Computer Assisted Simulations and Molecular Graphics Methods in Molecular Design
2. Proteinases and Receptor and Transport Proteins

F. M. L. G. STAMATO,[1] M. PAULINO,[2] R. GARRATT,[3] C. M. SOARES,[4] and O. TAPIA[4]
[1] *Chemistry Department, Federal University of Sao Carlos, Brazil*
[2] *Department of Quantum Chemistry, Universidad de la Republica, Uruguay*
[3] *Department of Physics and Informatics, Sao Paulo University, Sao Carlos 13560–970, Brazil*
[4] *Department of Physical Chemistry, Box 532, Uppsala University, 75121 Uppsala, Sweden*

(Received: 8 October 1994; in final form: 12 October 1994)

Abstract. A survey is presented of model building techniques and computer-assisted quantum-chemical and molecular-dynamics simulations, as applied to the study of protease and receptor design, to the determination of the properties of related effectors, agonist and antagonist molecules, and to the design of fatty-acid transport proteins and molecular carriers. Studies covering integral membrane protein design have been reviewed: they include porins, ion channels and G protein coupled receptors. In the transport molecule class, hydrophobic ligand transporters, such as serum and cellular retinoid binding proteins, have been reviewed.

Key words. Molecular modeling, receptor modeling, G protein-coupled receptors, transport molecules, retinoid binding proteins, fatty-acid binding proteins, carriers, porins, proteases.

1. Introduction

Computer assisted simulations and molecular graphics methods have reshaped the traditional methodology used in molecular engineering and, in particular, the process of molecular design and modeling. In the preceding paper [1], the bases of enzyme catalysis were surveyed and some principles of structure-based drug design [2] were discussed. The applications reviewed there concerned the molecular design of enzyme inhibitors and transition-state analogs. Theoretical and modeling studies of receptor and transport proteins are surveyed in the present paper.

Section 2 examines some aspects of the homology-based molecular modeling of serine proteases and its use in the understanding of substrate specificity. Section 3 presents a broad view of the modeling of receptors linked to second messenger production through G proteins (guanine nucleotide-binding regulatory proteins). This is one of the most important receptor systems discovered so far and, as pointed out by Blundell and coworkers [3], in spite of a high degree of sequence homology in the receptors themselves, there is an incredible variety of ligands which bind to them and are associated to a diversity of cellular effects. Section 4 is dedicated to examining the work done on transport molecules. Free-energy difference methods are becoming fairly successful in the study of a number of equilibrium situations. Binding energies and relative protein stability towards site

Molecular Engineering **4**: 375–414, 1995.

directed mutations are two interesting examples. An overview of the methodology and a few applications are presented in a companion paper (Part 3).

2. Protease Modeling

The need for a good starting model is probably never more evident than in the case of comparative protein model building based on one or more homologous structures. Such techniques are founded on the observation that proteins with similar sequences adopt similar three-dimensional structures and that the latter is more conserved throughout evolution than the former [4]. The majority of the methodologies currently employed are largely knowledge-based, the model being an extrapolation from its homologous counterpart [5]. Structure-dependent multiple sequence alignment, databases of loop conformations, side-chain rotamer libraries, torsional drivers, energy minimization and molecular dynamics simulations are among the battery of techniques most commonly applied to problems of this nature.

Perhaps the first example of such a study (although not described as such at the time) was the application of the recently determined structure of bovine chymotrypsin in explaining the primary substrate specificity of the homologous enzyme trypsin [6]. On comparing the sequences of the two enzymes in the regions which compose the S1 specificity pocket, it was observed that at position 189 (chymotrypsinogen numbering) a serine was substituted by aspartic acid in trypsin. This led to the suggestion that the known specificity of trypsin for hydrolysis on the C-terminal side of basic residues was due to the formation of an ion-pair in the enzyme–substrate complex between the aspartic acid and the P1 residue of the substrate (nomenclature according to [7]). The subsequent determination of the crystal structure of trypsin and its complexes with several inhibitors has borne out this prediction [8–10]. Similar sucesses have been achieved in the modeling of elastase [6] and more recently in the prediction of the substrate specificities of two enzymes from cytotoxic T lymphocytes which have been subsequently tested and verified experimentally [11, 12]. On the other hand, it has proved much more difficult to correctly predict more general aspects of protein structure using the same techniques. Models for α-lytic protease [13], *Streptomyces griseus* Trypsin [13, 14] and Rat Mast Cell Protease II [13], all demonstrated major errors, particularly in the prediction of loop conformations. In several of the most serious cases the origin of the error was clearly a misalignment of the two sequences [15].

It is perhaps instructive to ask why molecular modeling has proved at least partially sucessful in the prediction and explanation of substrate specificities in the serine proteases. The answer lies in what turns out to be the relative simplicity of the problem. For the purpose of examining the process of comparative model building it is useful to divide the model into three broad regions: (1) those regions of the two proteins which are well conserved in both sequence and structure; (2) those regions which are structurally and sequencially variable; and (3) those regions which show variation in sequence but which are structurally rather well conserved. Those parts of the structure which fall into the first category can normally be constructed with a high degree of confidence but generally supply little or no new information, since the model will normally turn out to be a replica

of the starting structure. Examples of such regions in the case of the serine proteases include the members of the catalytic triad and the β-strands of the barrels which have largely structural roles. On the other hand, the second group will generally contain a large part of what is most different (and therefore most interesting) about the structure being modeled but, due to the limitations of current techniques, is normally less reliable. The third group of residues can normally be built with a reasonable degree of confidence as the problem is largely limited to the modeling of new sidechains onto a conserved backbone. This is the case of the S1 specificity pocket of the mammalian serine proteases which are largely composed of structurally conserved regions of variable sequence. In these cases the residues which line the interior of the pocket are largely responsible for the specificity.

However, while the above reasoning may work well in many cases where the percentage sequence identity between the two molecules is reasonably high ($\geq$40%) it is clearly an oversimplification. The structures of several bacterial serine-proteases homologous to the trypsin family have been solved by X-ray crystallography and despite the low sequence identity (approximately 18%) with mammalian enzymes, the structures are remarkably similar [16]. However, one of the most notable structural rearrangements involves residues around position 189, part of the S1 pocket. As described above, in trypsin this position is occupied by the aspartic acid responsible for its primary specificity. In the bacterial enzymes, however, this residue is a 1/2-cystine which participates in a disulphide bridge and which points away from the S1 pocket! This structural rearrangement is due to a deletion in the bacterial enzymes with respect to the mammalians in the region 180 to 190. Thus, although the mammalian chymotrypsin and bacterial protease A from *Streptomyces griseus* shows similar substrate specificities, these have quite distinct structural origins, emphasizing the difficulty of modeling structures which have only low sequence similarity. Furthermore it seems to be a simplification to say that merely those residues lining the S1 pocket contribute to substrate specificity, as recently demonstrated by site-directed mutagenesis experiments aimed at converting trypsin into a chymotrypsin-like enzyme. The results clearly show that several loops not directly involved in forming the S1 pocket, but close to it, and the extended substrate binding site are also important in determining specificity of amide substrates [17].

We are clearly a long way from being able to correctly predict and engineer substrate specificity completely. A recent attempt to convert trypsin into a P1 acid specific protease by substitution of asp-189 by lysine instead produced an enzyme specific for P1-leucine. One possible explanation for this is that the lysine sidechain, which is longer than the native aspartic acid, is directed away from the specificity pocket.

Indeed, until recently no three-dimensional structure was available for a serine-protease with P1 acidic specificity, despite knowledge of the fact that such enzymes exist [18]. The first such crystal structure, that of the glutamic acid specific protease from *S. griseus* (SGPE) was described in 1993 [19]. Prior to its determination a model had been proposed by Barbosa *et al.* [20] and it is interesting to compare the two structures. The model was based on Proteases A and B from *S. griseus* and α-lytic protease from *Lysobacter enzymogenes* with which SGPE shares 59%,

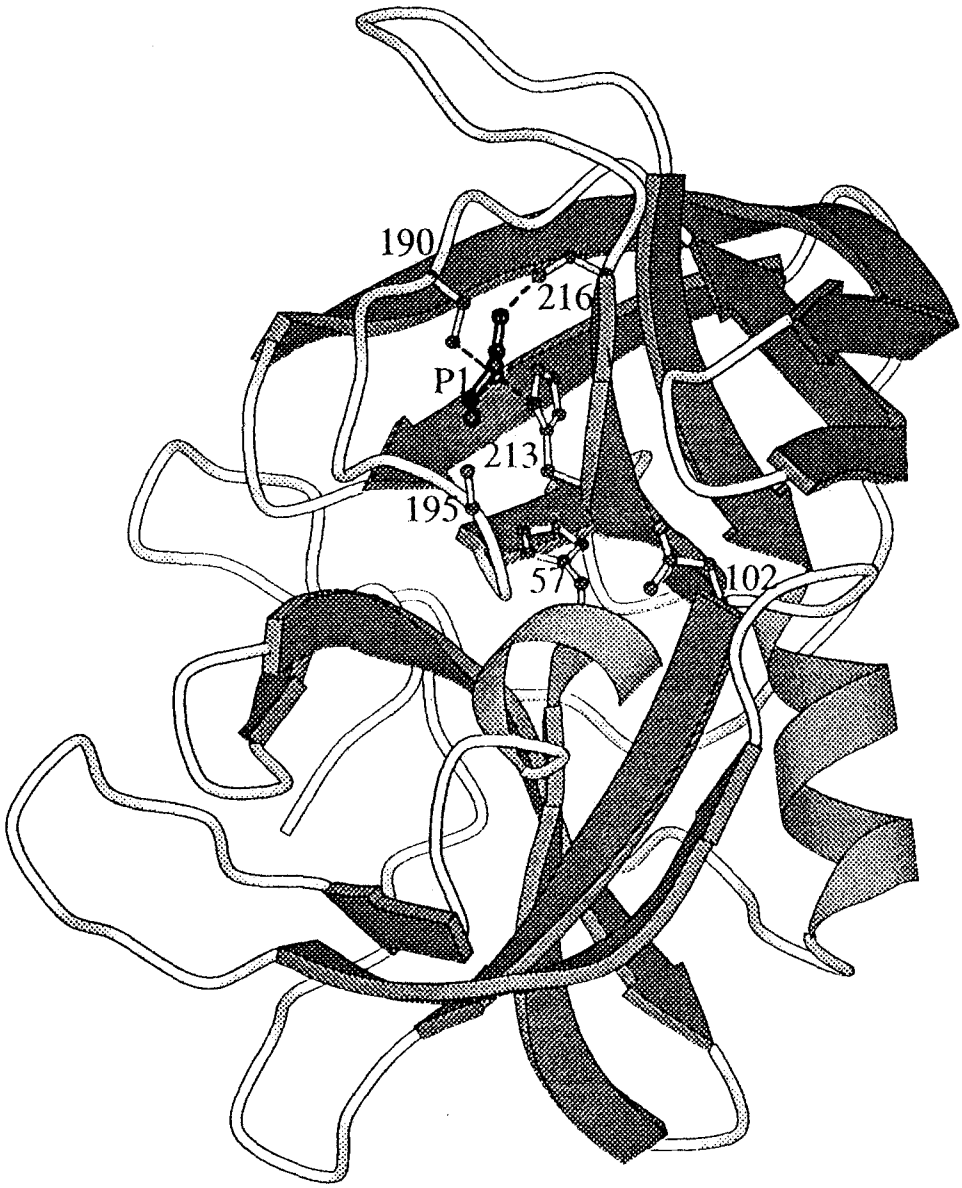

Fig. 1. Ribbon diagram of the model proposed for the glutamate specific endopeptidase from *S. griseus* [20]. The catalytic triad composed of His-57, Asp-102 and Ser-195 are shown together with those residues predicted to be involved in substrate binding, His-213, Ser-216 and Ser-190. The glutamate sidechain (P1) in the S1 pocket is shown in bold. These interactions have been largely confirmed experimentally with the exception that the histidine imidazole ring is flipped.

56% and 36% identity, respectively. It suggested that the carboxylate oxygens of the P1 glutamate would interact directly with the hydroxyls of Ser-190 and Ser-216 and with the $N_{\delta 1}$ of His-213 (Figure 1). The apparent importance of these three residues was supported by the observation that they are unique to SGPE

when compared to the three bacterial enzymes on which the model was based and which possess markedly different specificities. Of particular interest is His-213 which is conserved in other glutamate specific endopeptidases, such as V8 from *Staphylococcus aureus* [21] and GSE-BL from *Bacillus lichenformis* [22], suggesting that it plays a fundamental role in glutamate recognition. Otherwise these enzymes show only low sequence identity with SGPE.

The crystal structure of SGPE [19] was determined shortly after the publication of the model. It shows that the three residues predicted to be involved in glutamate recognition do indeed interact with the glutamate carboxylate group in a manner very similar to that proposed. The only significant difference is that $N_{\epsilon 2}$ rather than $N_{\delta 1}$ of His-213 is directly involved in the interaction. If the pK of the histidine were relatively normal, it would be neutral at the optimal pH of the enzyme ($\approx$8.3–9.0). Thus one of the problems that faced Barbosa *et al.* when constructing their model was the apparent absence of a positive charge within the S1 pocket in order to counterbalance the P1 glutamic acid. Their only suggestion was the presence of Lys-138 (which in other enzymes forms an ion-pair with Asp-194) at the base of the S1 pocket which might be at least partially involved in stabilizing the substrate via the O_γ of Ser-190. In fact, the crystal structure reveals an unexpected and novel triad of histidine residues, terminating in His-213, which runs right through the center of one of the β-barrels and is connected via structural water molecules to the N-terminus of the C-terminal α-helix (Figure 2). The authors suggest that this structure may be able to shuttle a partial positive potential through the centre of the barrel from the helix N-terminus to the S1 pocket with protons being shared by the members of the histidine triad.

Thus the model appears to have been only partially successful in its predictions, although the real importance of the histidine triad observed in the crystal structure remains to be unequivocally established. The conservation of His-213 in glutamate-specific enzymes clearly indicates its importance in glutamate recognition. However, it can equally clearly not be the whole story. Recent site-directed mutagenesis experiments on α-lytic protease in which a histidine at position 213 was introduced did not produce a glutamate specific enzyme [23]. Such an enzyme would not possess the histidine triad, of course; but there again, neither do any of the remaining glutamate-specific enzymes of known sequence. Furthermore, the recently determined structure of the picornaviral 3C cysteine protease [24] which possesses a three-dimensional structure similar to the trypsin-like serine proteases (but with the active site serine replaced by cysteine) also contains His-213 but is glutamine rather than glutamic acid specific [25]. We still have a long way to go in understanding the subtleties of substrate specificities even in the serine proteases, which represent one of the most fully understood families of enzymes.

3. Receptors

Receptors play an essential role in the regulation of several kinds of physiological processes, mainly those related to the central nervous, cardiovascular and endocrine systems [26–29]. It is commonly accepted that for a (bio)molecule to exert a desired physiological effect, it must interact with endogenous receptor molecules.

In a static view of the receptor structure, optimal interaction with it requires that the ligand must at least attain a complementary shape to the receptor binding

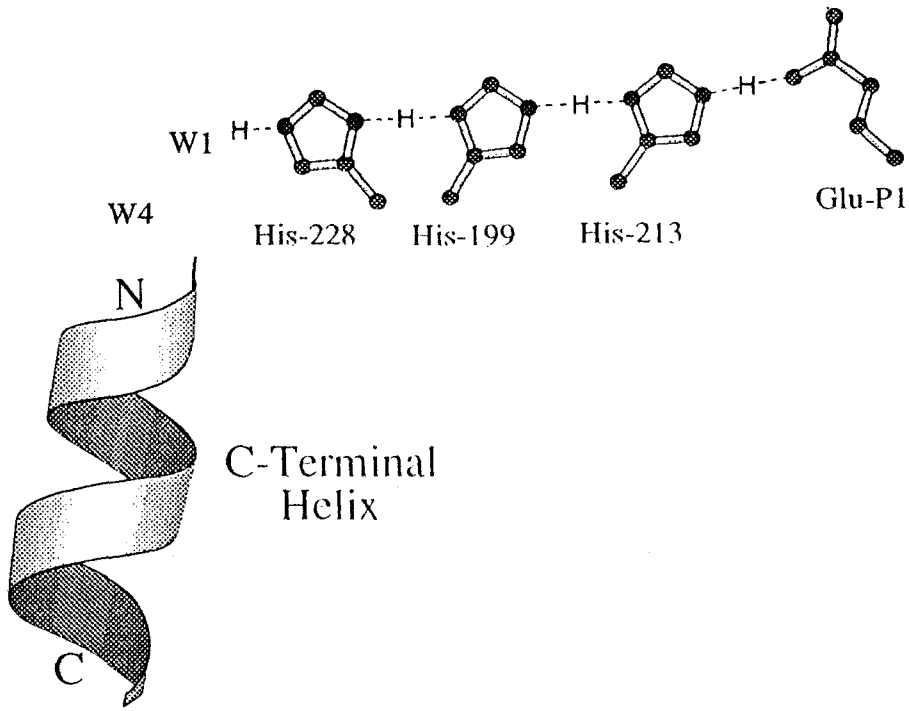

Fig. 2. The histidine triad observed in the crystal structure of the glutamate specific endopeptidase from *S. griseus* and which the authors propose to be responsible for glutamate specificity. The triad together with two water molecules would carry the partial positive charge from the N-terminus of the C-terminal helix to the S1 pocket where it stabilizes the glutamic acid at P1 of the substrate.

site. In a more dynamic view, receptor and drug should work on each other to attain the active complex conformation [26, 27]. The receptor–ligand binding process may be similar to the formation of an activated enzyme–substrate complex discussed in the first part of this paper [1]. Many factors are involved in attaining the active form: electrostatic, hydrophobic and hydrogen-bonding interactions, solvation–desolvation effect, and cooperative motion [30].

The prediction of molecular features required for a compound to induce a specific biological activity can be done on the basis of the known structure of the receptor. But this is not a simple task, however. Structural information, at the atomic level, is usually lacking on the receptor and effector molecules. Structural characterization of membrane receptors has been hampered, among other things, by experimental difficulties related to the nondegenerative purification and crystallization processes [31]. On the other hand, pharmacological, physiological and structure–activity studies demonstrate the existence of receptor subclasses that are activated by the same neurotransmitter [28].

In the absence of detailed structural knowledge about a receptor, theoretical methods have been used to: (1) assess the structural properties of the receptor required for its recognition and activation; and (2) build a receptor model or to design ligands which will hopefully bind in a selective manner to the receptor

[32, 33]. These methods exploit the available knowledge about molecules which elicit the desired biological effect, in order to characterize the structural and/or electronic factors that are recognized by the receptor. Based on the properties of interacting molecules and the results of these interactions, such methods also provide a better understanding of the key components of the activity of receptors at the molecular level. The identification of the molecular determinants of receptor activation may be related to classical pharmacological concepts such as efficacy and affinity [34].

Early theoretical work in the field of receptor modeling was mainly concerned with the study of structure–activity relationships using semiempirical quantum chemical calculations. In these studies, starting from extended sets of structurally related molecules displaying the same kind of pharmacological activity, one tries to define the molecular requirements for triggering the biological response elicited by the receptor. A receptor model is then built on the basis of a postulated electronic and conformational complementarity. Further simulations of the interactions of the lead compounds with the receptor model may be useful to characterize the nature of drug–receptor interactions and to explain pharmacological properties such as agonism, partial agonism and antagonism [35].

With the arrival of the present generation of workstations with enhanced CPU power and advanced graphics capabilities, computer-aided drug-receptor design has become a field of increasing activity. Graphical displays of molecular structures and properties are of great value in recognizing the steric factors which govern drug–receptor interactions, solvent accessible areas or structural changes elicited in animated dynamic phenomena. To this end, several computer-assisted methods have been developed to facilitate modeling studies, which in general benefit from the availability of large databases of 3D structures of biomolecules, in particular those of proteins [30]. The major pharmaceutical companies (and also many in the area of agrochemicals and materials) have made massive investments in this area, based on its potential application as a tool in deciding which molecules are the most appropriate to synthesize for specific roles [2, 36].

3.1. ION CHANNELS

The selective movement of ions across cell membranes occurs through ion channels. In general, they are pores formed in membranes through which ions are transported down electrochemical gradients. There are two types of channels; ligand-gated channels, opened by the binding of small molecules, and voltage-gated channels, opened by changes in voltage across the membrane [37, 38].

3.1.1. *Ligand-Gated Ion Channel Superfamily*

This group of molecules comprises three receptor families: the nicotinic acetylcholine (nACh), the γ-aminobutyric acid ($GABA_A$) and the glycine receptors believed to posses common topological features. Each member of this superfamily is composed of at least two distinct subunits, α and β, with similar sequences and hydrophobicity profiles. The receptors are pentameric oligomers and each oligomer is made up of four α-helices (M1–M4) spanning the membrane, as was

predicted for this superfamily of receptors in [37]. For the nACh receptor the topology is described as a five-subunit $\alpha_2\beta\gamma\delta$ assembly [38], each one formed by the four helices spanning the membrane. Glycine receptor is a pentameric oligomer composed of two subunits, the ligand binding α-subunit and the structural β-subunit [39].

The nACh receptor (nAChR) is opened by the neurotransmitter acetylcholine; it sets up a selective channel for Na^+ and K^+ ions. There are two types of nAChR. The peripheral receptors, found in the electric organs of several electric fish, and the vertebrate neuromuscular junction and the neuronal receptors, whose structures are less well understood. The peripheral receptors have been purified, reconstituted and cloned and the recombinant receptor has been functionally expressed [38]. The available data on the structural organization of peripheral nAChR from several species indicate that they are composed of four different polypeptide chains assembled into a pentameric transmembrane structure $\alpha_2\beta\gamma\delta$. The cholinergic binding site is in the extracellular domain of the α-subunit. Experimental evidence strongly suggest that the M2 helices contribute to the channel inner wall. Furthermore, labelling experiments have identified homologous serines in the five M2 helices and suggest that these residues face the center of the inner pore.

Some activity on the model building side was apparent in the 'eighties. An early model for the ACh receptor [40] was built using a 'holistic' approach, by taking into account all possible information available at that time. This model has been tested and been shown to be useful by measures of evolutionary persistence of exobilayer and ion channel active groups, the effects of genetically engineered deletions and point mutations, the response to agonists and antagonists, the behavior of reduced receptors and the possibility of producing rapid current fluctuations.

Two different α-cobratoxin-acetylcholine receptor (CTX/AChR) complex models were built on the basis of affinity labeling, toxin binding and antibody recognition data. The model analysis suggests that three different regions of the AChR α-subunit are involved in the acetylcholine binding site. The first regions contains Cys-128 and Cys-142 the second Cys-192 and Cys-193 (which are supposed to be involved in the formation of disulfide bridges at the binding site) and, the third, residues 185–196. The first model consisted of a β-pleated sheet spanning residues 136–142 and having a type II'β bend at Asp-138 and Gln-139 [41], while the second consisted of a dodecapeptide, corresponding to residues 185–196 of the α-subunit of *Torpedo* AChR [42]. This latter peptide, produced synthetically, was shown to contain the essential elements for the binding of ACh and neurotoxins to AChR as well as an unusual disulfide bond formed by adjacent Cys residues (192–193). These two models were refined by molecular mechanics and model building programs and compared in order to gain insight into the AChR binding site [42].

For the two models considered, the picture emerging was quite different both in regard to the shape of the model and to the nature of the ACh–CTX interactions. Model 1 had the shape of a compact protrusion interacting with CTX perpendicular to the direction of its long axis. Three residues at the tip of the AChR protrusion were found at interacting distances from four CTX residues. Model 2, instead, had the shape of a hook, with a long tail interacting with CTX

parallel to its long axis. In the latter case the AChR–CTX interactions were spread over the different parts of the molecules, covering regions of CTX which were not involved in the first model and without the extensive H-bonding and network of ionic interactions. This study also provided a third possibility for the binding region of ACh on its receptor, which consisted of the interior of the β-sheet.

A theoretical structural model of the channel inner wall of the nAChR has been proposed [43]. This model allowed for the calculation of the minimal interaxial distance among adjacent M2 helices and of a tilt angle of about 7 degrees with respect to the central axis of the helices. This relatively small tilt was proposed to generate only a small gap in the upper part of the pore wall, which can be easily closed by contact with a single other helix. Calculations of energy profiles and molecular electrostatic potentials showed that negatively charged Glu residues in the five helices are important for insuring the exit of the cation. The aptitude of the M1 helices to satisfy the contacts necessary with the M2 helices at the different heights of the model was also discussed in this paper. The authors concluded that the proposed model is compatible with known experimental features and allows for a rationalization of the role of the charged residues situated at the N- and C-termini of the M2 helices. In a more recent paper [44], energy profiles were calculated for a sodium atom placed in the channel of the model described above; four different mutants were considered. The relative energy location of the calculated profiles specified the role of each of the charged Glu residues placed in the anionic bottom of the M2 channel. In spite of the identity of the global charge of both channels in each pair, the structural analysis of the results gave a basis for understanding the differences observed in the conductance for the wild-type channel and the mutants. The striking correlation observed between the average relative energy location of the profiles and the conductance data appears to confirm the essential structural features of the theoretical model of the ion channel [45].

γ-Aminobutyric acid A ($GABA_A$) and glycine receptors are both chloride-ion channels opened (gated) by the inhibitory neurotransmitters GABA and glycine, respectively. Molecular cloning has revealed an extensive family of both $GABA_A$ and glycine receptor genes [38]. Both are receptors of the ligand-gated type, sharing many of the structural features of nAChR, in spite of being more similar to each other than to the nAChR [37]. A careful comparison of the amino acid sequences for both subunits of the $GABA_A$ receptor and one subunit of the glycine receptor identified four hydrophobic transmembrane α-helices and an exobilayer 'disulfide loop'.

Betz and coworkers [46] have mutated the $\alpha 1$ ligand-binding unit of the glycine receptor subunit around the amino acid position 160. This residue is an important determinant of agonist and antagonist binding to the receptor. Alignment of partial GlyR, $GABA_AR$ and nAChR subunit sequences show that the corresponding triads are FGY in GlyR $\alpha 1$, YAY in $GABA_AR$ $\alpha 1$, YGY in $GABA_AR$ $\beta 1$ and WTY in nAChR $\beta 1$, respectively (F = Phe, G = Gly, Y = Tyr, A = Ala.) The doubly mutated (F159Y, Y161F) $\alpha 1$ subunit has an affinity for β-alanine that is more potent than glycine. More interestingly, the mutated receptor was activated by GABA and D-serine amino acids while the wild type is not activated by these substances. For the singly mutant (F159Y) in GlyR $\alpha 1$ subunit, the exchange of a single hydroxyl group is sufficient to convert this receptor into a GABA-responsive

protein [46]. This finding is reinforced by the work of Amin and Weiss [47], who have identified two separate and homologous domains of the β-subunit, each of which contributes a tyrosine (Y157) and threonine essential for activation by GABA [48].

A partial model for the $GABA_A$ receptor [40] was derived from the nAChR model [49]; however, this model has been superseeded. More recent models for the transmembrane portions of some ligand- and voltage-gated channels were built assuming that the transmembrane region comprises primarily α-helices. In one such study [50], models of several peptides forming membrane channels were developed by using energy calculations with computer graphics and the Delphi and Amphi programs. Some of the questions addressed aimed at studying how these monomers aggregate on one surface of a membrane, how they insert across the membrane and how they form channels. The methods used to develop peptide channels were extended to the transmembrane region of the GABA receptor, which was built assuming that: (i) each subunit has four transmembrane helices (M1–M4); (ii) the receptor comprises two α and two β-subunits; (iii) polar faces of M1 and/or M2 line the pore; (iv) each tetramer has two channels; and, finally, (v) the tetramers aggregate. The glycine receptor was also modeled in this study.

3.1.2. *Voltage-Gated Ion Channels*

These species are sensitive to changes in voltage across the cell membrane. They can be broadly divided according to their ion selectivity into three main classes: Na^+, K^+ and Ca^{2+} channels. Several common structural features of voltage-gated ion channels were characterized by molecular cloning. Na^+ and Ca^{2+} channels consist of one large polypeptide containing four homologous repeats of six transmembrane α-helices, S1–S6. These helices are thought to form the ion channel. Helix S4, which contains a positively charged amino acid at each third position, acts as the voltage sensor. All the K^+ channels subunits differ from that general structure in being smaller and containing only one of the S1–S6 repeat domains. By analogy with the Na^+ and Ca^{2+} channels, four subunits are supposed to form the K^+ selective channel [38].

The selectivity of ion channels is an intriguing, open question. Some aspects of this problem have already been outlined above, in the discussion of ion carrier systems.

3.2. G-BINDING PROTEIN-COUPLED RECEPTORS

The members of the superfamily formed by the receptors coupled to the guanine-nucleotide-binding protein (G-binding protein-coupled receptors) can be identified by their characteristic seven hydrophobic membrane-spanning segments [31]. In spite of the diversity of functions they may exert, it is likely that these receptors have evolved from a common ancestral gene, thereby sharing basic structural and functional features [51]. They are typically present on cell surfaces, where they interact with extracellular ligands and undergo a conformational change that allows for the activation of a specific G protein.

G-protein coupling mediates a cascade of events produced after light absorption

by the visual pigment rhodopsin in the outer segment rod cell of the vertebrate retina that leads to membrane hyperpolarization and nerve excitation [51]. Henderson and coworkers have reported a projection structure of rhodopsin showing the configuration of the helices [52]. Structural elements mediating activation of the heterotrimeric G protein transducin were identified. Interestingly, the orientations of the helix axes do not coincide with those found for bacteriorhodopsin. This result may have an effect on model structures of G-protein-coupled receptors that have used BR (bacteriorhodopsin) as their template. The structural saga does not stop here. In 1993, Sigler and coworkers reported the 2.2 Å crystal structure of transducin-α complexed (Gα) with GTPγS [53]. The Gα proteins switch between inactive and active conformational states bound to GDP and GTP, respectively [54]. Receptors activate Gα by binding to the α·GDP·$\beta\gamma$ hetero trimer and catalysing release of GDP from the guanine-nucleotide-binding site of Gα. Thereafter, GTP binds, triggering dissociation of the $\beta\gamma$ complex and freeing α·GDP to stimulate the effector. Hydrolysis of the bound GTP *in situ* leads to an inactive α·GTP, thereby returning it to the α·GDP state. Rhodopsin binding and nucleotide exchange have been discussed. Exchange does not occur for transducin in the absence of activated rhodopsin [53]. There is a large helical domain that forms one wall of the nucleotide-binding cleft, thereby clogging the nucleotide. Noel *et al.* [53] suggest that the activated rhodopsin probably opens the nucleotide cleft.

Light acts as a peculiar substrate, not only due to its intrinsic quantum mechanical nature but because it produces 11-*cis*-retinal which is attached to the protein by a Schiff-base linkage into an electronically excited state. It is this latter which will be seen as the actual substrate which, via a conformational change, activates the cascade described above.

The G-protein coupled-receptor family includes receptors for neurotransmitters and peptides. Over 80 receptor sequences from a variety of species have been obtained by molecular-cloning techniques. Some of the receptors belonging to the superfamily are the adrenergic, octopamine, dopamine, serotonin, histamine, cannabioid, tachykinin, chemo-attractants, bombesin/GRP, neurotensin, angiotensin (mas), GH, opsin and other receptors. To this superfamily belongs the muscarinic acetylcholine receptor. Some of the model building and quantum chemical work is now reviewed.

3.2.1. 5-*HT Receptors*

The receptors for the neurotransmitter serotonin (5-hydroxytryptamine or 5-HT) comprise a plethora of receptor subtypes [55], some of which have had their sequences determined [31]. However, the details of their structure at the atomic level are not yet known. Current experimental evidence shows that the S-HT_{1A} and the 5-HT_2 receptors are different protein molecules and the effector mechanisms associated with each receptor are also different. While the 5-HT_{1A} receptor modulates adenylyl cyclase activity in the brain, the 5-HT_2 receptor stimulates phosphoinositide turnover [34]. It seems reasonable to assume that different recognition mechanisms might be involved in these receptors.

Weinstein and coworkers have carefully formulated models for identifying the structural and mechanistic basis of function, classification and ligand design for

5-HT receptors. In their early simulations, a heuristic approach was taken to study the molecular determinants for 5-HT_{1A} receptor recognition and activation. The electrostatic properties of 5-HT ligands, such as those expressed in the molecular electrostatic potentials (MEP) they generate, were identified as the major determinants of ligand selectivity at these receptors [56]. In relation to the positively charged end of the side chain of 5-HT_{1A} cognate, the directionality and shape of the MEP above the indole portion of 5-HT_{1A} were shown to predict the relative affinities of these compounds, as well as other classes of compounds including fused ring structures derived from tryptamine and ergoline derivatives [57]. The specific requirements for the spatial orientation of ligands on the receptor led to the proposal of a three-dimensional template, defined from the properties of the more rigid ligands [56].

A stacking model for the interaction between imidazolium cation (the heuristic receptor model) and a large number of 5-HT_{1A} congeners was proposed to study the interaction of 5-HT_{1A} receptors and compatible reactants [56]. The imidazolium cation was chosen to test the directionality hypothesis because of its known interaction with indole. Its charge distribution also exhibited the implied complementary property required for interaction with the MEP of the indole portion of 5-HT_{1A}. The quantum chemical simulations of this complex formation validated the modeling criteria based on MEP, by showing a clear energetic preference for specific mutual orientations of the stacking complexes. Furthermore, they also allowed insights into the consequences of ligand–receptor complex formation [57]. The interaction of the imidazolium cation with 5-HT_{1A} produces changes in its electron density distribution, which led to the proposal that it could donate a proton to a potential acceptor [58]. The interaction of a ligand, such as 5-HT_{1A}, with a proton transfer model (PTM) composed of the imidazolium cation interacting with a proton receptor site, was then simulated by quantum chemical calculation. For simplicity, an amino group was used as the proton receptor in the early calculations. These studies showed that in the absence of 5-HT_{1A}, the process of charge transfer from imidazolium to the receptor presents a high barrier and a minimal energetic driving force (i.e., the energy difference between reactants and products). The interaction set up by the approach of 5-HT_{1A} to the PTM lowers the barrier and generates an appreciable driving force for the proton transfer. The electrostatic field generated by the ligand determines a charge redistribution in the proton donor that reduced its proton affinity, allowing the transfer to start. This was proposed as a plausible molecular mechanism for the activation of 5-HT_{1A} receptors [58]. This model permitted discrimination between agonists (receptor activators) and antagonists (receptor blockers) by identifying those ligands able to produce the charge redistribution responsible for the positive energetic drive of the reaction.

Since this activation model for the 5-HT_{1A} receptor depends on the electrostatic field generated by the ligands and specific steric factors governing the mutual orientation of the ligands at the recognition site, the receptor environments may affect the activation. The inclusion of this effect required a simulation of the activation process inside the protein environment. Actinidin (a thiol protease) was selected as a model for the receptor, because it contains an appropriate juxtaposition of the groups that form the proton relay complex assumed in the above-

mentioned activation mechanism [59]. The structure of actinidine has been well characterized by X-ray crystallography at 1.7 Å resolution [60]. In the model, one of the proton transfer systems of actinidin consists of a His-62 hydrogen bonded to Cys-25 and another is made up of Asn-182 hydrogen bonded to the Nϵ2 of His-162. The indole group in the side chain of Trp-184, which could mimic ligands that are 5-HT derivatives, is in close proximity to this system and is aligned in a stacking geometry with respect to the imidazole ring of His-162.

The potential energy curve for proton transfer between Nδ1 of His-162 and Sγ of Cys-25, in which His was modeled by imidazolium and Cys-25 by methanethiol, was obtained by *ab initio* calculations employing a variety of basis sets, ranging from a minimal STO-3G to an extended 6-21G** basis set [59]. The qualitative trend was found to be independent of the basis set. The energy and charge distribution of the quantum motif (protein fragments represented quantum mechanically) in the classical electrostatic field generated by the protein surroundings was also calculated using either the Direct Reaction Field approach [61] or with a perturbation theory scheme [62]. Results from such calculations [34] showed that the six α-helices in actinidin contribute differently to the energy of proton transfer between His-162 and Cys-25. These results indicated the essential role that the helical motifs, predicted to exist in the structures of membrane-bound receptors, could play in receptor mechanisms. Proton transfer from His-162 to Asn-182 was also studied and proved to present greater sensitivity to the β-sheet structures in the protein [63]. The two proton relay model systems studied are not strongly affected by near neighbors, but the effect of the entire protein environment opposes proton transfer in both cases. Thus, the protein can prevent a spontaneous triggering of the receptor activation mechanism.

The calculations of the His-162/Asn-182 complex also showed that the ligands (simulated by the side chain of Trp-184 or replaced by other residues, with side chains simulating compounds with known activities at the 5-HT_{1A} receptor) do not equally affect the energy barrier for proton transfer. Within each structural group studied, the effect of the ligands in lowering the energy barrier follows an order compatible with the pharmacological data [34] and is determined by their effect on the proton donor-acceptor pair and not only in the moving proton. From this model, it follows that the receptor occupancy by a partial agonist does not produce a response since the energy barrier is still too high.

In order to explore the differences between the recognition processes at the 5-HT_{1A} and 5-HT_2 receptors, the determinants for recognition at the 5-HT_2 receptors were compared to those described above at the 5-HT_{1A} receptor [64]. To this end, the interactions of 5-HT and 6-HT with an indole molecule were explored. The indole molecule, which represents the side chain of a tryptophan residue, was taken as a model for the ligand-recognition site in the 5-HT_2 receptor. The interactions between 5-HT_2 selective ligands and the 5-HT_2 receptor were modeled by complexes between this ligand and tripeptides that are involved in the binding of some proteins to 5-HT [64]. The essential tripeptide seems to be composed of two aromatic residues (Phe and Trp, or Trp and Tyr) separated by a hydrophilic residue (Ser or Arg). The interaction potential (including the electrostatic and dispersion energies) between indole and 5-HT or 6-HT was scanned over an entire surface. The results obtained for the scan with 5-HT are quite different from those

with 6-HT. The scanning position that most clearly discriminates between 5-HT and 6-HT was selected for more detailed investigations of the variation in scan energies with the orientation angle. Here, the magnitude of the dispersion energy is generally greater than that of the electrostatic energy. It was also shown that the variations of these components of the energy are almost exactly out-of-phase. The electrostatic interaction was assumed to determine the preferred orientation and to establish a correlation with the selectivity, while the dispersion energy is related to the stability of the complex (e.g. to the ligand affinity). These results are in contradiction to those obtained previously for the model recognition site for the 5-HT_{1A} receptor, for which electrostatic factors governed both the preferred orientation and stabilization.

Since information on the protein sequences of 5-HT and other kinds of receptors are becoming available, it is now possible to combine them to build receptor models by using strategies which differ from the more heuristic one. For instance, it is known that for most G-protein coupled receptors, with the possible exception of the receptors for the glycoprotein hormones, the ligand is held in a hydrophobic pocket formed by the seven transmembrane helices. The comparison of results for different types of receptors shows that all helices but helix I are directly involved in the receptor–ligand interaction [51].

It was suggested that the G-binding proteins, in spite of a low sequence homology, share the topology of the seven transmembrane helices of bacteriorhodopsin. Increased sequence similarity was found by allowing for helix shuffling corresponding to minimal exon shuffling during evolution. A model for the 5-HT_{1A} receptor was then built [65] starting from the 3.5 Å resolution structure of BR, by making hypotheses concerning helix–helix interactions, their orientations and arrangements in bundles surrounded by lipid molecules.

Zhang and Weinstein have constructed a molecular model of 5-HT_2 receptor and carried out a series of molecular dynamics simulations of the receptor complexed to ligands [66]. BR was used as template. Although the MD trajectory appears to be non-equilibrated (the total energy as a function of time is still drifting) interesting results were obtained for the dynamics of the seven transmembrane helices. A signal transduction mechanism was proposed. It was also remarked by these authors [67] that the polarity of residues at certain positions in the transmembrane domain of this G-protein coupled receptor appears to be conserved. They seem to indicate the pattern of specific helix–helix packing of the helices. The reader is referred to the original work for further interesting details and for a description of the concept of polarity conserved positions [66, 67].

3.2.2. *Histamine H_2-Receptor*

Histamine [2-(4-imidazolyl)ethylamine] is a neurotransmitter known to act on three different receptors, namely H_1, H_2 and H_3, eliciting different biological responses in each case. These receptors have been characterized by using selective agonists and specific antagonists.

A large number of structure–activity relationships and mechanisms of action have been proposed for histamine, on the basis of its chemical and molecular properties, as well as those of its agonists and antagonists.

Kier's [68] early work, using extended Hückel type conformational calculations, attributed the H_1 and H_2 activities to two conformers (*trans* and *gauche*) of histamine; the former probably being the active form for the H_1 receptor agonist. The relative population of both conformers in solution (*trans*, 45% and *gauche*, 55%) was determined for methylhistamine from NMR spectroscopic measurements and EHT calculations [69]. The population ratio of these conformers was compared with their H_1/H_2 histamine receptor agonist activity ratio and they appeared to be unrelated [70].

From a conformational study of methylhistamines, the outstanding H_2-receptor selectivity toward 4-methylhistamine was attributed to its different conformational properties compared to those of histamine. The *trans* fully extended form of the side chain was proposed to be essential for productive interaction with the H_1 receptor [69–71]. Regions of maximum probabilities for the essential H_1 conformation were proposed [72].

Chemical differences in the receptor requirements were also examined; these endow histamine H_1- or H_2-receptors with agonist selectivity. The histamine monocation is likely to be the active form in both cases. The H_2-receptor agonists with a reasonable level of activity were assumed to be compounds able to undergo 1,3-prototropic tautomerism. Electron-withdrawing substituents in the 4 position of the imidazole ring change the relative tautomer concentrations, alter the electron density at the N atom and affect proton acidities. The active form of histamine for both receptors was proposed to be the N3–H tautomer [73].

The charge distribution of histamine monocation in its 'essential' conformation was obtained by integrating the square of the molecular *ab initio* wave function at the STO-3G basis set level. The results showed an evenly dispersed positive charge over the molecular skeleton and that pictures of receptors involving localized negative sites may be invalid. A detailed description of the electron distribution was given by Richards and Wallis [74].

The existence of the *trans* and *gauche* conformers for histamine H_2-receptor agonists was also obtained from CNDO/2 potential energy calculations; moreover, the *gauche* conformer was found to contain an internal hydrogen bond [75–77]. From these studies it was suggested that solvent effects could open this hydrogen bond and stabilize the *trans* conformation in such a way that an open conformation, nearly *trans*, was proposed to be the 'H_2-essential' conformation. A correlation between the mean charge of the imidazole ring of histamine derivatives, substituted in position 4, and the H_2 receptor agonist activity was also verified.

Structure–activity relationships for histamine H_2-receptor agonists and antagonists, belonging to two different structural series, were used to derive a proposed topography for the H_2-receptor from structural AM1 calculations [78, 79].

The acid–base properties of the imidazole nitrogen atoms and tautomeric preference for the monocation and neutral histamine were studied by using *ab initio* MO calculations at a minimal basis set level [80, 81]. A dynamic mechanism for H_2 activity and the H2 receptor was proposed, in which histamine appears in a proton relay system, similar to that found for the serine proteases. The model assumes that the receptor has three different active sites; histamine approaches the receptor as the N3–H tautomer of the monocationic (active) form. The cationic side chain interacts with a negative region of the receptor (site I) and the resulting

charge neutralization causes a shift in the tautomeric preference of the imidazole ring to N1–H. As a consequence of the attachment of histamine to the receptor, N1 attracts a proton from the proton-donor site (III) of the receptor, while the proton on N3 is transferred to the proton acceptor site (II). Theoretical calculations based on this mechanism show that it provides an explanation for the pharmacological activity of some histamine H_2-receptor agonists (2- and 4-methylhistamine [82] and triazole and tetrazole analogs of histamine [83]). The activation mechanism of this proposed H_2-receptor model was explored by studying the proton-relay mechanism. The negative binding site of the receptor (site I) was modeled with a hydroxide ion, ammonia represented the proton-acceptor site and an ammonium group was used to simulate the proton-donor ion. A hydroxide ion was alternatively used to represent site II [84].

Ab initio quantum chemical calculations at the 3-21G* basis set level were used by Weinstein and coworkers [85] to explore the interaction of dimaprit, (*S*-(3-(*N*,*N*-dimethyl amino)propyl)isothiourea) – a potent and specific H2-agonist – with the above molecular complex designed to model the specific recognition of histamine [84]. The H_2-receptor activation mechanism triggered by the binding of this ligand was also examined. The assumption was that the activation of this receptor by dimaprit followed the same mechanism proposed for histamine, in spite of the differences in their molecular structures. The methylisothiourea molecule was used as a model for the ligand. The stabilization of several isomeric forms of the isothiourea moiety of dimaprit in the receptor model was considered, including models for the monocationic and dicationic forms that are likely to occur under physiological conditions. The monocationic form was found to be the most likely physiologically active species. The energetics of the proton transfer from a receptor site to the ligand were evaluated in the presence and absence of other model sites in the receptor. The energetic contribution of ligand desolvation to the various steps of receptor binding and the activation mechanism was estimated from calculations of the enthalpy of solvation in water, here represented as a continuum dielectric. The results suggested that the proton transfer from the model receptor site required the sulfur in the isothiourea moiety of dimaprit to act as the proton acceptor in the activation mechanism. This proton transfer also depends on the receptor environment.

The chemical structures of dimaprit and NAGHA [2-(4-imidazolyl) ethylguanidine], a partial agonist, were used to improve this histamine H_2-receptor model [84], through a more realistic characterization of the receptor binding sites [35]. In this improved model, a formate anion, which is the smallest group that can simulate an aspartate or glutamate residue, was proposed both as the negative site at which the histamine cation is anchored at the receptor, and as the proton-acceptor site. An ammonium cation representing a lysine, arginine or histidine residue was used to simulate the proton-donor site. In this work, *ab initio* quantum chemical calculations at the STO-3G, 4-31G, 3-21G* and a 'mixed basis set' level were performed to characterize the interaction sites and the activation mechanism of the histamine H_2-receptor inside the receptor model. Geometry optimizations were carried out leading to stationary points on the potential energy surface; both minima and transition structures were evaluated. Semiempirical quantum chemical calculations using the AM1 Hamiltonian were used to optimize histamine geome-

trically inside the model. The molecular determinants for the activation process at the histamine H_2-receptor were assumed to be proton transfer from sites III (proton-donor site) to II (proton-acceptor site). This movement, mediated by the imidazole ring of the histamine, was found to be sequential in the potential energy surface. From an energetic point of view, both proton transfer processes were found to be feasible. However, the proton transfer from N3 to site II was the rate-limiting step in the recognition process, since it had the higher activation energy. The definition of the receptor sites essential for the biological activity can be used to explore the connections between the structure and the activity of compounds which act as agonists, partial agonists and antagonists. In this sense, this model also provided a basis to explain the molecular determinants of the biological activity of Na-guanylhistamine. The structural properties of the guanidinium group allow this compound to bind the H_2-receptor in two different modes. Only one of them was shown to prompt the proton transfer process and was, therefore, characterized as the binding mode displaying properties of an agonist; the other binding mode would elicit antagonist behavior. The properties of NAGHA as a weak antagonist and a partial agonist were examined by calculating the energies of ligand desolvation. It was concluded that the partial agonism of NAGHA is related to the ability of the molecule to interact with the histamine H_2-receptor in agonist and antagonist modes of similar affinity. Finally, the residues that form the active site of this receptor were inferred from the knowledge of the amino acid sequence of the gene that encodes the histamine H_2-receptor.

Two possible modes of interaction of dimaprit were proposed and studied by computer graphics, semi-empirical (MNDO) and non-empirical ($X\alpha$) quantum-chemical calculations [86]. In one case, of the two N atoms of the thioureum moiety of dimaprit one acts as proton acceptor in the proton transfer system and in the other case, the sulfur atom acts as the proton acceptor. To test these models, the optimized geometry of dimaprit was fitted to histamine in its fully extended conformation. Two types of fits were considered, so as to represent both interaction modes (S-fit and N-fit). Once the two conformations of dimaprit in the S- and N-fit were constructed, calculations were performed to obtain their energies and molecular electrostatic potentials. From a geometric point of view, both structural fits seem reasonable. However, from considerations based on the electrostatic potentials, the S-fit was favored. The sulfur atom in dimaprit presumably acts as the proton acceptor in the proton relay system needed to activate the histamine H_2-receptor.

The sites of the histamine H_2-receptor model already defined [35] were further employed to elucidate the pharmacological mechanism of action of some potent and selective histamine antagonists: cimetidine, tiotidine and ranitidine [87]. As a working hypothesis, it was assumed that these antagonists occupy the same receptor sites as agonists, thereby preventing the latter from occupying the active site and triggering the proton-relay mechanism which, in turn, prompts the pharmacological response. As in the previous work, quantum chemical calculations at the 4-31G and a 'mixed basis set' level were used to obtain the potential energy curve for proton transfer from site III to the antagonists; structural optimization of the drug-receptor complexes were performed with the AM1 Hamiltonian. The contribution of solvation energy to the stabilization of the various complexes was

studied as a discrimination factor to distinguish between the various ionic and tautomeric forms of the antagonist molecules in solution and the two possible modes of binding with the receptor. From the analyses of drug–receptor complex formation enthalpies, the monocationic form of the antagonists was found to be the most favorable species to interact with the receptor. The recognition pattern followed the same trends already observed for agonists: the protonated site of the molecule (imidazolium ion in cimetidine, guanidinium in tiotidine or substituted ammonium in ranitidine) interacted with the negative site of the receptor, while the non-basic part (cyanoguanidine in cimetidine and tiotidine and nitrodiaminoethene in ranitidine) was located between proton donor and acceptor sites. An energetic analysis of the interaction between antagonists and the receptor model, including the energies of ligand desolvation, showed that histamine cannot compete effectively with the studied antagonists for binding to the H_2-receptor. The predicted order of antagonism potencies, based on differences of enthalpy of formation (tiotidine > ranitidine > cimetidine) reproduced qualitatively the experimental rank order.

Previous models for the recognition and activation of the 5-HT and histamine H_2-receptors were used to design a compound to test the H_2-receptor, the structure of which was derived from the LSD framework [88]. The design strategy, to obtain a compound with potentially selective H_2 agonism, consisted of replacing the indole ring of 9,10-dihydro-LSD with an imidazole group. In the former compound, the recognition properties for 5-HT receptors were shown to be lost and the activation properties for 5-HT and H_2 receptors were absent [57]. The existence of the H_2 recognition and activation features in the proposed compound was verified by the computational studies of MEP (molecular electrostatic potential) and shifts in the tautomeric preference. The MEPs and the energies of the optimized derived structures were calculated at the *ab initio* Hartree–Fock level with the STO-3G basis set [89], while the optimization of the neutral and cationic forms of the N1 and N3 tautomers of the model compound was done at the MNDO level. The qualitative results showed that, regardless of the conformation, or group of conformations used, a shift in tautomeric preference was observed for the model compound which parallels that of histamine. The use of different conformational structures only affects the magnitude of the shift. Thus, the designed compound could act in the same way as histamine does in the activation process of the H_2-receptor.

An alternative model was proposed to explain both agonist and antagonist activities of histamine at the H_2-receptor by Kimura *et al.* [90]. This model is based on experimental evidence showing that histamine and its H_2-agonist, dimaprit, may interact with a triprotonated macrocyclic hexamine, (18)-ane-N_6, at physiological pH, to yield 1:1 complexes with simultaneous liberation of H^+. This interaction was assumed to mimic that found in histamine H_2-receptor-agonist and the resulting gastric acid secretion. This receptor model showed no affinity towards a histamine H_1-agonist. Furthermore, it was found to interact with some H_2-antagonists to form stable 1:1 complexes which competitively blocked H_2-receptors. In this latter case, since the side-chain of the H_2-antagonists are unprotonated at neutral pH, no acid liberation occurred upon complexation with the macrocyclic

cation. This model can thus account for known structural features distinguishing histamine H_1- and H_2-agonists, as well as histamine H_2-agonist and antagonists.

The theoretical work so far described focuses attention on the characterization of *equilibrium* geometry structures. Following a quite different approach, the potential energy hypersurface of the N3–H histamine monocation was determined by *ab initio* methods at the STO-4G level using analytical gradient techniques [91]. The hypothesis suggested in this work was that the histamine receptor is better suited to bind conformers having a geometry ressembling transition structures (TS) rather than equilibrium geometry. To trigger a biological response, the TS must connect to a minimum deep enough to be able to perform significant work on the receptor. The deep minimum may act as a virtual attractor. The STO-4G hypersurface had such features, several stationary points on the potential energy hypersurface of the N3–H histamine monocation were located (potential energy minima and first-order saddle points). One of the deep minima, a nearly *trans* preferred conformer M2, was found, in agreement with previous results obtained with a STO-3G approximation by Topiol *et al.* [80]. In the conformer hyperspace near to this minima, three transition structures were found. One of them is energetically accessible from the commonly populated conformers of histamine; it connected the *trans* and *gauche* conformational regions. This structure had the bond order for one of the three possible N—H bonds lower than the others. This fact suggested an 'incipient' formation of an internal hydrogen bond. The deep minima corresponded to a chelated structure, where a proton was incorporated in a six membered ring. This 'scorpio' structure may be related to the chelated complex proposed in Kimura's mechanism for the activation of the histamine H_2-receptor [91, 92]. The hypersurface of histamine has now been calculated by Hernandez-Laguna and coworkers at a much higher level of theory. The results obtained previously have been globally confirmed and more details have appeared. Finally, the intrinsic basicity of this molecule has been recently established [93].

3.2.3. *Adrenergic Receptor Family*

Maloney, Huss and Lybrand have applied computer-aided model building techniques to generate a three dimensional structure for hamster β_2-adrenergic receptor protein based on computer assisted model building techniques [94]. The model building began with the assumption that the receptor protein consists of a transmembrane domain that is most likely a helix bundle. The first stage of the process was the detailed analysis of the primary sequence of the β_2-receptor in order to identify the possible membrane helical segments. Experimental data was used as constraints to guide the model building process. The model reported represents a plausible three-dimensional structure of this G protein-coupled receptor [94].

Using bacteriorhodopsin as a template, Wollmer and coworkers [95, 96] have modeled seven hydrophobic regions of β_2-AR, human β_2-AR, chemotactic receptor (C5aR), human C5a receptor. A thorough discussion of model building is given by these authors and a novel approach is presented which is based on so-called interaction energy differences. The tertiary structure of the seven helix receptor is conceived as a hypothetical folding mechanism using the same reasoning

as the one introduced by Popot [97] in his modeling of membrane proteins (see Section 4.3.1). The resulting model appears to be in fair agreement with existing data from structure–function studies [95].

3.2.4. *Dopamine D_2-Receptor*

Two central dopamine (DA) receptors (D_1 and D_2) have been characterized [98], of which the D_2 subtype is particularly involved in motor and endocrinological diseases. In an attempt to define the molecular requirements for dopaminergic activity, several topographic models were proposed [99, 100] and reviewed by Katerinopoulos and Schuster [101]. More recently, the receptors D_3 and D_4 have been discussed by several authors in connection with possible sources for schizophrenia [102–104].

A common structural model for central nervous system drugs and their receptors was proposed by Lloyd and Andrews [105]. It consists of two aromatic 'receptor' points, a nitrogen atom and a nitrogen 'receptor point'. The molecular requirements for individual receptor classes were defined by the relative location of secondary binding groups.

Similar techniques were used to generate a three dimensional receptor model of the dopamine D_2 receptor from computer graphics analyses of D_2 agonists [106]. In this work, 33 compounds with known pharmacological activity were studied: some of them displayed high potency, while others presented very low D_2 agonist activity. Four potent D_2 agonists were used to define a primary pharmacophore for the D_2 receptor. The classical potential energy was calculated to find all low-energy conformations for each molecule, by varying up to three torsion angles. Based on the previous structural model [105], hypothetical receptor points (R_1, R_2 and R_3 and a nitrogen atom) were built for each compound studied. Here, R_1 and R_2 represent hydrophobic bonding to the receptor and R_3 represents a hydrogen bond to an electronegative atom on the receptor. The four most active compounds were superimposed using these four points as guides to minimize the root mean square distance deviations. Coordinates were obtained for the primary pharmacophore. The low-energy conformers of the other compounds, belonging to eight distinct structural classes of D_2 agonists, were then superimposed on the primary pharmacophore to help locate secondary binding sites. These were identified as two lipophilic clefts, an area of steric bulk, a region to hydrogen bond meta hydroxy groups and a critical region accepting methoxy and halogen substituents, but not hydroxy substituents.

Recent work aimed at defining the 3D requirements leading to Na^+-dependent or independent D_2 receptor binding used molecular graphic design coupled with PCILO (perturbative configuration interaction with localized orbital representation [89]) and crystallographic data, to first investigate four dopamine D_2 receptor antagonists showing Na^+-dependent binding [107]. Two of them are 'rigid' Na^+-dependent antagonists, while the other two are 'flexible'. Finally, comparisons were made between these compounds and other Na^+-independent derivatives. Three putative pharmacophoric element – a nitrogen lone pair, a phenyl ring, and a carbonyl moiety – were found to be similarly oriented in all the Na^+-dependent drugs. Most probably, these particular 3D structural elements and associated

properties explain their selectivity towards the D_2 receptor. Conversely, for Na^+-independent analogs, the two latter pharmacophoric elements play a subordinate role, but two *p*-electron regions are systematically localized on the other side of the molecule. The first is a phenyl group and the second may be either a carbonyl function, a cyano group or a phenyl ring, depending on the structure of the compounds under study.

A 3D model of the dopamine D_2 receptor was constructed from the rat D_2 receptor amino acid sequence [108]. The sequences of 13 other G protein-coupled neurotransmitter receptors were aligned. The modeling was based on five assumptions: (1) membrane-spanning domains of G protein-coupled neurotransmitter receptors are α-helices; (2) each transmembrane helix contains 27 amino acids; (3) membrane-spanning segments of G protein-coupled neurotransmitter receptors have similar localizations in the aligned sequences; (4) the most polar surface areas of the transmembrane helices form a central core; and, finally, (5) G protein-coupled neurotransmitter receptors have a common ligand-binding site. The computer built model was further refined by molecular dynamics simulations and molecular mechanics calculations; it has seven α-helical transmembrane segments that form a central core with a putative ligand-binding site. Two residues, postulated to be involved in agonist binding (Asp-80 in helix 2 and Asn-390 in helix 7), were placed adjacent to each other, at a distance that perfectly accommodated the energy minimized *anti* conformation of the dopamine molecule. This result was in accordance with the hypothesis that this might be the active conformation. The calculated molecular electrostatic potential was mainly negative on the synaptic side of the receptor model and around the aspartate residues lining the central core positive on the cytoplasmatic side. It was therefore assumed that protonated ligands were attracted to the binding site by electrostatic forces and that protonated agonists may induce conformational changes in the receptor, leading to G-protein activation, by increasing the electrostatic potential near Asp-80. The docking of dopamine into a postulated binding site was examined using MD simulations. The results indicated that the protonated amino group became oriented toward negatively charged aspartate residues in helices 2 and 3, whereas the dopamine molecule fluctuated rapidly between different *anti* and *gauche* conformations during the simulation; the results support the so called 'zipper' type of mechanism for dopamine-receptor interactions [108].

Conformational analysis with the MM2 molecular mechanics method were performed by Petersson and Liljefors [109] on four different types of phenylbenzamides and a phenylpyrrole analogue which display high affinity for the dopamine D_2 receptor using a previously described D_2 receptor-interaction model [110]. Energetic consequences of small deviations from coplanarity of the phenyl ring and the amide moieties of the compounds were investigated with *ab initio* STO-3G calculations. The low-energy conformations of the compounds were least-square fitted to the previous model by using three different atoms or points in each of the molecules to perform the fitting: the center of the phenyl ring, the amino nitrogen group and a point 2.8 Å away from the amino N atoms, in the direction of the N lone-pair, that simulated an H-bond to the receptor. All four types of benzamides accommodated in this model are in low-energy conformations. Two different receptor sites for *N*-alkyl substituents were also suggested.

The influence of solvation in drug-receptor interactions was examined for a series of 3-benzazepines interacting with the dopamine D_1-receptor. An inverse linear relationship was found between the free energy of solvation and the logarithm of the affinity of the ligands. However, it was suggested that electrostatic properties might also modulate affinity [111]. Two methods designed to calculate free energies of solvation of ions and neutral molecules, AM1-SM1 [112, 113] and DelPhi were used in this study.

4. Transport Molecules

4.1. THE SUPERFAMILY OF HYDROPHOBIC LIGAND TRANSPORTERS

The fatty acid binding proteins (FABPs) belong to a superfamily of proteins that encapsulate and solubilize hydrophobic ligands. All members of this superfamily seem to use the same kind of barrel topology in their ligand binding, with the first β-strand shared by two orthogonal sheets. However some differences in their topology and other structural features give rise to two families: that of serum Retinol Binding Protein (RBP) defined by weak sequence similarity among their members and the cellular retinoid binding proteins that shows better internal sequence similarity [114].

4.1.1. *Serum Retinoid Binding Protein Family*

It was for the retinol binding protein that molecular dynamics simulations were first employed to test stability of the apo form towards thermal fluctuations [115–117]. The X-ray structure of the holo form, retinol in its complex with RBP [118], was known but the apo form resisted crystallization. The suggested reason for such a failure was that the fold of apo-RBP was unstable. The MD simulations suggested that the fold of apo-RBP was as stable as the holo-form. Recently, the apo and holo structures of a rat fatty acid binding protein, a structurally related molecule, were determined with X-ray crystallography. The experimental results showed that the β-structure remains essentially invariant in both molecules [119, 120]. The holo and apo structures of human and bovine RBP, bovine β-lactoglobulin and bovine odorant-binding protein have been compared in their sequence and 3D structure [121]. After determining the X-ray structure, these authors confirmed the 3D structure obtained by themselves from molecular model building techniques. Thus, the substrate does not appear to drastically affect the structure of the apo protein, minor changes in the loop arrangement can be detected.

The fold of this family has a simple topology. Its main feature is a β-barrel onto which stacks a four-turn α helix. Eight antiparallel strands form the barrel so that part of the first strand is flanked by the last. This results in two orthogonal β-sheets with the first strand common to both. The barrel topology results in two possible entrances/exits. One is closed off by the tight packing of five phenylalanine rings around the side chain of Met-53. The retinol molecule sits in the center of this barrel with the β-ionone ring innermost. At the entrance, which is formed by residues from three loops, the alcohol group of the ligand is slightly exposed to solvent.

For other family members, such as β-lactoglobulin and odorant-binding protein, the central core, α-helix and C-terminal loop can be aligned, while the entrance loops vary in size and conformation. Although the binding site for the retinol in the RBP structure is unambiguous, it seems that there may be a secondary binding site. Interestingly, the sequence-conserved residues among the members of the RBP family are not located on the inner surface of the barrel and thus they cannot interact with the ligand inside the barrel.

Model building techniques have been used to predict the structure of a molecular complex between RBP and its carrier protein. This followed the theoretical prediction of the stability of the apo-protein [122]. In the plasma cells, a complex is formed between holo-RBP and the thyroxine-binding transthyretin tetramer (previously known as prealbumin). This complex is then recognized by a cell surface receptor, which delivers the vitamin to the cell. Since RBP probably interacts with more than one intracellular protein when receiving the retinol in the liver hepatocytes, the molecule is involved in at least four different interaction types. The 3D model for the complex provides a molecular basis to describe the experimental results [123].

4.1.2. *Cellular Retinoid Binding Protein Family*

This family shows better internal sequence similarity than does the serum RBP family. The structure consists of a barrel with the up-and-down topology which is similar to the barrel of the RBP family. However, there are 10 antiparallel strands instead of eight. Again, the structure has orthogonal sheets where the first strand is shared between them.

A difference exists in the relative positioning of α-helices with respect to the barrel in the two families. In the cellular family, one end of the barrel is closed by a helix-turn-helix motif connecting the first two strands. A partially accessible opening to the interior of the protein is formed by residues in one α-helix and two turns connecting β-strands. This zone has been postulated to be the entry portal for fatty acids which are bound inside the barrel. In the serum family, however, it has been suggested that the α-helix of the RBP, which is situated at the opposite side of the opening, could be correlated with some intermolecular interaction at a cellular receptor [115, 117].

The cellular retinoic acid binding protein (CRABP) and the cellular carrier protein for retinol (CRBP) may act as shuttles, transferring the retinoids from the cytoplasm to nuclear sites. Three dimensional modeling of the complex between retinoic acid-binding protein (CRABP) and retinoic acid suggests that binding of the ligand is mediated by interaction between the carboxyl group of retinoic acid and two charged arginines (111 and 131) whose side chains project into the interior of the barrel of the protein [124]. To assess the contribution of these amino acids to protein–ligand interaction, amino acid substitutions were made by site-specific mutagenesis. Like the wild-type CRABP, the mutant proteins are composed mainly of β-strands, as determined by circular dichroism in the presence and absence of ligand, and thus they are most likely folded into the same compact barrel structure as the wild-type protein. Mutants in which Arg-111 and Arg-131 are replaced by glutamine bind retinoic acid with significantly lower affinity than

the wild-type protein, suggesting that these two residues contribute to ligand binding. Thermostability is increased in the mutant, suggesting that repulsive forces between positively-charged arginine residues confer conformational flexibility on the native protein permitting retinoic acid to enter the binding pocket. Such effects simultaneously cause a loss in stability.

Several groups have reported a number of structures belonging to this family, including rat intestinal and liver I-FABP [125], chicken liver FABP and bovine heart FABP [126]. An apo-FABP from bovine heart has been studied with NMR spectroscopy [127]. The 3D structure of the complex of fatty acid binding protein from bovine heart (cFABP) with endogenous lipid was determined [128]. The crystal structure of rat holo I-FABP:palmitate, with a single molecule of bound palmitate, has been refined to 2 Å resolution using a combination of least squares methods, energy refinement and molecular dynamics and comparisons made with those of apo-I-FABP [119, 129]. The results show that the folding pattern is very similar for the apo and holo forms. The refined structures of *Escherichia coli*-derived rat intestinal fatty acid binding protein with bound oleate and that without bound ligand (apo I-FABP) have also been reported [120, 130]. Comparisons of the structure of I-FABP:oleate with a recently refined 1.19 Å model of apo I-FABP, and an earlier 2.0 Å model of I-FABP:palmitate, revealed a remarkable degree of similarity in the positions of their main chain and side chain atoms and in the conformations of the bound oleate and palmitate.

The crystal structure of recombinant murine adipocyte lipid-binding protein (ALBP) has been determined and refined to 2.5 Å [131]. ALBP is phosphorylated by the insulin receptor kinase upon insulin stimulation. The crystal structure of recombinant murine ALBP has been determined and refined to 2.5 Å with a final *R*-factor for the model of 18%. Crystalline ALBP has a conformation which is essentially identical to that of intestinal fatty acid binding protein and myelin P2 protein (also a member of the cellular hydrophobic ligand protein family). It contains a number of bound and implied unbound water molecules and shows no large obvious portal to the external environment. The cavity of ALBP, which by homology is the ligand-binding site, is formed by both polar and hydrophobic residues among which is found Tyr-19 which is phosphorylated by the insulin receptor kinase [132], suggesting that this region must display considerable conformational flexibility.

The 3D structure of P2 protein from peripheral nervous system myelin was determined at 2.7 Å resolution by X-ray crystallography [114, 133]. This protein consists of a single, flattened globular domain that is a 10-stranded up-and-down β-barrel similar to that of the cellular fatty acid binding proteins.

It has been confirmed that the holo-P2 structure may contain oleic or palmitic acid. The carboxyl group of the ligand is in the center of the barrel interacting with two buried arginine residues and a tyrosine. Because these arginines are well conserved for members of the family binding acidic ligands, but are glutamines in the cellular retinoid binding proteins (CRBP/CRBPII) molecules, it has been suggested they play an important role in determining the ligand specificity of the CRBPs [133].

EgDF1 protein was indirectly identified from *Echinococcus granulosus*, a parasite causing hydathidosis, which is one of the more important endemic illnesses in

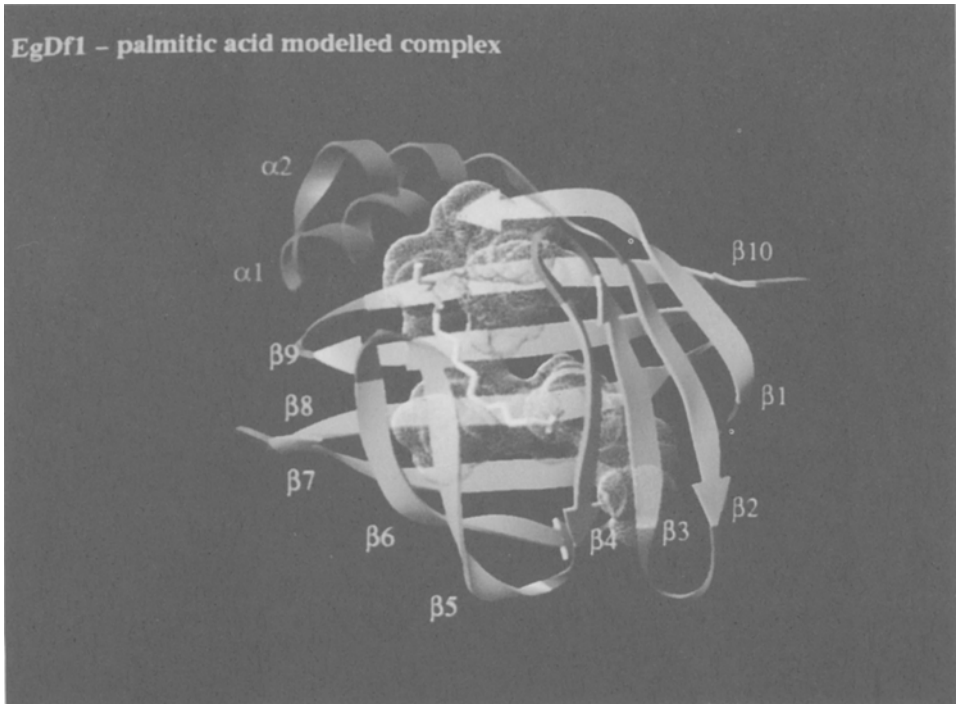

Fig. 3. Ribbon picture of the model-built EgDf1 in its complex with modeled palmitate.

tropical countries. The molecular and cellular bases of *E. granulosus* development are still largely unknown. Nevertheless, recent progress in developmental biology, showing common regulatory processes between distant organisms, and the increasing knowledge concerning the development of other platyhelminths, provided important clues for studying the molecular basis of *E. granulosus* development [134]. EgDf1 has a molecular weight of 17 ± 2 kD. Sequence analysis revealed that the protein presents 53% sequence similarity with respect to RBP and 72% (29.8% identity) with respect to I-FABP and a 40% identity with respect to the primary sequence of P2 myelin.

A first model was built for EgDf1 by taking the secondary structure of I-FABP. Figure 3 shows a ribbon picture of our model structure of EgDf1 with palmitate. The construction of a second model is in progress, based on the structure of P2 myelin. The modeled structures are being studied with respect to their stability to thermal effects with MD simulations. In general, the secondary structure from I-FABP was conserved, namely, a two β-barrel with 10 sheet strands and two α-helices. These are found at the same place in all members of the family. Between the fifth and sixth β-sheet strands, there is a loop which is shorter than that found in the original structure. If this structure is confirmed, a local change in the ligand-entrance region can be expected. The structure of palmitate from I-FABP was

docked into the model of EgDf1 and a similar procedure was adopted in order to model a V-shaped fatty acid (tentatively oleic acid) from the P2 myelin structure [133], to produce a model for the EgDf1:oleate complex. While differences are apparent, the oleate and palmitate can be equally well bound inside the β-barrel. MD simulations and free energy calculations are in progress.

A second application of model building techniques to this family of molecules is in the development of vaccines against schistosomiasis and fascioliasis. Like *Echinoccocus*, the parasites *Schistosoma mansonsoni* which infects an estimated 200 million individuals worldwide and *Fasciola hepatica* responsible for the infection of more than a million cattle and sheep, also produce fatty-acid binding proteins which belong to this same family. The molecule from *S. mansoni* has been named Sm14 [135] and is a highly protective antigen capable of reducing the level of infection of both parasites in experimental animals [136]. Thus Sm14 and its homologue from *F. hepatica* apparently show strong immunological cross-reactivity. These observations have led to the suggestion that Sm14 may serve as a useful bivalent antihelminth vaccine in the control of both diseases.

Individuals naturally infected with *S. mansoni* present a strong antibody response to recombinant Sm14 but give no signs of autoimmune disease, despite the fact that Sm14 shares $\approx$40% sequence identity with several human fatty acid binding proteins, including that from cardiac muscle [137] and the P2 myelin protein from peripheral nerve [138]. What may at first sight seem curious is that the level of sequence identity between the two parasite molecules, which share strong immunological cross-reactivity, is not markedly greater than the $\approx$40% described above. However, when one examines exactly which residues are conserved in both cases one observes markedly different patterns. Figure 4 shows the percentage sequence identity within a 25 residue scanning window for several comparisons of Sm14 with human fatty acid binding proteins and also between the two parasite molecules (in bold). The figure clearly demonstrates that the principal differences between the human proteins and Sm14 lie within the C-terminal third of the molecule whilst this same region shows an average sequence identity for the parasite molecules of $\approx$47%. This observation strongly suggests that this region of the molecule may play an important role in determining immunological cross-reactivity between the two parasites. Figure 5 shows the model of the Sm14 molecule and indicates those residues which are conserved in the two parasite sequences but which are most different from the human sequences. Two-thirds of these residues appear within the C-terminal region described above and the residues as a whole can be seen to cluster in two regions, one at each end of the β-barrel. The authors suggest that these are the principal epitopes of the Sm14 molecule responsible for the immunological cross-reactivity observed experimentally. The model is also finding application in the selection of peptides for testing as possible antigens in an alternative approach to vaccine development.

4.2. CARRIERS AND ION CHANNEL FORMING MOLECULES

The biological functions of ion-carriers, such as small ionophores, depend on the ability a given molecule has to selectively bind a particular ionic species, in preference to others. Differential solvation effects are the basis for such differ-

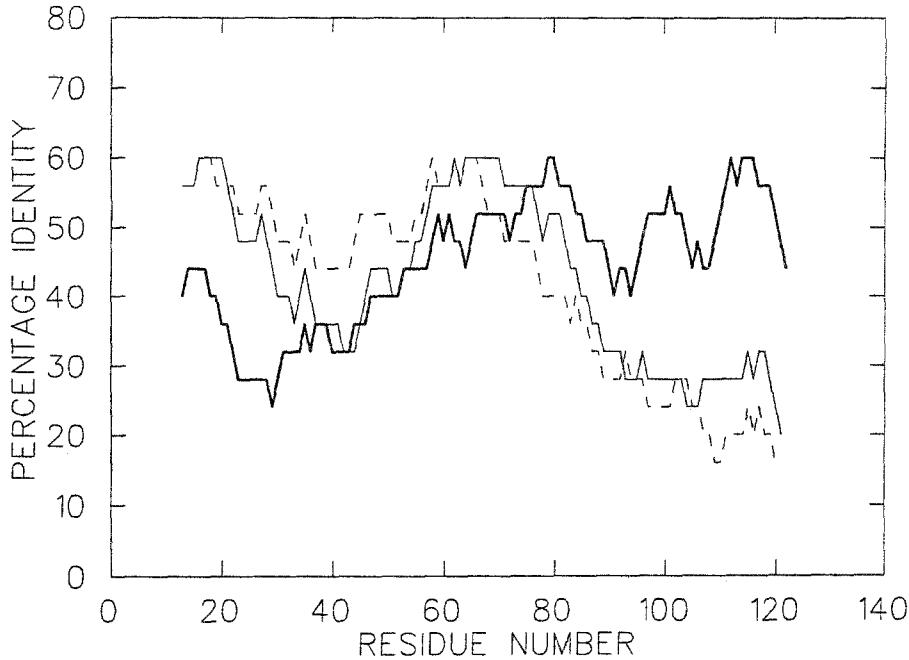

Fig. 4. Sequence identity plot comparing the fatty acid binding proteins of *Schistosoma* (Sm14) with *Fasciola* (bold), Sm14 with human cardiac FABP (solid line) and Sm14 with human P2 protein from periferal nerve (dashed line). The percentage identity is calculated within a 25 residue sliding window and plotted as a function of sequence. The two parasite sequences are seen to be most similar in the C-terminal third of the molecule where the human and Sm14 sequences differ most.

ences. Structure and selectivity issues have recently been discussed for calcium channels by several authors [139, 140]. Potassium channels have recently been reviewed [141]. Statistical mechanical studies of microscopic models are required to elicit the structural factors underlying the selectivity phenomenon. Such studies demand an adequate description of the interactions in metal-containing biological systems, e.g. metalloenzymes and metalloproteins. As the descriptive and predictive power of MD simulations and empirical quantum-chemical methods lies in the reliability of the potential energy functions used to describe interatomic interactions and in the quality of the parameters used in the force fields, a thorough comparison with experimental data is mandatory.

A set of empirical interaction potential functions were developed using the free energy perturbation approach, for various metal cations, especially designed to permit the theoretical treatment of metal-containing biological systems. The systems of interest involved solutions and solvated biomolecules. Attention was focused on models of ion interaction that can reproduce and predict properties in these systems. To this end, the experimentally observed hydration free energies and radial distribution functions were used to calibrate ion interaction potentials. Parameters for alkali and alkaline-earth ions were derived, as well as a simple model for transition metals that have *d*-orbital valence electrons. The resulting

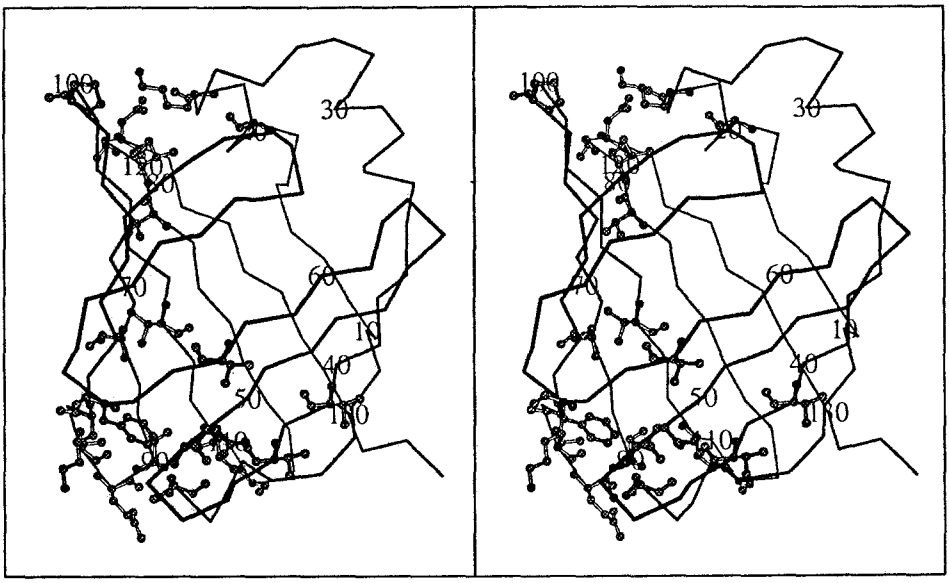

Fig. 5. Stereo figure of the model for Sm14. The residues indicated by the ball-and-stick representation are those observed to be identical in the parasite sequences (from *Schistosoma* and *Fasciola*) but which are not conserved in human sequences. They can be seen to cluster at the two poles of the barrel and probably correspond to the major epitopes.

potentials reproduced the available experimental data. These potential energy functions were applied to evaluate free energies of biomolecules in a number of test cases, involving metal ion catalyzed proton transfer reactions in solutions and in carbonic anhydrases. Ion binding to host molecules (valinomycin and gramicidin A) has been previously reported by Åqvist [142].

4.2.1. *Valinomycin*

Valinomycin i.s a small ion-carrier peptide having the ability to selectively and tightly bind small monovalent cations in several solvents. For alkaline cations, the selectivity order is $Rb^+ > K^+ > Cs^+ > Na^+ > Li^+$; note the anomalous position of the K^+ ion in this series. The primary sequence of valinomycin is cyclo-(-D-Hyl-D-Val-L-Lac-L-Val-)$_3$, where D-Hyl is D-hydroxyisovaleric acid and L-Lac is L-lactic acid. Depending on the dielectric constant of the solvent, the valinomycin molecule wraps around a cation using six ester carbonyl oxygens, thereby forming a folded structure stabilized by six intramolecular H-bonds between all the —NH groups and the amide carbonyl oxygens. One face of this structure is formed by D-Hyl and the other by L-Lac [142–144]).

Different experimental studies on ion binding to valinomycin agree very closely in the selectivity ratios in low-dielectric constant solvents. This implies that the ion–valinomycin complexes are essentially isosteric and that the relative complexation energy only depends on the ion–host interactions and not on solvent interactions.

The evaluation of relative complexation energies, and hence the selectivity of valinomicyn, would only require calculations *in vacuo*. The relative free energies of complexation between these cations and valinomycin were calculated as a test of the formerly derived parameters for alkaline metals. The correct selectivity sequence and the relative free energies were both well reproduced in these *in vacuo* calculations by Åqvist [142–144].

The differential binding energies obtained in water show that the conformation of the ion–valinomycin complexes changes according to the polarity of the solvent. MD free energy perturbation simulations were therefore used to examine the ion-binding energetics and selectivity of valinomycin in methanol [142–144], a solvent for which an extensive set of experimental data is available. The starting structure was the K^+-valinomycin complex. The ion parameters, taken from the already determined set, were then slowly perturbed so that the ion gradually changed from K^+ to Na^+ and then to Cs^+. GROMOS potential functions were used to perform the calculations. The methanol molecules were represented by three atoms with the CH_3 group taken as an extended atom. The relative and absolute solvation energies of the valinomycin–ion complexes in methanol were initially calculated to assess model reliability; this check is important since the parameter set was derived from *in vacuo* calculations. The resulting values were in reasonable agreement with experimental estimates. The ion-selective properties of the Cs^+, Rb^+, K^+ and Na^+ complexes with valinomicyn were then examined by performing a number of free energy perturbation simulations for different values of the partial charge on the six ligand carbonyl groups. This simulation aimed at studying the dependence of selectivity on the dipole moment of the ion ligands. The experimentally observed values were qualitatively reproduced by a carbonyl partial charge of about $\pm 0.58e$. This value tor the C=O dipole moment is also that which minimizes the structural deviation from the crystal conformation. As the C=O dipole moment was increased, the dependence of selectivity on the dipole moment of the ion ligands was found to obey early selectivity theories. Besides, the carbonyl partial charge, which reproduced the selectivity properties of valinomicyn in these simulations, was also quite similar to that obtained in the early *in vacuo* calculations ($q = \pm 0.55e$); this close agreement corroborates the notion of isostericity. On the other hand, since the available force fields were not parametrized for the type of ion–carbonyl interactions studied in this work, the results of the simulations were used to determine carbonyl parameters that reproduced the observed selectivity in a quantitative way. The optimal dipole moment which emerged from these calculations is somewhat larger than that normally used in molecular mechanics force fields and indicates that there is a significant polarization of the carbonyl groups by the field from the ion. The simulations also suggested that the unloading of ions would mainly occur through the lactic acid face of the molecule, since it does not provide as effective a shield against attack from the solvent as the more hydrophobic isohydroxyvaleric acid face.

4.2.2. *Ca^{2+} Binding Proteins*

The calcium binding proteins are a class of highly charged proteins that bind calcium ions with high affinity and selectivity. A large group of Ca^{2+} binding

proteins are characterized by the presence of specialized elements of protein structure, composed of amino acid sequence motifs forming characteristic helix-loop-helix structures that have been named 'EF-hands'. Calmodulin (CaM) and troponin C (TNC) belong to this family. In these molecules, the helix-loop-helix motif is found in sets of interacting pairs [145], which form near globular domains at the N- and C-termini of the molecule.

Calmodulin is a regulatory protein of widespread occurrence that regulates the activity of a large number of cellular proteins. Cellular proteins activated by calmodulin may act as enzymes, pumps, channels, regulators of gene expression and transducers of intracellular signals and depends on its specific binding [146]. Both CaM and TNC bind four Ca^{2+} ions per molecule (two in the N-terminal domain and two in the C-terminal domain). The Ca^{2+}-saturated form of calmodulin is the active form [147]. In order to trigger activation, CaM must bind Ca^{2+} at specific sites formed by charged oxygen bearing amino-acid side chains and polar groups coming from the loop and first turn of the second helix of the helix-loop-helix motif (EF-hand). The two proteins have different affinities for Ca^{2+} ions. In troponin C, there are two high-affinity and two low-affinity sites, while in CaM all the sites exhibit similar affinities, which lie between the two ranges of affinity present in TNC [148].

The selectivity for Ca^{2+} ions and the requirements for Ca^{2+}-dependent conformational change are key elements in the Ca^{2+}-dependent mechanism of interaction between CaM and the proteins it activates [146]. To elucidate the basis for Ca^{2+} selectivity exhibited by calcium-binding proteins and ion carriers, the relative solvation energies of Ca^{2+} and Mg^{2+} ions in complex with cyclo(-L-Pro-Gly-)3 (here indicated as PG_3) were estimated from molecular dynamics free energy difference simulations. This small syntethic peptide binds Ca^{2+} with an affinity comparable to that of the natural proteins. The results for the 1:1 and 1:2 complexes show that the ion selectivity of PG_3 can be traced back to the difference in the solvation energies of the competing ions in water. They explain why Mg^{2+} forms 1:1 complexes with PG_3, whereas with Ca^{2+} 1:2 complexes are obtained. The structural basis for the differential ion solvation depends on how well the protein accommodates the ion with respect to water [149].

The crystallographic structure, in both enzymes, has a characteristic dumbell shape. The two lobes are formed by the calcium-binding domains connected by a linking helical structure which is exposed to the solvent. However, recent studies suggest that the structure of these molecules in solution differs from that of the crystal, the former being more compact. This same type of structural reorganization is found when CaM is complexed to proteins [148] and peptides which represent the CaM recognition domains of target proteins [150–152].

MD simulations were carried out to examine the structural changes that CaM and TNC can undergo [146, 148]. In the case of calmodulin, the structures emerging from the MD simulations were significantly more compact than the initial ones; they yielded a nearly globular structure which was in good agreement with experimental observations made for complexes of the enzyme with peptides. This process was observed in all the trajectories; each one of which was started with different initial conditions. In this process, two interior pockets that contained a number of hydrophobic residues become exposed to the surrounding medium.

Therefore, the simulation enabled identification of specific residues involved in the process of compaction and pointed to specific CaM residues participating in ligand binding. Discrete steps in the mechanism of structural rearrangement from the extended structure found in the crystal to this more compact structure were also identified, implying that three Arg residues mediate the bending of the central α-helix. Structural and energetic considerations suggested a dynamic H-bonding pattern around the Arg residues, in which the side chain of Arg was engaged in a clockwise turn, while continuously changing its network of hydrogen bonds. The stabilization of the intermediate steps in the helix bending process, through the H-bond network, determined the direction of the compaction process. Initial model building studies, to locate potential binding sites of ligands such as trifluoperazine (TFP), indicated that the compacted structure had functional significance, thereby explaining the selective binding of molecules such as TFP in the N-terminal domain.

MD simulations designed to explore the differences in the degree of compactness between the structures of CaM and TNC molecules in the crystal and in solution have also been reported [148]. These simulations explored the energetic accessibility of the modified structures and their possible functional role. To account for solvent effects, the environment was modeled with a distance-dependent dielectric permittivity and discrete water molecules were also included, with starting positions identified in the crystals of CaM and TNC. Methods of macromolecular structure analysis, including linear distance plots, distance matrices and a matrix representation of hydrogen bonding, were used to analyze the nature, the extent, and the source of the dramatic structural differences between the computed structures of the molecules and their conformations in the crystal. The analysis of the longest simulation showed that the intradomain structure of the molecule was conserved, however, the crystallographically observed dumbell structure changed, due to a bending in the center of the central tether helix connecting the two domains which moved into close proximity. The dynamic existence of a non negligible fraction of the compacted structure in aqueous solutions of CaM was also suggested.

Molecular mechanics and computer-graphical methods were used to study the interaction between calmodulin and a series of seven phenotiazine drugs and penfluridol, aiming at identifying their binding sites [147]. These drugs were shown to inhibit the reaction of calmodulin with many enzymes, by binding to CaM directly. A molecular model of the α-helical sequences of calmodulin was built on the basis of its crystallographic structure. Energy calculations were performed for the protein model, the drugs and the protein-drug complexes and the energies displayed a good correlation with experimental inhibition data. Two possible binding sites were derived for phenotiazines and one binding site for penfluridol. The principal binding forces are provided by contacts between acidic amino acids and the electronegative substituents of the aromatic rings.

4.3. MEMBRANE PROTEINS

To our knowledge, structural information is available on only four types of membrane proteins. Two of them have been determined at high resolution by X-

ray crystallography: *Rhodobacter viridis* photosynthetic reaction center (PRC) by Deisenhofer and Michel [153]; and porins from *Rhodobacter capsulatus* [154], two from *E. coli* [155], and one from *R. blastica* [156]. Electron crystallography at a lower resolution has been used to determine two other 3D structures; bacteriorhodopsin by Hendersson *et al.* [157] and light-harvesting chlorophyll *a*/*b*–protein complex associated with photosystem II, LHC-II [158]. The 6 Å map of LHC-II has now been refined and presented as an atomic model based on a three-dimensional map at 3.4 Å resolution [159] and a review was recently reported by Kühlbrandt [160] where structure and function aspects are described.

The main feature that emerges from these structures (porin excepted) is the presence of polypeptide chains assembled as α-helices crossing one or more times the bacterial plasma membrane. However, in the trimeric porins, each monomer forms a beta-sheet barrel whose axis is perpendicular to the membrane.

4.3.1. *α-Helical Crossing Structures*

Experimental data about the energetic of insertion of hydrophobic polypeptides in membranes, suggest that the α-helices found in BR, PRC and LHCII might be established even in the absence of the protein core. Each hydrophobic transmembrane helix behaves as an independent folding domain and a two step theoretical folding model has been proposed by Popot [97]. The first step corresponds to the insertion of tranmembrane segments by hydrophobic interactions and their helicoidal folding via hydrogen-bond interactions. In this model, the surrounding lipid cage forces these segments to adopt a well-defined, three-dimensional structure. The next step consists in assembling them as a helical bundle. However, we cannot yet be sure that this scheme is universally followed by all the membrane proteins. There are at least two main reasons against such a possibility: first, the number of known structures is still insufficient to allow for reliable general conclusions to be drawn; and second, the exception represented by the structure of porins shows that α-helical hydrophobic bundles are not the only fold that can be assembled in a membrane.

BR structure and more recently rhodopsine maps have been used to help model G protein-coupled receptors as described in the preceding section. Recent work on membrane water channel proteins (aquaporin) has been reviewed by Engel *et al.* [161].

4.3.2. *β-Barrel Crossing Structures*

The outer membrane of Gram-negative bacteria plays an important role in the intrinsic resistance of these organisms to antibiotics by decreasing the flow of antimicrobial agents into the cell. Porins are channel-forming proteins in the outer membrane of Gram-negative bacteria as well as mitochondria and chloroplast. These molecules primarily act as molecular sieves letting through their channel hydrophilic low molecular mass compounds of less than 600 to 700 Da. Porins form trimers that are relatively stable towards high temperature and detergents [162]. In Gram-negative bacteria they serve as diffusional pathways for nutrients,

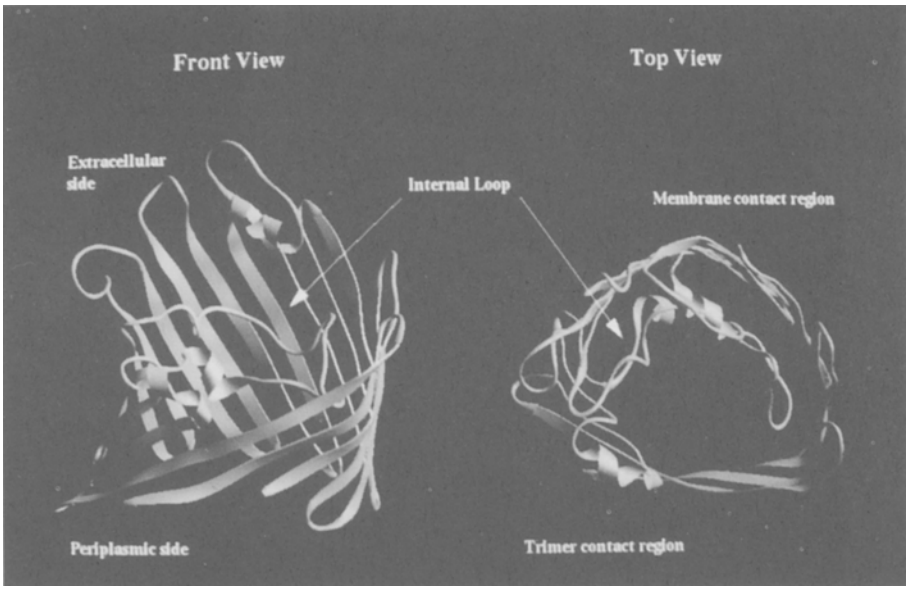

Fig. 6. Ribbon view of porin from *R. capsulatus*. The coordinates were kindly provided by Pr G. E. Schulz.

waste and antibiotics. The folding pattern of the *R. capsulatus* molecule is depicted in Figure 6.

Mycobacteria are Gram-positive bacteria which are major human pathogens and highly resistant to a wide range of antibiotics. The bacteria responsible for Tuberculosis and Leprosy are two examples. The lipid layer wall of these cells acts as a strong permeability barrier in a manner analogous to that of the Gram-negative bacteria and this cell wall is similarly resistant to antibiotics. Benz and coworkers [163] have recently shown the existence of porins in the cell wall of mycobacteria. For a general review the reader is refered to Cowan's paper [164].

The building of models of this class of proteins from scratch has been a difficult problem. Welte *et al.* [165] made a prediction of the general structure of the Gram-negative *E. coli* OmpF and PhoE from their sequences using the X-ray structure from *R. capsulatus* as a template. Although their primary sequences show little overall similarity, sequence alignment was obtained by using comparative hydrophilicity profiles. Thereafter, the folding of OmpF and PhoE were derived by using FRODO [166] and energy minimizing procedures. Recently, Benz and coworkers have proposed a model for the topology of PhoE which is very similar to the structure of porin from *R. capsulatus* [167]. The barrel motif was confirmed for the *E. coli* porins while differences could be correlated with their differing internal loop structures [155, 164].

A molecular dynamics study of the stability and plastic properties of the interior loop L3 in *R. capsulatus* porin has been carried out in our laboratory [168] using the so called NIS force field. The results reported showed a number of trends

agreeing with experimental data. The changes of average structure and fluctuation stability when the trimer fluctuation pattern is compared with that of the monomer suggested an important increase in molecular stability of the molecular complex. Early studies of β-barrel structures, as they are found in the retinol binding protein [117], showed that the stability of the model *in vacuo* is consistent with the structure determined crystallographically. It is experimentally well established that the trimer is fairly stable either to an increase in temperature or to detergent concentration. The monomer–monomer interactions, according to these calculations, would noticeably contribute to the stability of the trimer. The simulations suggest that the molecular determinant of the pore opening–closing mechanism would be located on a patch of residues in the loop L3. A number of conformational substates for this loop were detected.

More recently, Soares *et al.* [169] have made simulations of the voltage gating phenomenon. The single monomer model of Björkstén *et al.* [168] was used. There, in order to keep the molecular topology observed in the X-ray structure while simulating one single monomer, all the main chain atoms, except those belonging to the L3 loop, are position restrained. The use of position restraints allows the study of the dynamics of L3 without undesirable interference from deformations caused by the absence of the membrane and the remaining oligomers. The restraints do not permit deformation, but allow fluctuations around average positions. In addition, all side chain atoms are allowed to move freely, which in fact means that the L3 loop is studied in the fluctuating field of the whole molecule.

Voltage gating is seen as a perturbation of the electrostatic screening inside the porin pore, where, by the influence of the membrane potential gradient, water and counter-ion distribution can be slightly displaced from their equilibrium distribution. The fluctuation pattern of ion distribution in the solution around and inside the porin is assumed to be changed by changes in external voltage. The ions are the most mobile in the solution, thus an ionic gradient would arise, leading, among other things to changes in the equilibrium ionic distribution inside porin. One would expect that such a perturbation disrupts the screening of charged groups by counter ions (the equilibrium situation). If this happens, the effects of an asymmetric charge distribution in the pore would create an attractive force, pulling the L3 loop towards the positive patch. Thus, it would be the change in the electric field which creates the conditions for a conformational change in the L3 loop. This is the hypothesis that was put to the test, by studying the effects of changes in the electric field inside the pore by means of an electrostatic perturbation [169].

The model proposed above can be simulated by perturbing the screening electrostatics of ionizable groups inside the pore. Under these conditions, a localized conformational change takes place, involving twelve out of the 44 residues of the L3 loop. The pore is reduced to a sixth of its size in the open state. Interestingly, the conformational change could be achieved with a small perturbation and is shown to be reversible once this is switched off (relaxation process). Another type of behavior, predominating at higher temperatures, was also found for the loop. This involved an additional conformational change in the loop segment which completely closed the pore, but was not reversible under the simulation conditions.

5. Discussion

An overview has been presented of model building and computer assisted simulations of proteinases, receptor and transport proteins. Special emphasis has been given to structural aspects. The work reviewing proteinases extend and complement the report made in our previous paper [1]. Recently, metalloproteinases superfamilies and drug design have been overviewed by Blundell [170]. An overview of the spectacular progress achieved in G protein-coupled receptors modeling was presented here. Recent structural evidence backing allosteric behavior for these proteins is reviewed by Galzi and Changeux [171].

The production of new protein structures with X-ray and multidimensional NMR methods is gaining momentum and real coordinates for these receptors will be available in the near future. Site-directed spin labeling is being used to track structure and dynamics in membrane proteins, the technique has been extended to include time-resolved detection of structural changes with high spatial resolution [172, 173].

In the linear future, one would expect structural methods to produce detailed atomic information of molecular complexes of increasing size. For instance, the trend in structure determination of large molecular complexes is illustrated by the recent report on the 2.8 Å resolution of F_1-ATPase structure from bovine heart mitochondria [174]. This is the largest asymmetric structure ever to be solved [175]. One may expect in the near future fairly detailed structural information that will help understand structure/function relationships in G protein-coupled receptors [176]. In fact, Sprang and coworkers [177] have reported structures of active conformations of the α subunit $G_{i\alpha 1}$ and a discussion of the molecular mechanism of GTP hydrolysis was presented. One of the important points is the existence of a metastable GTP-bound conformational state of the α subunit that may act as an information carrier. The ability of a given molecular system to produce a receptor response is likely due to the fact that the conformational state of the effector or drug is in a neighborhood of a saddle point of index 1. This would help to trigger some other molecular conformational change in the drug–receptor complex. This scenario is in agreement with the ideas developed in previous papers from this group [1, 178–180].

In the approach advocated by Tapia and coworkers, the structural elements and information content on the particular chemical interconversion process are derived from the saddle points of index 1 (SPi-1) [178–180]. The transition vector is the eigenvector associated to the unique negative eigenvalue obtained after diagonalization of the Hessian matrix (second derivatives of the total energy with respect to geometric parameters). The surprising feature is that the SPi-1 calculated *in vacuo* for appropriate model systems **overlaps geometrically** with transition state analogues obtained from X-ray data. For transition vectors displaying the reactive fluctuation pattern that can be associated to the particular chemical interconversion step, and for those cases where the SPi-1 geometry overlaps with relevant X-ray data, the structural and dynamical information residing in its quadratic zone is relevant not only for describing the process *in vacuo* but also at the active site of the enzyme [178–180]. The substrate then binds to the enzyme with a geometry resembling that of a precursor complex. Precursor and successor complexes have

structures closely related to the geometry of the SPi-1: they do not correspond to either reactants or products at infinite intermolecular distance. The geometric overlap between transition state analogues attached to the enzyme and the SPi-1 calculated for molecular systems *in vacuo* has been found for new systems such as dihydrofolate reductase, lactate dehydrogenase [181] and glutathione reductase. It seems apparent that the substrate–enzyme system can be described as a process of reciprocal protein structural activation and substrate molding from the normal conformation towards a conformation as near as possible the precursor complex of the chemical interconversion step. This molding confers activity on the whole system. It is this very principle that was proposed to be in action for the drug–receptor complex [91, 182]. The active configuration of a given drug system would likely be related to conformations found in the neighborhood of saddle points of index one rather than minimum energy conformation. One can assume, as an hypothesis, that this may be a general principle.

Acknowledgements

This work has received important financial support from SAREC. C. M. S. gratefully acknowledges Junta Nacional de Investigação Científica e Tecnológica (Brazil) for financial support. O.T. thanks NFR for sustained support. We thank M. Campillo and L. Pardo for communicating their work on receptor modeling before publication.

References

1. O. Tapia, M. Paulino, and F. Stamato: *Mol. Eng.* **3**, 377 (1994).
2. J. Robertus: *Nature Str. Biol.* **1**, 352 (1994).
3. D. Donnelly, M. S. Johnson, T. L. Blundell, and J. Saunders: *FEBS Lett.* **251**, 109 (1989).
4. M. Bajaj and T. L. Blundell: *Ann. Rev. Biophys. Bioeng.* **13**, 453 (1984).
5. T. L. Blundell, D. Carney, S. Gardner, F. Hayes, B. Howlin, T. Hubbard, J. Overington, D. A. Singh, B. L. Sibanda, and M. Sutcliffe: *Eur. J. Biochem.* **172**, 513 (1988).
6. B. S. Hartley: *Phil. Trans. Roy. Soc. Lond.* **B257**, 77 (1970).
7. I. Schechter and B. Berger: *Biochem. Biophys. Res. Comm.* **27**, 157 (1967).
8. M. Kreiger, L. M. Kay, and R. M. Stroud: *J. Mol. Biol.* **83**, 209 (1974).
9. W. Bode and P. Schwager: *J. Mol. Biol.* **98**, 693 (1975).
10. M. E. P. Murphy, J. Moult, R. C. Bleackley, H. Gershenfeld, I. L. Weissman, and M. N. G. James: *Proteins: Struct. Funct. Genet.* **4**, 190 (1988).
11. S. Okada, C. K. Kam, L. Narasimhan, M. Poe Blake, J. T., O. Krahenbuhl, J. Tschopp, and J. C. Powers: *Biochemistry* **30**, 2217 (1991).
12. A. D. McLachlan and D. M. Shotton: *Nature (London)* **229**, 202 (1971).
13. J. Greer: *J. Mol. Biol.* **153**, 1027 (1981).
14. L. Jurásek, R. W. Olafson, P. Johnson, and L. B. Smillie: *Miami Winter Symp.* Miami, pp. 93 (1976).
15. R. J. Read, G. D. Brayer, L. Jurásek, and M. N. G. James: *Biochemistry* **23**, 6570 (1984).
16. L. T. J. Delbaere, W. L. B. Hutcheon, M. N. G. James, and W. E. Thiessen: *Nature* **257**, 758 (1975).
17. L. Hedstrom, L. Szilagyi, and W. J. Rutter: *Science* **255**, 1249 (1992).
18. G. R. Drapeau, Y. Boil, and J. Houmard: *J. Biol. Chem.* **247**, 6720 (1972).
19. L. Nienaber, K. Breddam, and J. J. Birktoft: *Biochemistry* **32**, 11469 (1993).
20. J. A. R. G. Barbosa, R. C. Garratt, and J. W. Saldanha: *FEBS Letts.* **324**, 45 (1993).
21. G. R. Drapeau: *Can. J. Biochem.* **56**, 534 (1978).

22. I. Svendsen and K. Breddam: *Eur. J. Biochem.* **204**, 165 (1992).
23. L. D. Graham, K. D. Haggett, P. A. Jennings, D. S. Le Brocque and R. G. Whittaker: *Biochemistry* **32** (1993).
24. M. Allaire, M. M. Chernaia, B. A. Malcolm, and M. N. G. James: *Nature* **369**, 72 (1994).
25. D. A. Jewell, W. Swietnicki, B. M. Dunn, and B. A. Malcolm: *Biochemistry* **31**, 7862 (1992).
26. H. P. Rang and M. M. Dale: *Pharmacology* (2nd Ed.), Churchill Livingstone, Edinburgh (1991).
27. L. Stryer: *Biochemistry*, Freeman, New York (1988).
28. M. F. Hibert, S. Trumpp-Kallmeyer, A. Bruinvels, and J. Hoflack: *Mol. Pharmacology* **40**, 8 (1991).
29. W. B. Pratt and P. Taylor (eds.): *Principles of Drug Action. The Basis of Pharmacology* (3rd Ed.), Churchill Livingstone Inc., New York, p. 836 (1990).
30. J. B. Moon and W. J. Howe: *Proteins* **11**, 314 (1991).
31. A. R. Kerlavage: *Curr. Op. Struct. Biol.* **1**, 394 (1991).
32. J. M. Baldwin: *EMBO J.* **12**, 1693 (1993).
33. L. Pardo, J. A. Ballesteros, R. Osman, and H. Weinstein: *Proc. Natl. Acad. Sci. (USA)* **89**, 4009 (1992).
34. H. Weinstein and R. Osman: *Neuropsychopharmacology* **3**, 397 (1990).
35. L. Pardo, J. Giraldo, M. Martin, and M. Campillo: *Mol. Pharmacol.* **40**, 980 (1991).
36. W. G. Richards (ed.): *Computer-Aided Molecular Design*, IBC Technical Services Ltd., London, p. 266 (1989).
37. D. Gordon: *Curr. Opin. Cell Biol.* **2**, 695 (1990).
38. F. A. Stephenson: *Curr. Op. Struct. Biol.* **1**, 569 (1991).
39. S. Froehner: *Nature* **366**, 719 (1993).
40. E. M. Kosower: *FEBS Letters* **231**, 5 (1988).
41. L. Smart, H. W. Meyers, R. Higenfeld, W. Saenger, and A. Maelicke: *FEBS Letters* **178**, 644 (1984).
42. M. Shibata and R. Rein: 'A Computer Modeling Study of Acetylcholine Receptor–Ligand Interactions', in R. Rein and A. Golombek (ed.), *Computer-Assisted Modeling of Receptor–Ligand Interactions. Theoretical Aspects and Applications to Drug Design*, Alan R. Liss, Inc., New York, p. 39 (1989).
43. S. Furois-Corbin and A. Pullman: *Biochim Biophys. Acta* **984**, 339 (1989).
44. S. Furois-Corbin and A. Pullman: *Biophys. Chem.* **39**, 153 (1991).
45. D. A. Dougherty and D. A. Stauffer: *Science* **250**, 1558 (1990).
46. V. Schmieden, J. Kuhse, and H. Betz: *Science* **262**, 256 (1993).
47. J. Amin and D. S. Weiss: *Nature* **366**, 565 (1993).
48. D. Colquhoun and M. Farrant: *Nature* **366**, 510 (1993).
49. E. M. Kosower: *Eur. J. Biochem.* **168**, 431 (1987).
50. H. R. Guy and G. Raghunathan: *Progr. Clin. Biol. Res.* **289**, 231 (1989).
51. P. A. Hargrave: *Curr. Op. Struct. Biol.* **1**, 575 (1991).
52. G. F. X. Schertler, C. Villa, and R. Henderson: *Nature* **362**, 770 (1993).
53. J. P. Noel, H. E. Hamm, and P. B. Sigler: *Nature* **366**, 654 (1993).
54. H. R. Bourne: *Nature* **366**, 628 (1993).
55. H. Weinstein and R. Osman: *Simulations of Ligand–Receptor Interactions as Guides for Design*, Munskgaard, Copenhagen (1989).
56. H. Weinstein, R. Osman, S. Topiol, and J. P. Green: *Ann. N.Y. Acad. Sci.* **367**, 434 (1981).
57. H. Weinstein, R. Osman, and A. P. Mazurek: *Simulations of Molecular Stereoelectronic Mechanisms for the Interaction of Hallucinogens and Indole Derivatives at 5-HT Receptors*, CRC Press, Florida, p. 199 (1987).
58. R. Osman, S. Topiol, L. Rubenstein, and H. Weinstein: *Mol. Pharmacol.* **32**, 699 (1987).
59. G. A. Mercier, R. Osman, and H. Weinstein: *Prog. Clin. Biol. Res.* **289**, 399 (1989).
60. E. N. Baker: *J. Mol. Biol.* **141**, 441 (1980).
61. J. P. Dijkman, R. Osman, and H. Weinstein: *Int. J. Quant. Chem.* **35**, 241 (1989).
62. G. A. Mercier, J. P. Dijkman, R. Osman, and H. Weinstein: 'Effects of Macromolecular Environments on Proton Transfer Processes: The Calculation of Polarization', in R. Carbo (ed.), *Quantum Chemistry: Basic Aspects, Actual Trends*, Elsevier, Amsterdam, p. 577 (1988).
63. H. Weinstein, R. Osman, and G. Mercier: *Recognition and Activation of a 5-HT Receptor by*

Hallucinogens and Indole Derivatives; US-HHS, NIDA: Maryland, Vol. 90, NIDA Research monograph, p. 243 (1988).
64. L. A. Rubenstein and R. Osman: *Theochem* **81**, 321 (1991).
65. J. A. Ballesteros and H. Weinstein: *Biophys. J.* **62**, 107 (1992).
66. D. Zhang and H. Weinstein: *J. Med. Chem.* **36**, 934 (1993).
67. D. Zhang and H. Weinstein: *FEBS Lett.* **337**, 207 (1994).
68. L. B. Kier: *J. Med. Chem.* **11**, 441 (1968).
69. C. R. Ganellin, E. S. Pepper, G. N. Port, and W. G. Richards: *J. Med. Chem.* **16**, 610 (1973).
70. C. R. Ganellin, G. N. J. Port, and W. G. Richards: *J. Med. Chem.* **16**, 616 (1973).
71. C. R. Ganellin: *J. Med. Chem.* **16**, 620 (1973).
72. L. Farnel, W. G. Richards, and C. R. Ganellin: *J. Med. Chem.* **18**, 662 (1975).
73. G. J. Durant, C. R. Ganellin, and M. E. Parsons: *J. Med. Chem.* **18**, 905 (1975).
74. W. G. Richards and J. Wallis: *J. Med. Chem.* **19**, 1250 (1976).
75. Y. G. Smeyers, F. J. Romero-Sanchez, and A. Hernandes-Laguna: *J. Mol. Struct.* (*Theochem*) **123**, 431 (1985).
76. Y. G. Smeyers, F. J. Romero-Sanches, and A. Hernandes-Laguna: 'Quantum mechanical study of the activity of some H_2-receptors of histamine', in: J. K. Seydel (ed.), *QSAR and Strategies in the Design of Bioactive Compounds*, Verlag Chemie, Wenheim, p. 374 (1985).
77. Y. G. Smeyers, A. Hernandes-Laguna, C. Munos-Caro, and F. J. Romero-Sanches: *J. Chim. Phys.* **84**, 663 (1987).
78. F. M. L. G. Stamato, M. A. Perez, and E. Longo: *J. Mol. Struct.* (*Theochem*) **210**, 447 (1990).
79. F. M. L. G. Stamato, E. Longo, M. A. Perez, M. A. P. Guimaraes, Jr., and Y. G. Smeyers: *J. Mol. Struct.* (*Theochem*) **254**, 505 (1992).
80. S. Topiol, H. Weinstein, and R. Osman: *J. Med. Chem.* **27**, 1531 (1984).
81. H. Weinstein, D. Chou, C. L. Johnson, S. Kang, and J. P. Green: *Mol. Pharmacol.* **12**, 738 (1976).
82. P. Reggio, S. Topiol, and H. Weinstein: *J. Med. Chem.* **29**, 2412 (1986).
83. A. P. Mazurek, R. Osman, and H. Weinstein: *Mol. Pharmacol.* **31**, 345 (1987).
84. H. Weinstein, A. P. Mazurek, R. Osman, and S. Topiol: *Mol. Pharmcol.* **29**, 28 (1986).
85. L. Pardo, A. P. Mazurek, R. Osman, and H. Weinstein: *Int. J. Quant. Chem. Quant. Biol. Symp.* **16**, 282 (1989).
86. E. E. J. Haaksma, G. M. D. O. den Kelder, H. Timmerman, P. Veriioijs, and W. Ravenek: *Progr. Clin. Biol. Res.* **289**, 361 (1989).
87. J. Giraldo, M. Martin, M. Campillo, and L. Pardo: *Mol. Pharmacol.* **42**, 373 (1992).
88. S. Topiol and M. Sabio: *J. Computer-Aided Mol. Design* **5**, 263 (1991).
89. G. Náray-Szabó, P. R. Surján, and J. G. Angyán: *Applied Quantum Chemistry*, Akadémiai Kiadó, Budapest (1987).
90. E. Kimura, T. Koike, and M. Kodama: *Chem. Pharm. Bull.* **32**, 3569 (1984).
91. O. Tapia, R. Cardenas, Y. G. Smeyers, A. Hernandez-Laguna, J. J. Randez, and F. J. Randez: *Int. J. Quantum Chem.* **38**, 727 (1990).
92. A. Hernandez-Laguna, Z. Cruz-Rodriguez, Y. G. Smeyers, G. A. Arteca, J.-L. M. Abboud, and O. Tapia: *Theochem* in press (1995).
93. A. Hernandez-Laguna, J.-L. M. Abboud, R. Notario, H. Homan, and Y. G. Smeyers: *J. Am. Chem. Soc.* **115**, 1450 (1993).
94. K. MaloneyHuss and T. P. Lybrand: *J. Mol. Biol.* **225**, 859 (1992).
95. D. Röper, E. Jacoby, P. Krüger, M. Engels, J. Grötzinger, A. Wollmer, and W. Strassburger: *J. Recept. Res.* **14**, 167 (1994).
96. J. Grötzinger, M. Engels, E. Jacoby, A. Wollmer, and W. Strassburger: *Prot. Eng.* **4**, 767 (1991).
97. J. L. Popot: *Modelisation de proteines membranaires*, J.-P. Frénoy (ed.), Les Cahiers IMABIO CNRS, CNRS, Paris, Vol. 7, p. 25 (1993).
98. J. C. Stoof and J. W. Kebabian: *Life Sci.* **35**, 2281 (1984).
99. P. Seeman, M. Watanabe, D. Grigoriadis, J. L. Tedesco, S. R. George, U. Svensson, J. L. G. Nilsson, and J. L. Neumeyer: *Mol. Pharmacol.* **28**, 391 (1985).
100. T. Liljefors and H. Wilkstrom: *J. Med. Chem.* **29**, 1896 (1986).
101. H. E. Katerinopoulos and D. I. Schuster: *Drugs of the Future* **12**, 223 (1987).
102. G. Taubes: *Science* **265**, 1034 (1994).

103. L. L. Iversen: *Nature* **365**, 393 (1993).
104. P. Seeman, H.-C. Guan, and H. H. M. Van Tol: *Nature* **365**, 441 (1993).
105. E. J. Lloyd and P. R. Andrews: *J. Med. Chem.* **29**, 453 (1986).
106. D. T. Manallack and P. M. Beart: *J. Pharm. Pharmacol.* **40**, 422 (1988).
107. S. Collin, D. P. Vercauteren, D. Vanderveken, G. Evrard, and F. Durant: *J. Comput. Aided Mol. Des.* **3**, 39 (1989).
108. S. G. Dahl, Ø. Advardsen, and I. Sylte: *Proc. Natl. Acad. Sci. USA* **88**, 8111 (1991).
109. I. Petersson and T. Liljefors: *J. Med. Chem.* **35**, 2355 (1992).
110. T. Liljefors and K. P. Bogeso: *J. Med. Chem.* **31**, 306 (1988).
111. I. Alkorta, H. O. Villar and J. J. Perez: *J. Comput. Chem.* **14**, 620 (1993).
112. C. J. Cramer and D. G. Truhlar: *J. Am. Chem. Soc.* **113**, 8305 (1991).
113. C. J. Cramer and D. G. Truhlar: *Science* **256**, 213 (1992).
114. T. A. Jones, S. W. Cowan, M. E. Newcomer, and T. Bergfors: *Frontiers in Drug Research*, Alfred Benzon Symposium No. 28, p. 267 (1990).
115. O. Tapia and J. Åqvist: 'Molecular dynamics as a tool for structural and functional predictions: The retinol binding protein and chloroplast C-terminal fragment of the L12 ribosomal protein cases', in: *Computer-Assisted Modeling of Receptor-Ligand Interactions: Theoretical Aspects and Applications to Drug Design*, Alan R. Liss Inc., p. 55 (1989).
116. P. Sandblom, J. Åqvist, T. A. Jones, M. E. Newcomer, W. F. van Gunsteren, and O. Tapia: *Biochem Biophys. Res. Comm.* **139**, 564 (1986).
117. J. Åqvist, P. Sandblom, T. A. Jones, M. E. Newcomer, W. F. van Gunsteren, and O. Tapia: *J. Mol. Biol.* **192**, 593 (1986).
118. M. E. Newcomer, T. A. Jones, J. Åqvist, J. Sundelin, U. Eriksson, L. Rask, and P. A. Peterson: *Embo J.* **3**, 1451 (1984).
119. J. C. Sacchettini, J. Gordon, and L. J. Banaszak: *Proc. Natl. Acad. Sci. USA* **86**, 7736 (1989).
120. G. Scapin, J. I. Gordon, and J. C. Sacchettini: *J. Biol. Chem.* **267**, 4253 (1992).
121. H. L. Monaco and G. Zanotti: *Biopolymers* **32**, 457 (1992).
122. J. Åqvist and O. Tapia: *J. Mol. Graph.* **5**, 30 (1987).
123. J. Åqvist and O. Tapia: *J. Mol. Graph.* **10**, 120 (1992).
124. J. Zhang, Z. Liu, and T. A. Jones: *Proteins* **13**, 87 (1992).
125. J. C. Sacchettini, J. I. Gordon, L. J. Banaszak, and J. Leonard: *J. Biol. Chem.* **26**, 5815 (1988).
126. G. Scapin, P. Spadon, M. Mammi, G. Zanotti, and H. L. Monaco: *Mol. Cell. Biochem.* **98**, 95 (1990).
127. C. Luecke, D. Lassen, H. J. Kreinenkamp, F. Spener, and H. Rueterjams: *Eur. J. Biochem.* **210**, 901 (1992).
128. A. Mueller-Fahrnow, U. Egner, T. A. Jones, H. Ruedel, F. Spener, and W. Saenger: *Eur. J. Biochem.* **199**, 271 (1991).
129. J. C. Sacchettini, J. I. Gordon, and L. J. Banaszak: *J. Mol. Biol.* **208**, 327 (1989).
130. J. C. Sacchettini, G. Scapin, D. Gopaul, and J. I. Gordon: *J. Biol. Chem.* **267**, 23534 (1992).
131. P. R. Young: *Prot. Eng.* **5**, 117 (1992).
132. M. K. Buelt, Z. Xu, L. J. Banaszak, and D. A. Bernlohr: *Biochemistry* **31**, 3493 (1992).
133. T. A. Jones, T. Bergfors, J. Sedzik, and T. Unge: *EMBO J.* **7**, 1597 (1988).
134. A. Esteves, B. Dallagiovanna, and R. Ehrlich: *Mol. Biochem. Parasit.* **58**, 215 (1993).
135. D. Moser, M. Tendler, G. Griffiths, and M.-Q. Klinkert: *J. Biol. Chem.* **226**, 8447 (1991).
136. M. Tendler, C. A. Brito, N. Serra-Freire, C. M. Diogo, M. S. Almeida, A. C. B. Delbem, R. C. Garratt, W. Savino, F. Silva, N. Katz, and A. J. G. Simpson: *Proc. Natl. Acad. Sci.* submitted (1994).
137. G. D. Offner, P. Brecher, W. B. Sawlivich, C. E. Costello, and R. F. Troxler: *Biochem. J.* **252**, 191 (1988).
138. M. Suzuki, K. Kitamura, Y. Sakamoto, and K. J. Uyemura: *Neurochem.* **39**, 1759 (1982).
139. G. Yellen: *Nature* **366**, 109 (1993).
140. J. Yang, P. T. Ellinor, W. A. Sather, J.-F. Zhang, and R. W. Tsien: *Nature* **366**, 158 (1993).
141. L. Yeh Jan and Y. Nung Jan: *Nature* **371**, 119 (1994).
142. J. Åqvist, O. Alvarez, and G. Eisenman: *J. Phys. Chem.* **96**, 10019 (1992).
143. J. Åqvist, O. Alvarez, and G. Eisenman: 'Computer modelling of ion binding sites in proteins', in: A. Pullman (ed.): *Membrane Proteins: Structures, Interactions and Models*, Kluwer, Dordrecht (1992).

144. J. Åqvist: *J. Mol. Struct. (Theochem)* **256**, 135 (1992).
145. J. Haiech, M. C. Kilhofers, T. J. Lukas, T. Craig, D. M. Roberts, and D. M. Watterson: *J. Biol. Chem.* **266**, 3427 (1991).
146. J. L. Pascual-Ahuir, E. L. Mehler, and H. Weinstein: *Mol. Eng.* **1**, 231 (1991).
147. H. D. Höltje and M. Hense: *J. Computer-Aided Mol. Design* **3**, 1011 (1989).
148. E. L. Mehler, J. L. Pascual-Ahuir, and H. Weinstein: *Prot. Eng.* **4**, 625 (1991).
149. F. Sussman and H. Weinstein: *PNAS* **86**, 7880 (1989).
150. M. Ikura, G. M. Clore, A. M. Gronenborn, G. Zhu, C. B. Klee, and A. Bax: *Science* **256**, 632 (1992).
151. W. E. Meador, A. R. Means, and F. A. Quiocho: *Science* **257**, 1251 (1992).
152. W. E. Meador, A. R. Means, and F. A. Quiocho: *Science* **262**, 1718 (1993).
153. J. Deisenhofer and H. Michel: *Science* **245**, 1463 (1989).
154. M. S. Weiss and G. E. Schulz: *J. Mol. Biol.* **227**, 493 (1992).
155. S. W. Cowan, T. Schirmer, G. Rummel, M. Steiert, R. Ghosh, R. A. Pauptit, J. N. Jansonius, and J. P. Rosenbush: *Nature* **358**, 727 (1992).
156. A. Kreusch, A. Neubüser, E. Schiltz, J. Weckesser, and G. E. Schulz: *Protein Science* **3**, 58 (1994).
157. R. Hendersson, J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing: *J. Mol. Biol.* **213**, 899 (1990).
158. W. Kühlbrandt and D. N. Wang: *Nature* **350**, 130 (1991).
159. W. Kühlbrandt, D. N. Wang, and Y. Fujiyoshi: *Nature* **367**, 614 (1994).
160. W. Kühlbrandt: *Curr. Op. Struct. Biol.* **4**, 519 (1994).
161. A. Engel, T. Walz, and P. Agre: *Curr. Op. Struct. Biol.* **4**, 545 (1994).
162. E. Schiltz, A. Kreusch, U. Nestel, and G. E. Schulz: *Eur. J. Biochem.* **199**, 587 (1991).
163. J. Trias, V. Jarlier, and R. Benz: *Science* **258**, 1479 (1992).
164. S. W. Cowan: *Curr. Op. Struct. Biol.* **3**, 501 (1993).
165. W. Welte, M. S. Weiss, U. Nestel, J. Weckesser, E. Shiltz, and G. E. Schulz: *BBA* **1080**, 271 (1991).
166. C. C. Cambillau and E. Horjales: *J. Mol. Graph.* **5**, 174 (1987).
167. M. Struyvé, J. Visser, H. Adriaanse, R. Benz, and J. Tommassen: *Mol. Microbiol.* **7**, 131 (1993).
168. J. Björkstén, C. M. Soares, O. Nilsson, and O. Tapia: *Prot. Eng.* **7**, 487 (1994).
169. C. M. Soares, J. Björkstén, and O. Tapia: *Protein Eng.* **7**, in press (1995).
170. T. L. Blundell: *Nature Struct. Biol.* **1**, 73 (1994).
171. J.-L. Galzi and J.-P. Changeux: *Curr. Op. Struct. Biol.* **4**, 554 (1994).
172. J. L. Spudlich: *Nature Struct. Biol.* **1**, 495 (1994).
173. W. L. Hubbel and C. Altenbach: *Curr. Op. Struct. Biol.* **4**, 566 (1994).
174. J. P. Abrahams, A. G. W. Leslie, R. Lutter, and J. E. Walker: *Nature* **370**, 621 (1994).
175. R. L. Cross: *Nature* **370**, 594 (1994).
176. D. Donnelly and J. B. C. Findlay: *Curr. Op. Struct. Biol.* **4**, 582 (1994).
177. D. E. Coleman, A. M. Berghuis, E. Lee, M. E. Linder, A. G. Gilman, and S. R. Sprang: *Science* **264**, 1405 (1994).
178. Tapia and J. Andrés: *J. Mol. Struct. (Theochem)* submitted (1995).
179. O. Tapia, J. Andrés, and V. S. Safont: *J. Phys. Chem.* **98**, 4821 (1994).
180. O. Tapia, J. Andrés, and V. S. Safont: *J. Chem. Soc. Faraday Trans.* **90**, 2365 (1994).
181. J. Andrés, V. Moliner, and V. S. Safont: *Chem. Soc. Faraday Trans.* **90**, 1703 (1994).
182. O. Tapia and O. Nilsson: 'Molecular dynamics computer modelling and protein engineering', in: J. Bertrán (ed.), *Molecular Aspects of Biotechnology: Computational Models and Theories*, Kluwer Academic Publishers, p. 123 (1991).