cavities within the core of a protein is rapid. *Nature Struct. Biol.* 3:516–521.

Findlay, J., and E. Eliopoulos. 1990. Three-dimensional modelling of G protein-linked receptors. *Trends Pharmacol. Sci.* 11:492–499.

Garcia, K. C., P. M. Ronco, P. J. Verroust, A. T. Brünger, and L. M. Amzel. 1992. Three-dimensional structure of an angiotensin II-fab complex at 3 Å: hormone recognition by an anti-idiotypic antibody. *Science.* 257:502–507.

Gershengorn, M. C., and R. Osman. 1996. Molecular and cellular biology of thyrotropin releasing hormone receptors. *Physiol. Rev.* 76:175–191.

Gershengorn, M. C., and C. N. Thaw. 1991. Regulation of thyrotropin-releasing hormone receptors is cell type specific: comparison of endogenous pituitary receptors and receptors transfected into non-pituitary cells. *Endocrinology.* 128:1204–1206.

Go, N., T. Noguti, and T. Nishikawa. 1983. Dynamics oa a small globular protein in terms of low-frequency vibrational modes. *Proc. Natl. Acad. Sci. USA.* 80:3696–3700.

Han, B., and A. H. Tashjian. 1995a. Importance of extracellular domains for ligand binding in the thyrotropin releasing hormone receptor. *Mol. Endocrinol.* 9:1708–1719.

Han, B., and A. H. Tashjian. 1995b. Identification of Asn289 as a ligand binding site in the rat thyrotropin releasing hormone (TRH) receptor as determined by complementary modifications in the ligand and receptor: a new model for TRH binding. *Biochemistry.* 34:13412–13422.

Horiuchi, T., and N. Go. 1991. Projection of Monte Carlo and molecular dynamics trajectories onto the normal mode axes: human lysozyme. *Proteins.* 10:106–116.

Ichiye, T., and M. Karplus. 1991. Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins*. 11:205–217.

Kyle, D. J., S. Chakravarty, J. A. Sinsko, and T. M. Stormann. 1994. A proposed model of bradykinin bound to the rat B2 receptor and its utility for drug design. *J. Med. Chem.* 37:1347–1354.

Laakkonen, L. J., F. Guarnieri, J. H. Perlman, M. C. Gershengorn, and R. Osman. 1996. A refined model of the thyrotropin releasing hormone (TRH) receptor binding pocket: novel mixed mode Monte Carlo/stochastic dynamics simulations of the complex between TRH and TRH receptor. *Biochemistry.* 35:7651–7663.

Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.

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

McGregor, M. J., S. A. Islam, and M. J. E. Sternberg. 1987. Analysis of the relationship between side-chain conformation and secondary structure in globular proteins. *J. Mol. Biol.* 198:295–310.

Morris, A. L., M. W. MacArthur, E. G. Hutchinson, and J. M. Thornton. 1992. Stereochemical quality of protein structure coordinates. *Proteins.* 12:345–364.

Nicholls, A. 1992. GRASP: graphical representation and analysis of surface properties. Columbia University, New York.

Perlman, J. H., A. D. Colson, R. Jain, B. Czyzewski, L. A. Cohen, R. Osman, and M. C. Gershengorm. 1998. Role of the extracellular loops of the thyrotropin-releasing hormone (TRH) receptor. Evidence for an initial interaction with TRH. *Biochemistry*. In press.

Perlman, J. H., A. O. Colson, W. Wang, K. Bence, R. Osman, and M. C. Gershengorn. 1997. Interactions between conserved residues in transmembrane helices 1, 2 and 7 of the thyrotropin-releasing hormone receptor. *J. Biol. Chem.* 272:11937–11942.

Perlman, J. H., L. Laakkonen, R. Osman, and M. C. Gershengorn. 1994a. A model of thyrotropin-releasing hormone (TRH) receptor binding pocket: evidence for a second direct interaction between transmembrane helix 3 and TRH. *J. Biol. Chem.* 269:23383–23386.

Perlman, J. H., L. Laakkonen, R. Osman, and M. C. Gershengorn. 1995a. Distinct roles for arginines in transmembrane helices 6 and 7 of the thyrotropin-releasing hormone receptor. *Mol. Pharmacol.* 47:480–484.

Perlman, J. H., L. J. Laakkonen, F. Guarnieri, R. Osman, and M. C. Gershengorn. 1996. A refined model of the thyrotropin releasing hormone (TRH) receptor binding pocket: experimental analysis and energy minimization of the complex between TRH and TRH receptor. *Biochemistry*. 35:7643–7650.

Perlman, J. H., D. R. Nussenzveig, R. Osman, and M. C. Gershengorn. 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., C. N. Thaw, L. Laakkonen, C. Y. Bowers, R. Osman, and M. C. Gershengorn. 1994b. Hydrogen bonding interaction of the thyrotropin-releasing hormone (TRH) with transmembrane tyrosine 106 of the TRHR receptor. *J. Biol. Chem.* 269:1610–1613.

Perlman, J. H., W. Wang, D. R. Nussenzveig, and M. C. Gershengorn. 1995b. A disulfide bond between conserved extracellular cysteines in the thyrotropin releasing hormone receptor is critical for binding. *J. Biol. Chem.* 270:24682–24685.

Rini, J. M., U. Schulze-Gahmen, and I. A. Wilson. 1992. Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science*. 255:959–965.

Seno, Y., and N. Go. 1990a. Deoxymyoglobin studied by the conformational normal mode analysis. I. Dynamics of globin and the heme-globin interaction. *J. Mol. Biol.* 216:95–109.

Seno, Y., and N. Go. 1990b. Deoxymyoglobin studied by the conformational normal mode analysis. II. The conformational change upon oxygenation. *J. Mol. Biol.* 216:111–126.

Shenkin, P. S., and D. Q. McDonald. 1994. Cluster analysis of molecular conformations. *J. Comp. Chem.* 15:899.

Stanfield, R. L., and I. A. Wilson. 1995. Protein-peptide interactions. *Curr. Opin. Struct. Biol.* 5:103–113.

Straub, R. E., G. C. Frech, R. H. Joho, and M. C. Gershengorn. 1990. Expression cloning of cDNA encoding the mouse pituitary thyrotropin-releasing hormone receptor. *Proc. Natl. Acad. Sci. USA.* 87:9514–9518.

Sylte, I., O. Edvardsen, and S. G. Dahl. 1996. Molecular modeling of UH-301 and 5-HT1a receptor interactions. *Protein Eng.* 9:149–160.

Teeter, M. M., and D. A. Case. 1990. Harmonic and quasi-harmonic descriptions of crambin. *J. Phys. Chem.* 94:8091–8097.

van Aalten, D. M. F., A. Amadei, A. B. M. Linssen, V. G. H. Eijsink, G. Vriend, and H. J. C. Berendsen. 1995a. The essential dynamics of thermolysin: confirmation of the hinge bending motion and comparison of simulations in vacuum and water. *Proteins.* 22:45–54.

van Aalten, D. M. F., J. B. C. Findlay, A. Amadei, and H. J. C. Berendsen. 1995b. Essential dynamics of the cellular retinol binding protein: evidence for ligand induced conformational changes. *Protein Eng.* 8:1129–1135.