G protein-coupled receptors as challenging druggable targets: insights from *in silico* studies

Stefano Moro,*[a] Magdalena Bacilieri,[a] Francesca Deflorian[a] and Giampiero Spalluto[b]

Received (in Montpellier, France) 18th November 2005, Accepted 9th January 2006
First published as an Advance Article on the web 9th February 2006
DOI: 10.1039/b516389g

The successful identification of hundreds of G protein-coupled receptors (GPCRs) represents the single greatest opportunity for novel drug development today. The crystal structure of rhodopsin provides the first information on the three-dimensional structure of GPCRs, which now supports homology modeling studies and structure-based drug-design approaches. We review our recent work on adenosine receptors, a family of GPCRs. Focusing our attention on A_3 adenosine receptor, we have demonstrated that the reciprocal integration of different theoretical and experimental disciplines can be very useful for the successful design of new, potent and selective receptor ligands.

G protein-coupled receptors: intriguing molecular machines

The G protein-coupled receptors (GPCRs) family can be considered one of the most important, intriguing, and challenging molecular machines designed by mother nature. GPCRs represent the most efficient signaling system used by cells to establish relationships with the external environment. They constitute a very large family of heptahelical, integral membrane proteins that mediate a wide variety of physiological processes, ranging from the transmission of the light and odorant signals to the mediation of neurotransmission and hormonal actions.[1,2] Indeed, they control the activity of enzymes, ion channels and transport of vesicles principally *via* the catalysis of the GDP–GTP exchange on heterotrimeric G proteins (Gα–βγ).[3,4] However, it has been known for a long time that GPCRs interact directly not only with G proteins, but also with several other important proteins involved in the control of cellular homeostasis such as arrestins,[5,6] or PDZ domain-containing proteins (Fig. 1).[7] In particular, these arrestins bind specifically to GPCRs phosphorylated by G protein-coupled receptor kinases (GRKs),[8] an interaction which participates in the homologous desensitization of the receptor by disturbing their coupling to G proteins.[8] However, the arrestin proteins, which were initially discovered due to their role in GPCR desensitization, serve equally important roles in regulating internalization and alternative signaling events.[8]

The classical view of GPCR–G protein coupling stoichiometry is one receptor for one G protein. It is now generally accepted that G protein-coupled receptors (GPCRs) can exist as dimers or as part of larger oligomeric complexes.[8,9] Increasing evidence suggests that a dimer is the minimal functional structure, but considerable variation exists between reports of the effects of agonist ligands on quaternary structure.[9,10] Many studies have intimated the existence of heterodimeric GPCR pairings.[9,10] Key questions that remain to be addressed effectively include the prevalence and relevance of these in native tissues and the implications of heterodimerization for pharmacology and, potentially, for drug design.[11]

G protein-coupled receptors are the largest family of cell-surface receptors involved in signal transmission with some estimated 800 human receptor genes (≈3% of the human genome). Several classification systems have been used to sort out this superfamily. Some systems group the receptors by how their ligand binds, and others have used both physiological and structural features. One of the most frequently used systems uses clans (or classes) A, B, C, D, E, and F, and

[a] *Molecular Modeling Section, Dipartimento di Scienze Farmaceutiche, Università di Padova, Via Marzolo 5, I-35131 Padova, Italy. E-mail: stefano.moro@unipd.it*
[b] *Dipartimento di Scienze Farmaceutiche, Università degli Studi di Trieste, Piazzale Europa 1, I-34127 Trieste, Italy*

Stefano Moro is Associate Professor of Medicinal Chemistry at the University of Padova, Italy. He received a B.Sc. degree in Medicinal Chemistry and a Ph.D. degree in Physical Organic Chemistry from the University of Padova, Italy. He was a Postdoctoral Fellow at the National Institutes of Health in Bethesda, USA and Visiting Professor at ETH of Zurich, Switzerland. He is the chief of the Molecular Modeling Section at the University of Padova, Italy. His main interest concerns the application of computational approaches to medicinal chemistry. He loves G protein-coupled receptors and, in particular, adenosine receptors. He is author of more than 100 papers published in international journals and monographs. He is the recipient of several national and international awards.

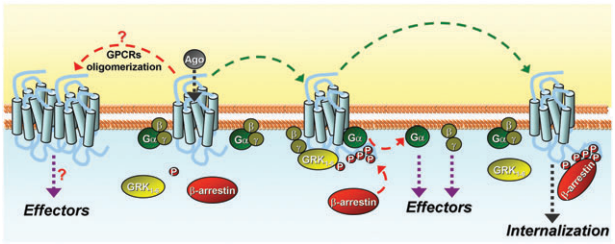

Fig. 1 Signal transduction pathways associated with the activation of the human adenosine receptors. Abbreviations: α, α-subunit of G protein; βγ, βγ-subunits of G protein; ATP, adenosine triphosphate; cAMP, cyclic adenosine monophosphate; DAG, diacylglycerol; G_i, G_i family of G proteins; G_s, G_s family of G proteins; G_o, G_o family of G proteins; G_q, G_q family of G proteins; IP3, inositol (1,4,5)-trisphosphate; P, phosphate moiety; PKC, protein kinase C; PLC, phospholipase C.

subclans are assigned using roman number nomenclature.[12,13] This A–F system is designed to cover all GPCRs, in both vertebrates and invertebrates. Some families in the A–F system do not exist in humans. Examples of this are clans D and E, which represent fungal pheromone receptors and cAMP receptors, family IV in clan A, which is composed of invertebrate opsin receptors, and clan F, which contains archaebacterial opsins. The overall classification of the GPCRs has been hampered by the large sequence differences between mammalian and invertebrate GPCRs. The GPCRs in *Drosophila melanogaster* show in many cases little resemblance to those in mammals.[14] Certain species show also a high difference in the numbers of receptor genes in different classes. *Caenorhabditis elegans*, a worm, has, for example, developed a remarkable number of chemosensory (olfactory) GPCRs related to the creature's specific lifestyle. Those chemosensory receptors, as well as the olfactory receptors in *D. melanogaster*, do not show any clear resemblance to the olfactory receptors in humans.

Recently, Fredriksson and co-workers collected a large set of GPCR sequences in the human genome and performed multiple phylogenetic analyses.[15,16] They identified more than 800 human GPCR sequences and simultaneously analyzed 342 unique functional nonolfactory human GPCR sequences with phylogenetic analyses. Fredriksson's classification subdivided GPCRs into five main families, named glutamate, rhodopsin, adhesion, frizzled/taste2, and secretin, forming the GRAFS classification system (Fig. 2). The rhodopsin family is the largest and forms four main groups with 13 sub-branches. Positions of the GPCRs in chromosomal paralogon regions indicate the importance of tetraploidizations or local gene duplication events for their creation. They also searched for "fingerprint" motifs using Hidden Markov Models delineating the putative inter-relationship of the GRAFS families. Fredriksson and co-workers have shown several common structural features indicating that the human GPCRs in the GRAFS families share a common ancestor. Remarkably, this study represents the first overall map of the GPCRs in a single mammalian genome.[15,16]

Many features of GPCR structure and function have been reviewed recently.[1,2,17,18] Here, we will highlight some fundamental features that may expand upon the classical view of GPCR structure and function. Sequence comparison between the different GPCRs revealed the existence of different receptor families sharing no sequence similarity. However, all these receptors have in common a central core domain consisting of seven transmembrane helices (TM1 through TM7) connected by three intracellular (IL1, IL2 and IL3) and three extracellular (EL1, EL2 and EL3) loops. Two cysteine residues (one in TM3 and one in EL2) which are conserved in most GPCRs, form a disulfide link which is possibly crucial for the packing and for the stabilization of a restricted number of conformations of these seven TMs. Aside from sequence variations, GPCRs differ in the length and function of their N-terminal extracellular domain, their C-terminal intracellular domain and their intracellular loops (Fig. 3). Each of these domains provide very specific properties to these various receptor proteins.

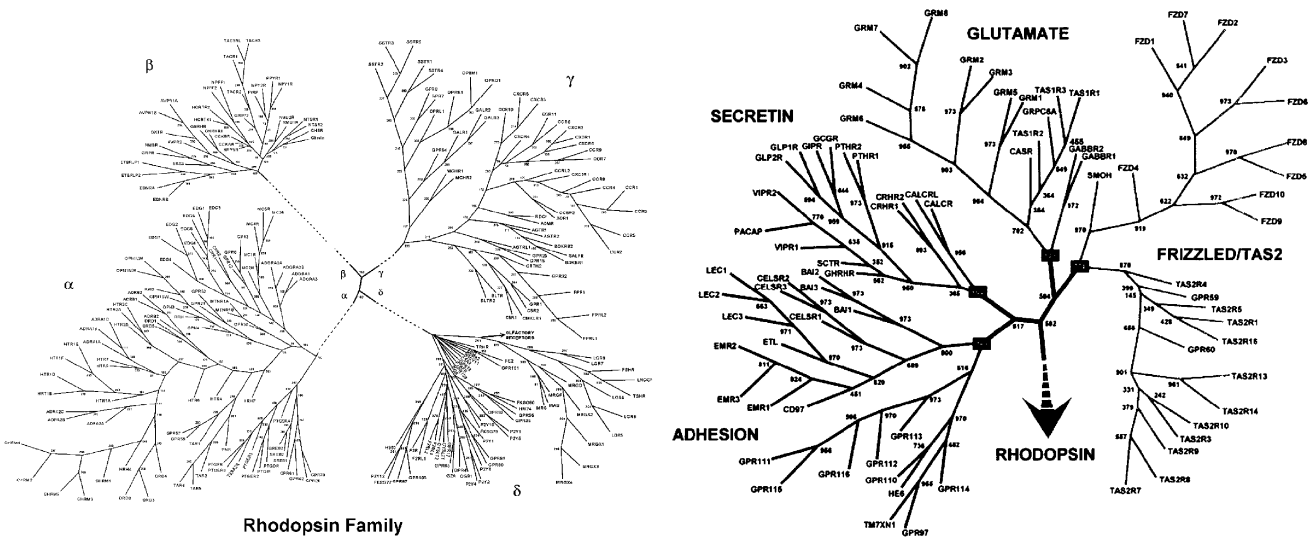

Fig. 2 *On the left*: phylogenetic relationship between the GPCRs (TM1–TM7) in the human genome. *On the right*: the phylogenetic relationship between GPCRs (TM1–TM7) in the human rhodopsin family.

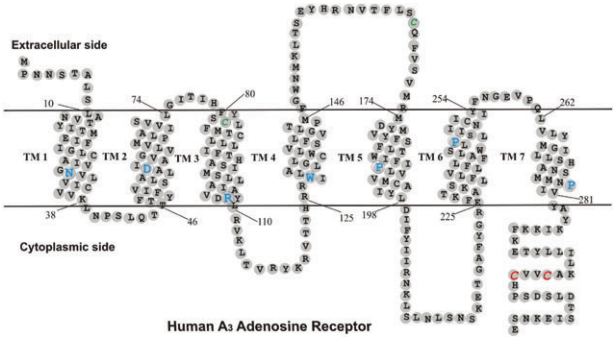

Fig. 3 Schematic representation of the membrane topology of the human A_3 adenosine receptor. Each of the 7 TMs have at least one characteristic residue (blue colour), which is found among the majority of family A receptors (Asn30(1.50); Asp58(2.50); Arg108(3.50); Trp135(4.50); Pro189(5.50); Pro245(6.50); and Pro279(7.50)). Disulfide bridge formation between Cys83 (TM3) and Cys166 (EL2) (green colour), palmitoylation sites (Cys300 and/or 303, red colour) in the C terminus.

Disregulation of GPCRs has been found in a growing number of human diseases,[19,20] and GPCRs have been estimated to be the direct or indirect targets of nearly half of the drugs used in clinical medicine today.[21,22] Thus, understanding how GPCRs function at the molecular level is an important goal of biological research. Nowadays, GPCR-based drug discovery has focused on the development of more selective and/or potent molecules that act at the receptor's orthosteric site (*i.e.* the binding site utilized by the endogenous agonist).[22,23] However, significant challenges remain. For example, some GPCR families display high sequence conservation within the orthosteric binding domain across receptor subtypes, making the development of subtype-selective orthosteric ligands difficult.[22,23] Furthermore, new cell-based functional assays, which are replacing the radioligand binding assay as the primary method for GPCR high-throughput screening (HTS), are now identifying molecules that modify receptor signaling while exerting a negligible effect in standard radioligand binding assays.[22] In addition to the classic orthosteric site, therefore, it is likely that GPCRs contain other binding sites that mediate biologically relevant functions. One promising class is the GPCRs' allosteric modulators,[24,25] which are defined as ligands that bind to a site on the receptor that is topographically distinct from the orthosteric binding site but regulate the action of agonists or antagonists. Moreover, the mechanisms controlling ligand binding, activation and signal transduction of the GPCRs–G protein system as well as the mechanisms required for defining the specificity of the receptor–G protein-effector interaction and the efficiency and regulation of signal transduction are highly complex and multifactorial.[1–4] Knowledge and mapping of the structural determinants and requirements for optimal receptor function are of paramount importance for understanding the molecular basis of ligand action and receptor function in normal and abnormal conditions. Deciphering structure–function relationships in GPCRs will promote computer-aided drug discovery by elucidating the binding mode(s) of known ligands into their receptor binding-sites and identify the pharmacophores involved.

Based on experience in this field, the structure-based approach to GPCR drug discovery in the absence of the real structures requires a multi-disciplinary approach, where molecular models represent a structural context to efficiently integrate experimental data and inferences derived from molecular biology, biophysics, bioinformatics, pharmacology and organic chemistry methods.[22,23,26] Although not always achievable, the success of a synergistic effect among these disciplines is highly dependent on the experimental design. Synergy is best achieved when mutations are structurally interpretable, structural hypotheses are experimentally testable, ligands are well characterized pharmacologically, and the necessary chemical modifications of the ligands are available. The extent to which these conditions are met define the quality of the information derived to guide the lead compound optimization process. In the majority of reports that describe computer simulations of GPCRs, the purpose of model construction was to describe the binding pocket for the ligand within the receptor. Models of ligand–receptor complexes were constructed incorporating experimental evidence of the specific receptor amino acid residues with which the ligands were found to interact. These models are being used, for example, in efforts to discover GPCR agonist or antagonist ligands. Another important use of GPCR models is to generate hypotheses regarding the binding and signaling functions of these receptors to be tested experimentally.[22,23,26] We believe that the most effective use of models, however, is as part of an iterative, bidirectional approach in which models are used to generate hypotheses that are tested by experimentation and the experimental findings are, in turn, used to revise and refine the model. The success of this approach is due to the synergistic interaction between theory and experiment. We will emphasize the latter approach in this review.

The evolution of the field of computer-aided design of GPCR ligands (both agonists and antagonists) has depended on the availability of suitable molecular receptor templates.[25–29] In fact, due to technical difficulties, which complicate experimental X-ray crystallography and NMR structure determination of GPCRs, the 3D structure of most GPCRs is still unknown. The only known GPCR structure, a 2.8 Å resolution structure of rhodopsin, was published only recently by Palczewski and collaborators.[30] The structure of rhodopsin is so far a suitable template for the resting state (resembling an *antagonist-like state*) of most family A GPCRs.[25–29] It is still very questionable if Palczewski's bovine rhodopsin structure could also be an appropriate template for the active state (*agonist-like state*) of the receptor. Even if several scientists have used Palczewski's bovine rhodopsin structure as the starting point to model the active state of different GPCRs, we know very well that the seven TM region constitutes the core domain of these receptors, and a change in conformation of this domain is probably responsible for receptor activation.[31] The change in conformation of the core domain generally affects the conformation of the IL2 and IL3 intracellular loops (which are directly linked to TM3 and TM6, respectively) that constitute one of the key sites for G protein recognition and activation. Although it has been proposed that a similar change in conformation of the core domain is associated with GPCR activation, diverse molecular mechanisms have been selected during evolution

to allow the natural ligand to induce this change in conformation. For this reason, we decided to focus our attention on the corresponding resting state of GPCRs until future crystallographic efforts are able to elucidate the structure of rhodopsin, or other GPCRs, in their active state.

As anticipated, rhodopsin-based homology modeling represents a widely used and well-consolidated approach to visualize GPCR three-dimensional models.[25–29] Modeling ligand–receptor complexes, as provided by the crystallographic structure of rhodopsin, enables the study of specific interactions at higher resolution. To evaluate which computational approaches are most suitable to model ligand–receptor interactions on a structural template, we have adopted the following ligand–receptor-based integrated strategy. (a) GPCR three-dimensional model(s) are constructed and refined using a conventional rhodopsin-based homology modeling approach. At this stage, the model can be used to probe specific interactions at the atomic level detected using site-directed mutagenesis analysis. In fact, the ligand–receptor interactions need to be identified in order to guide docking of the ligand into the receptor pocket. (b) Generation of 3D receptor-driven pharmacophore(s). Docking the structure of lead compounds onto a receptor model enables the exploration of the potential effects of chemical modifications of the compounds, and thus facilitates the lead optimization process. (c) Generation of 3D receptor-driven quantitative structure–activity relationship (3D-QSAR). We recognized that receptor-based 3D-QSAR can be used as an alternative and robust computation tool to predict the binding affinity of all compounds docked into the binding site. (d) Selection, synthesis and pharmacological characterization of new A_3 receptor antagonist candidates. This is the last step, but not the least, of our integrated approach (Fig. 4).

The rational design of new human A_3 antagonists as a key study

A_3 adenosine receptors (ARs) belong to the adenosine receptor family of GPCRs, which consists of four distinct subtypes, A_1, A_{2A}, A_{2B} and A_3. ARs are ubiquitously expressed in the human body.[32] Many cells express several AR subtypes, although in different densities. All subtypes, including the A_3 receptor, have been cloned from a variety of species including rat and human.[32] Species differences for A_3 receptors are larger than for other AR subtypes, particularly between rodent and human receptors (only 74% sequence identity between rat and human A_3 amino acid sequence). This results in different affinities of ligands, particularly antagonists, for rat *versus* human A_3 receptors. A_3 adenosine receptors are negatively coupled to adenylate cyclase *via* $G_{i2,3}$.[32,33] Coupling to $G_{q/11}$ leading to a stimulation of phospholipase C has been described.[34] Coupling of A_3 adenosine receptor to phospholipase D has also been demonstrated.[34] A_3 adenosine receptor stimulation can lead to activation of ERK1/2. In fact, A_3 adenosine receptor agonists stimulate PI3K-dependent phosphorylation of Akt leading to the reduction of basal levels of ERK1/2 phosphorylation, which in turn inhibits cell proliferation.[35] After exposure to agonist, A_3 adenosine receptors undergo rapid desensitization *via* phosphorylation by G protein receptor kinase (GRK2) at the intracellular terminal chain (particularly at threonine 318 on the rat receptor).[36]

The A_3 adenosine receptor, which is the most recently identified adenosine receptor is implicated in a variety of important physiological processes.[33] Activation of A_3 adenosine receptors increases the release of inflammatory mediators, such as histamine, from rodent mast cells,[37] and inhibits the production of tumor necrosis factor-α.[38] The activation of the A_3 adenosine receptor is also suggested to be involved in immunosuppression and in the response to ischemia of the brain and heart.[39] It is becoming increasingly apparent that agonists or antagonists of A_3 adenosine receptors have potential as therapeutic agents for the treatment of ischemic and inflammatory diseases.[40]

As we will summarize in the paper, the development in particular of antagonists for the A_3 receptors has so far been directed by traditional medicinal chemistry.[41] However, in the last few years we optimized a multidisciplinary integrated approach to speed up the discovery and the structural refinement of new potent and selective A_3 receptor antagonists.[23,29,42]

Following the above described computational philosophy, we have recently revisited the 3D model of the human A_3 receptor in its resting state (antagonist-like state), using the high-resolution crystal structure of bovine rhodopsin as a template, which can be considered a further refinement in building the hypothetical binding site of the A_3 receptor antagonists.[23,29,42] The flow-chart of our rhodopsin-based homology modeling is reported in Fig. 5.

Special care had to be given to the second extracellular (EL2) loop, which has been described in bovine rhodopsin to fold back over transmembrane helices, and, therefore, limit the size of the active site. As for the high-resolution structure of rhodopsin, also the human A_3 model reveals a seven-helix bundle with a central cavity surrounded by helices 1–3 and 5–7 (see Fig. 6). Helix 4 is not part of the cavity wall and makes contacts only with helix 3. The central cavity ($\cong$670 Å^3) is accessible from the cytosol, but the hairpin (EL2) between helices 4 and 5 prevents access from the periplasm. This hairpin lies between the helices, roughly parallel to the

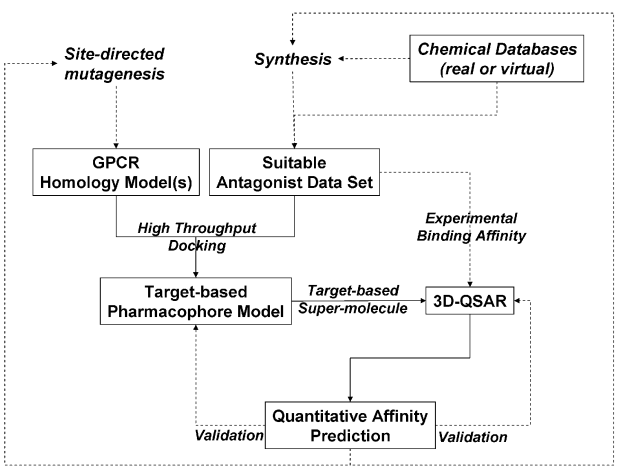

Fig. 4 Flow-chart describing our integrated target-based and ligand-based drug design approach.

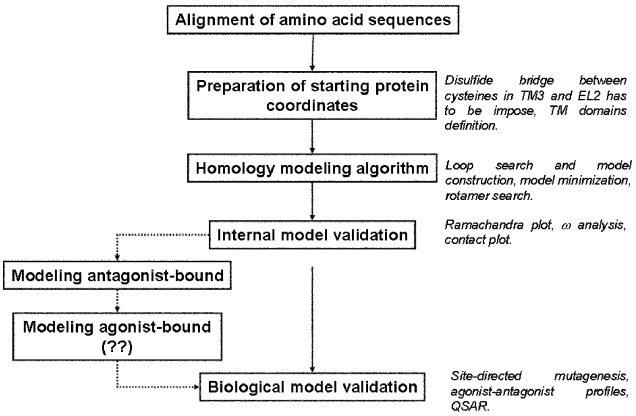

Fig. 5 Rhodopsin-based homology modeling protocol.

membrane surface. It has contacts with side chains of most of the helices.

The most prominent contact is a disulfide bridge to helix 3 (TM3). The ligand (antagonist) recognition seems to occur in the upper region of the TM helical bundle. TMs 3, 5, 6 and 7 seem to be crucial for the recognition of both agonists and antagonists. To test the quality of our theoretical model, we decided to mutate several residues of the human A_3 adenosine receptor within TMs 3 and 6 and EL2, which have been predicted by previous molecular modeling to be involved in the ligand recognition, including His95, Trp243, Leu244, Ser247, Asn250, and Lys152.[43–45] The N250A mutant receptor lost the ability to bind both radiolabeled agonists and antagonists. The H95A mutation significantly reduced the affinity of both agonists and antagonists. In contrast, the K152A (EL2), W243A (6.48) and W243F (6.48) mutations did not significantly affect the agonist binding but decreased antagonist affinity by 3–38 fold, suggesting that these residues were critical for the high affinity of A_3 adenosine receptor antagonists.[45] The A_3 agonist 2-chloro-N^6-(3-iodobenzyl)-5′-*N*-methylcarbamoyladenosine stimulated phosphoinositide turnover in the wild-type but failed to evoke a response in cells expressing W243A and W243F mutant receptors, in which agonist binding was less sensitive to guanosine 5′-γ-thiotriphosphate than in wild-type.[45] Thus, although not important for agonist binding, Trp243 was critical for receptor activation. The results were successfully interpreted using our rhodopsin-based model of ligand–A_3 receptor interactions.[45]

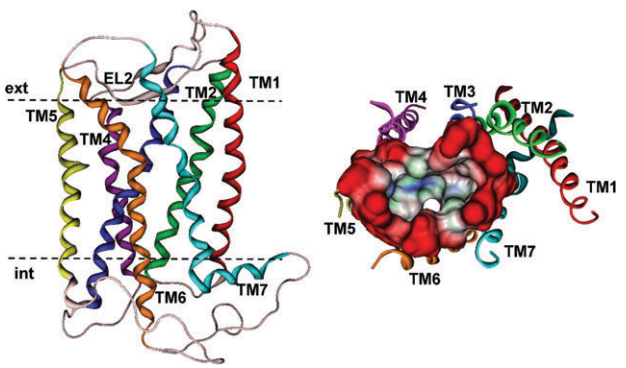

Fig. 6 *On the left*: general topology of the human A_3 adenosine receptor obtained using a rhodopsin-based homology modeling. *On the right*: molecular surface of the transmembrane antagonist binding cavity.

As already mentioned, the best approach to identify ligand–receptor interactions relies on the structure–activity relationship (SAR), supported by experimental data. Our approach for the identification of specific ligand–receptor interaction sites for a ligand and receptor pair of interest are based on correlating the effects of minor modifications of the ligand with mutations of the receptor. In the last five years, many investigations have been conducted for searching potent and selective human A_3 adenosine antagonists. In this field many different classes of compounds have been proposed, possessing high affinity (nM range) and with a broad range of selectivity. Using a high-throughput docking strategy, we have recently docked more than 350 known human A_3 antagonists in the corresponding putative ligand binding site. A consensus binding motif among all known antagonists has been found, and a novel Y-shaped 3D-pharmacophore model has been proposed.[20,42] Our pharmacophore model nicely fits inside the TMs region occupying the upper region of the helical bundle. His95 (TM3) and Ser247 (TM6) appear to be very critical for the recognition of the antagonist structures. In fact, a major structural difference between the hypothetical binding sites in these receptor subtypes is that the A_3 receptor does not contain the histidine residue in TM6 common to all A_1 (His251 in h_A_1) and A_{2A} (His250 in h_A_{2A}) receptors.[46,47] This histidine has been shown to participate in both agonist and antagonist binding to A_{2A} receptors. In the A_3 receptor this histidine in TM6 is replaced with a serine residue (Ser247 in h_A_3). Another strong hydrogen bonding interaction is possible with Asn250 (TM6). Also this asparagine residue, conserved among all adenosine receptor subtypes, was found to be important for ligand binding. An important and peculiar hydrophobic pocket delimited by non-polar amino acids, Leu90 (TM3), Leu246 (TM6), and Ile268 (TM7), is also addressable in our binding site model. Another important and highly conserved region, probably stabilized by π–π interactions, is located between Phe168 (EL2) and Phe182 (TM5), respectively. This region seems to be another important pharmacophore feature of our binding motif. Interestingly, the amino acids corresponding to Leu90 and Phe182 in the human A_{2A} receptor were found to be essential for the binding of both agonists and antagonists.[46,47] The last crucial pharmacophore region is mostly hydrophobic and characterized by three non-polar amino acids: Ile98 (TM3), Ile186 (TM5), Leu244 (TM6). Also in this case, the results were successfully interpreted using our rhodopsin-based model of ligand-A_3 receptor interactions. Generating a 3D pharmacophore for the receptor of interest enables the identification of key residues and physicochemical properties that are likely the molecular determinants for ligand recognition. Structure–activity relationships among available ligands for a given receptor provide equivalent pharmacophores at the ligand level. The parallel development of ligand-based and receptor-based 3D pharmacophores, developed independently from each other, enables a self-consistency check search for spatial complementarities between them based on favourable

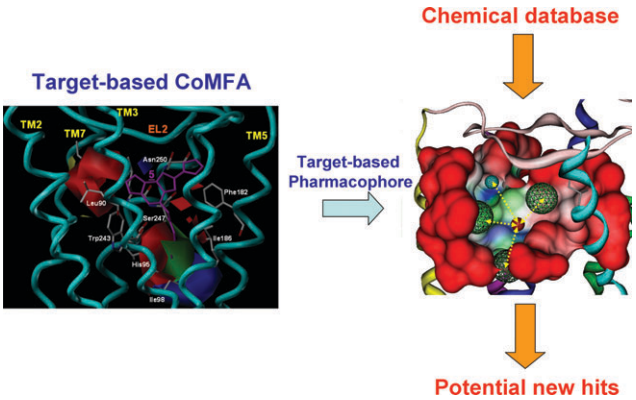

Fig. 7 *On the left:* side view of the best docked model of a very potent and selective pyrazolo-triazolo-pyrimidine derivative in complex with the human A_3 receptor. The side chains of some crucial important residues in proximity (≤ 5 Å) to the docked pyrazolo-triazolo-pyrimidine molecule are highlighted and labeled: Leu90 (TM3), His 95 (TM3); Phe182 (TM5), Ile186 (TM5); Trp243 (TM6); Ser247 (TM6), Asn250 (TM6), Ser271 (TM7), His272 (TM7), Ser275 (TM7). The steric and the electrostatic contour plots, obtained from the CoMFA analysis, are included into ligand binding cavity. The yellow and the blue polyhedra correspond to regions of the field that are predicted to decrease the A_3 receptor affinity, while the green and the red regions are predicted to increase binding affinity. *On the right*: side view of the human A_3 receptor TMs recognition region. The receptor-based pharmacophore model has been used as a filtering tool of chemical libraries in different virtual screening studies.

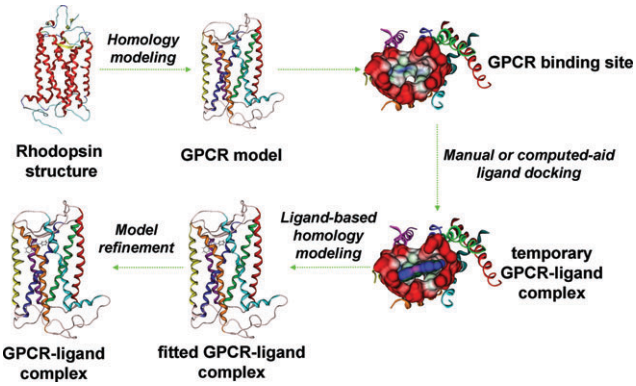

Fig. 8 Flow-chart of the ligand-based homology modeling technique considering the evolution of a conventional homology modeling algorithm.

physicochemical interactions. Currently, we are using this novel Y-shaped 3D-pharmacophore model to guide the selection of compound libraries (real and virtual libraries) for screening purposes as shown in Fig. 7.

We have also focused our attention on an alternative approach to computationally explore the multi-conformational space of the antagonist-like state of the human A_3 receptor. In fact, the specificity (complementarity) of the ligand–receptor recognition process has been regarded for a long time as a critical feature of the receptor concept, even when the nature of the receptors was entirely unknown. Following this assumption, it is reasonable to expect that both pharmacological agonist and antagonist-like states can be described by several different conformational receptor states depending on the nature of both ligand and receptor. Indeed, considering different ligand–receptor complementarities, we might explore different conformations of the same pharmacological state. We have developed a *ligand-based homology modeling* technique as a new approach to simulate the reorganization of the receptor induced by the ligand binding.[48] Briefly, the ligand-based homology modeling technique is an evolution of a conventional homology modeling algorithm based on a Boltzmann-weighted randomized modeling procedure adapted from Levitt.[49] *Ligand-based* tools are very useful when one wishes to build a homology model in the presence of a ligand docked to the primary template. In this specific case both model building and refinement take into account the presence of the ligand in terms of specific steric and chemical features. A schematic representation of the ligand-based homology protocol is shown in Fig. 8. Starting from the conventional rhodopsin-based homology model and applying our ligand-based homology modeling implementation we have generated another two antagonist-like conformational states of the human A_3 receptor in which the ligand recognition cavity has been expanded from 670 Å^3 to 770 and 1120 Å^3, respectively. Currently, we are using these multi-conformational states of human A_3 receptor to guide the selection of compound libraries (real and virtual libraries) for virtual screening purposes. We consider this strategy an alternative computational tool to explore the possible multi-conformational space associate with the antagonist-like state of the human A_3 receptor.

We have also recently proposed, an alternative strategy to theoretically estimate the binding affinity of different human A_3 receptor antagonists. We decided to combine a target-based approach, such as high-throughput docking, and a 3D ligand-based methodology to improve our capability to design new potent and selective human A_3 antagonists. We have used both Comparative Molecular Field Analysis (CoMFA),[42] and autocorrelated Molecular Electrostatic Potential (autoMEP) vectors in combination with the conventional Partial Least Square (PLS) analysis as "alternative" strategies to estimate the binding affinity of all known human A_3 antagonists.[50,51] CoMFA is a widely used methodology to develop molecular models which enable a better understanding of molecular activity. Typically, a database of molecules with known properties, the training set, is suitably aligned in 3D space according to various methodologies. Superimposition is one of the most crucial steps in CoMFA.[42] Once one has chosen the alignment, charges are calculated for each molecule at an appropriate level of theory. Steric and electrostatic fields are consequently derived for each molecule by the interaction with a probe atom on a series of grid points surrounding the aligned database in 3D space.[42] These field energy terms are then correlated with a property of interest by the use of partial least squares with cross-validation, giving a measure of the predictive power of the model.[42]

On the other hand, we have recently used the autocorrelation vectors as molecular descriptors for the PLS analysis.[50,51] Interestingly, autocorrelation vectors allow in fact to compare molecules (and their properties) with different structures and

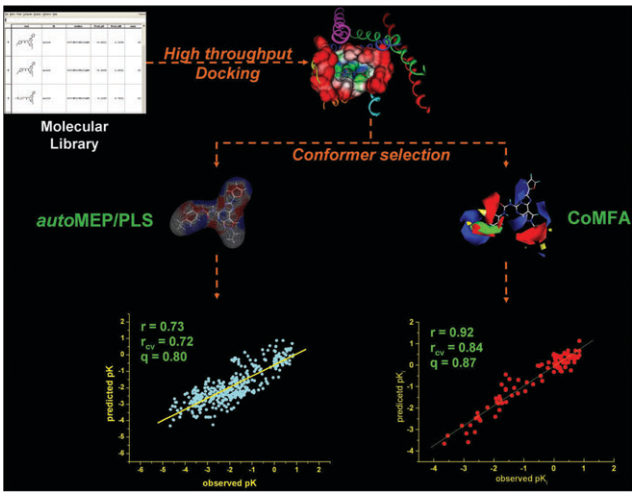

Fig. 9 Flow-chart describing our CoMFA and autoMEP/PLS 3D-QSAR approaches.

with different spatial orientation without any previous alignment. In particular, Molecular Electrostatic Potential (MEP), projected on Connolly's molecular surface, was computed and its information has been encoded in autocorrelation vectors. The 3D spatial distribution and the values of the electrostatic potential are in fact largely responsible for the binding of a substrate to its receptor binding site. In our studies, we decide to perform high-throughput docking to analyze the hypothetical binding conformation of more than three hundred human A_3 antagonists. After analyzing the docking results, we utilized the superposition of all low energy-docked conformations as structural input (super-molecule) for a classical 3D-CoMFA or for the calculation of the autoMEP vectors as input of a PLS analysis (Fig. 9). Using both strategies, we have obtained good quantitative models able to correctly predict all the binding affinities of already known human A_3 antagonists.

To validate our 3D-QSAR models, we have designed several new potential antagonists with estimated activities from the micromolar to the sub-nanomolar range. Impressively, all predicted *K*i were very closed to the experimental values.[42,50,51] This is a nice example in which an integrated approach between a very consolidated target-based approach, such as molecular docking, and a solid quantitative ligand-based methodology such as CoMFA or autoMEP/PLS can really help to better understand the chemical rules of ligand–receptor recognition from one side, and to speed up the discovery process of new potent and selective ligands on the other.

Perspective

GPCRs represent a fabulous assembly of nanomachines. The insightful understanding of their structures and functions is one of the most challenging objectives of all biosciences. However, several important questions will require additional structural information at higher resolution. From a drug discovery point of view, one of the most intriguing is: what is the precise structural basis of ligand specificity in a particular receptor, and how can the basic seven-helical structure be tuned to bind such a large and chemically-diverse spectrum of ligands? We do believe that the successful answer to these questions will ultimately result from a synergistic interaction between theory and experiment, and we consider our finding in adenosine receptor subtypes a nice example of the applicability of this integrated approach.

Acknowledgements

The molecular modeling work coordinated by S. Moro has been carried out with financial supports of the University of Padova, Padova (Italy) and the Italian Ministry for University and Research (MIUR), Rome, Italy.

References

1 R. J. Lefkowitz, *Trends Pharmacol. Sci.*, 2004, **25**, 413.
2 J. Bockaert and J. P. Pin, *EMBO J.*, 1999, **7**, 1723.
3 G. R. Post and J. H. Brown, *FASEB J.*, 1996, **10**, 741.
4 T. B. Patel, *Pharmacol. Rev.*, 2004, **56**, 371.
5 P. H. McDonald and R. J. Lefkowitz, *Cell. Signalling*, 2001, **13**, 683.
6 R. J. Lefkowitz and S. K. Shenoy, *Science*, 2005, **308**, 512.
7 B. Brone and J. Eggermont, *Am. J. Physiol. Cell Physiol.*, 2005, **288**, 20.
8 T. Metaye, H. Gibelin, R. Perdrisot and J. L. Kraimps, *Cell. Signal.*, 2005, **17**, 917.
9 G. Milligan, *Mol. Pharmacol.*, 2004, **66**, 1.
10 S. C. Prinster, C. Hague and R. A. Hall, *Pharmacol. Rev.*, 2005, **57**, 289.
11 M. Filizola and H. Weinstein, *Curr. Opin. Drug Discovery Dev.*, 2005, **8**, 577.
12 T. K. Attwood and J. B. Findlay, *Protein Eng.*, 1994, **7**, 195.
13 L. F. Kolakowski, *Recept. Channels*, 1994, **2**, 1.
14 J. V. Broeck, *Arch. Insect Biochem. Physiol.*, 2001, **48**, 1.
15 R. Fredriksson, M. C. Lagerstrom, L. G. Lundin and H. B. Schioth, *Mol. Pharmacol.*, 2003, **63**, 1256.
16 H. B. Schioth and R. Fredriksson, *Gen. Comp. Endocrinol.*, 2005, **142**, 94.
17 R. E. Stenkamp, D. C. Teller and K. Palczewski, *Arch. Pharm. (Weinheim, Ger.)*, 2005, **338**, 209.
18 C. R. McCudden, M. D. Hains, R. J. Kimple, D. P. Siderovski and F. S. Willard, *Cell. Mol. Life Sci.*, 2005, **62**, 551.
19 B. K. Rana, T. Shiina and P. A. Insel, *Annu. Rev. Pharmacol. Toxicol.*, 2001, **41**, 593.
20 M. D. Thompson, W. M. Burnham and D. E. Cole, *Crit. Rev. Clin. Lab. Sci.*, 2005, **42**, 311.
21 D. P. Behan and D. T. Chalmers, *Curr. Opin. Drug Discovery Dev.*, 2001, **4**, 548.
22 T. Klabunde and G. Hessler, *ChemBioChem*, 2002, **3**, 928.
23 S. Moro, F. Deflorian, G. Spalluto and K. A. Jacobsob, *Chem. Commun.*, 2003, **21**, 2949.
24 L. T. May, V. A. Avlani, P. M. Sexton and A. Christopoulos, *Curr. Pharm. Des.*, 2004, **10**, 2003.
25 J. Presland, *Curr. Opin. Drug Discovery Dev.*, 2005, **8**, 567.
26 O. M. Becker, S. Shacham, Y. Marantz and S. Noiman, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 11304.
27 K. Kristiansen, *Pharmacol. Ther.*, 2004, **103**, 21.
28 F. Fanelli and P. G. De Benedetti, *Chem. Rev.*, 2005, **105**, 3297.
29 S. Moro, G. Spalluto and K. A. Jacobson, *Trends Pharmacol. Sci.*, 2005, **26**, 44.
30 K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto and M. Miyano, *Science*, 2000, **289**, 739.
31 L. Hunyady, G. Vauquelin and P. Vanderheyden, *Trends Pharmacol. Sci.*, 2003, **24**, 81.
32 B. B. Fredholm, A. P. IJzerman, K. A. Jacobson, K. N. Klotz and J. Linden, *Pharmacol. Rev.*, 2001, **53**, 527.
33 P. Fishman and S. Bar-Yehuda, *Curr. Top. Med. Chem.*, 2003, **3**, 463.

34 M. Parsons, L. Young, J. E. Lee, K. A. Jacobson and B. T. Liang, *FASEB J.*, 2000, **14**, 1423.
35 S. Merighi, A. Benini, P. Mirandola, S. Gessi, K. Varani, E. Leung, S. Maclennan and P. A. Borea, *J. Biol. Chem.*, 2005, **280**, 19516.
36 T. M. Palmer and G. L. Stiles, *Mol. Pharmacol.*, 2000, **57**, 539.
37 S. Rorke and S. T. Holgate, *Am. J. Respir. Med.*, 2002, **1**, 99.
38 K. A. Jacobson, *Trends Pharmacol. Sci.*, 1998, **19**, 184.
39 T. W. Stone, *Adv. Exp. Med. Biol.*, 2002, **513**, 249.
40 J. Linden, *Annu. Rev. Pharmacol. Toxicol.*, 2001, **41**, 775.
41 C. E. Muller, *Curr. Top. Med. Chem.*, 2003, **3**, 445.
42 S. Moro, P. Braiuca, F. Deflorian, C. Ferrari, G. Pastorin, B. Cacciari, P. G. Baraldi, K. Varani, P. A. Borea and G. Spalluto, *J. Med. Chem.*, 2005, **48**, 152.
43 M. E. Olah, K. A. Jacobson and G. L. Stiles, *J. Biol. Chem.*, 1994, **269**, 24692.
44 Z. G. Gao, J. E. Van Muijlwijk-Koezen, A. Chen, C. E. Muller, A. P. Ijzerman and K. A. Jacobson, *Mol. Pharmacol.*, 2001, **60**, 1057.
45 Z. G. Gao, A. Chen, D. Barak, S. K. Kim, C. E. Muller and K. A. Jacobson, *J. Biol. Chem.*, 2002, **277**, 19056.
46 Q. Jiang, B. X. Lee, M. Glashofer, A. M. van Rhee and K. A. Jacobson, *J. Med. Chem.*, 1997, **40**, 2588.
47 Q. Jiang, A. M. Van Rhee, J. Kim, S. Yehle, J. Wess and K. A. Jacobson, *Mol. Pharmacol.*, 1996, **50**, 512.
48 S. Moro, M. Bacilieri, F. Deflorian and G. Spalluto, *Curr. Pharm. Des.*, 2006, in press.
49 M. Levitt, *J. Mol. Biol.*, 1992, **226**, 507.
50 S. Moro, M. Bacilieri, C. Ferrari and G. Spalluto, *Curr. Drug Discovery Technol.*, 2005, **2**, 13.
51 S. Moro, M. Bacilieri, B. Cacciari and G. Spalluto, *J. Med. Chem.*, 2005, **48**, 5698.

