UNIVERSITÀ DI PISA

Dottorato di ricerca in *Scienza del farmaco e delle sostanze bioattive*

Computational tools for the study of the structure-property relationship and design of new biologically active compounds.

Relatore

Chiar.mo Prof. Adriano Martinelli

Dottorando

Dott. Tiziano Tuccinardi

CHIM/08

2003

Index

2

Preface

The aim of this PhD course was to explore a broad overview on the topic of the Structure-Property Relationship (SPR) with a strong emphasis on the pratical aspects.

The basic principle the study of SPR is based on is that all the chemical, physicochemical, biological or other functional properties of a molecule are dependent on its structure. Usually the reasons for undertaking a SPR study are either the characterization of the properties of a molecule (interpretation or razionalization phase) or the attempt to design a new structure in order to obtain a desidered property (prediction phase). The first choice is, in any case, a necessary step for obtaining the second one.

The interest in razionalizing and optimizing the process of SPR stems from considerations of cost-effectiveness. Since the costs of chemical research are very high, the need for making the process of understanding the relationships between structure and function more efficient has a very high priority in the strategy of all chemical companies. A classical example is the one of drug research: attempting to design a new drug by trial and error is obviously a very inefficient approach (even though in some forms this is a method still in use), while knowing which are the pharmacophoric elements of a basic molecular structure allows for a contained number of specific modifications. This, in turn is more likely to provide the desired properties and a great saving in time and resources.

Therefore it is of critical importance to utilize all the theoretical methods that will allow the access or the extraction of all possible information hidden in the individual facts about structure and properties of molecule. In this view I could say that the common denominator of this PhD course is in fact information extraction and information manipulation.

Data in chemical research, and in particular in drug discovery, is varied and oftentimes very complex. In drug discovery one has to make sense of different type of data such as structural, biological, physico-chemical, pharmacological, toxicological and so on, which, ultimately have to be associated to a single molecular structure. In order to sort

out these data and extract appropriate information, a number of tools have been devised on computers and workstations in the form of different programs; the reader will find that many of these tools and methods have been used during this PhD course.

More in details in Chapter 1 the homology modeling of the adenosine receptors was explored and accompanied to the pharmacophoric analysis and synthesis of new compounds.

In Chapter 2 the analysis of the MMP-inhibitor interaction led us to implement the Amber Forcefield, and the following docking analysis allowed the design of new selective inhibitors.

The modeling of the activate form of the cannabinoid receptors (Chapter 3) corresponded to an attempt for going away from the homology modeling procedures; together with the goal of obtaining a quantitative model from an automated docking study.

In Chapter 4 the study of ligand-estrogen receptor interaction was developed exploring the free energy calculation, while finally in the last Chapter the angiotensin receptor AT1 construction led us to propose a new binding orientation for the non-peptide antagonists, using the 3D-QSAR approach as validation and predictive method.

CHAPTER 1

ADENOSINE RECEPTORS

1.1 MODELING OF THE A$_1$ AND A$_{2A}$ HUMAN ADENOSINE RECEPTORS:

1.1.1 Introduction.

Adenosine receptors are members of the superfamily of G protein-coupled receptors (GPCR). As such, they are single polypeptide chains possessing seven hydrophobic transmembrane-spanning segments that couple to an effector molecule through a trimeric G protein complex.

To date, four adenosine receptor subtypes have been cloned; these include the A$_1$ and A$_3$ receptors, which inhibit adenylyl cyclase, and the A$_{2a}$ and A$_{2b}$ receptors, which stimulate adenylyl cyclase.[1]

In the present study, we focused our attention on the interaction of ligands with the A$_1$ and A$_{2a}$ARs; the A$_1$AR was initially cloned from the thyroid gland of dogs,[2] and was later isolated from several other species.[3,4] The A$_1$AR is a protein of 326 aminoacids, which has a high affinity for N^6-substituted adenosine analogues. Several highly selective A$_1$AR compounds are available, including the agonist N^6-cyclopentyladenosine (CPA).[5] The carboxyl terminus of the A$_1$AR is shorter than that of the A$_{2a}$AR, whereas the amino terminus is longer. At the aminoacid level, the A$_1$ and A$_{2a}$ARs are generally 60% identical within the transmembrane domains.[3,4,6] Unlike the A$_1$AR, the A$_{2a}$AR has a high affinity for 5'-substituted adenosine agonists and a low affinity for N^6-substituted compounds.[7] Highly selective A$_{2a}$AR agonists are available, and include the compound CGS-21680.[8]

These receptors are important pharmacological targets in the treatment of a variety of conditions such as asthma, neurodegenerative disorders, psychosis and anxiety, chronic inflammatory diseases and many other physiopathological states that are believed to be associated with changes in adenosine levels.[9-11]

In particular the development of agonists for the adenosine A$_1$AR receptor, able to mimic the central inhibitory effects of adenosine (and so inhibiting neurotransmitter release), may therefore be clinically useful as neuroprotective agents. On the contrary A$_1$AR selective antagonists have been developed as antihypertensives and potassium-

saving diuretics, cognition enhancers and useful therapeutics for the alleviation of the symptoms of Alzheimer's disease.[12,13]

Furthermore A$_{2A}$AR agonists are potentially useful for the treatment of cardiovascular diseases, such as hypertension, ischemic cardiomyopathy, inflammation, and atherosclerosis,[14,15] and A$_{2A}$AR antagonists have been proposed as novel therapeutics for Parkinson's disease and may also be active as cognition enhancers, neuroprotective and antiallergic agents, analgesics, and positive inotropics.[16-18]

Consequently, selective and potent agonists or antagonists at the adenosine receptor subtypes are needed for therapeutic intervention; however, a clear picture of how these receptors bind their various ligands has not emerged yet. A knowledge of the 3D structure of adenosine receptors could be of great help in the task of understanding their function and in the rational design of specific ligands. However since GPCRs are membrane-bound proteins, high-resolution structural characterization is still an extremely difficult task. For this reason, great importance has been placed on molecular modelling studies, and, in particular in the last few years, on homology modelling techniques. The publication of the first high-resolution crystal structure for rhodopsin,[19] a GPCR superfamily member, provides the option of homology modelling to generate 3D models based on detailed structural information. With the aim of achieving a better understanding of experimental results, in the present study we constructed the A$_1$AR and A$_{2a}$AR three-dimensional model of the seven α-helical transmembrane domains, using bovine rhodopsin as a template .

To test these models, we carried out the docking of certain A$_1$ and A$_{2a}$AR selective agonists.

1.1.2 Computational methods.

All the calculations were carried out by means of the Batchmin programme of the MACROMODEL suite,[20] using the Amber forcefield and making the dielectric "distance-dependent" constant equal to 4.0. The molecular dynamics (MD) simulations were performed at 300° K, with a timestep of 1.0 fs and an equilibration time of 40 ps, while all the minimizations were carried out by means of 2000 Steepest Descent steps, followed by Conjugate Gradient until a convergence of 0.05 kJ/Å•mol.

The graphic manipulations and visualizations were performed by means of the

Maestro,[20] WebLabViewer,[21] and Chimera[22] programs, while the quantum mechanical calculations were carried out using the Gaussian03 program.[23]

1.1.2.1 Amino acid numbering. To refer to specific amino acids sequences, the numbering system suggested by Ballesteros and Weinstein was used.[24] The most highly conserved residue in each transmembrane helix (TMH) was assigned a locant value of 0.50, and this number was preceded by the TMH number and followed in parentheses by the sequence number. The other residues in the helix were given a locant value relative to this.

1.1.2.2 Molecular modeling All the information regarding the primary structure of human A_1 and A_{2a} receptors, and their subdivision into transmembrane, cytoplasmatic and extracellular domains, was obtained from the GPCR Data Bank.[25] We modelled only the TM domains, since the function of the loops has still not been defined. Although site-directed mutagenesis suggests a role for adenosine receptor loops, and in particular for the second extracellular (E2) ones, it remains unclear whether the E2 loop is in direct contact with ligands, or whether it contributes to the overall physical architecture of the receptor protein.[26-28] After defining the primary structure, the secondary and tertiary ones were defined by using the structure of bovine rhodopsin as a template.[19] The receptor-template superimposition was carried out maintaining the maximum analogy between the receptors, and choosing the regions with a conserved or semi-conserved sequence. The alignment was studied on several adenosine receptors by means of the ClustalW program[29] using the blosum algorithm, with a gap open penalty of 10 and a gap extension penalty of 0.05. As shown in Figure 1, the alignment was guided by the highly conserved amino acid residues, including the D/ERY motif (D/E3.49, R3.50, and Y3.51), the two proline residues P4.50 and P6.50 and the NPXXY motif in TM7 (N7.49, P7.50, and Y7.53).[30] We used the 3D X-ray crystallographic structure of bovine rhodopsin registered in PDB (1F88) as a direct template to construct the 7 TM helical structure of the A_1AR, using the Maestro programme to substitute the residues of rhodopsin not conserved in the receptor; subsequently, each model helix was capped with an acetyl group at the N-terminus and an N-methyl group at the C-terminus. In order to consider the mutagenesis data regarding TM3, it was necessary to rotate the third helix of the receptor by approximately 90°.

```
                                      TM1
AA1R_HUMAN   -------MPPSISAFQAAYIGIEVLIALVSVPGNVLVIWAVKVNQALRDATFCFIVSLAV
AA1R_MOUSE   -------MPPYISAFQAAYIGIEVLIALVSVPGNVLVIWAVKVNQALRDATFCFIVSLAV
AA2AR_HUMAN  ----------MPIMGSSVYITVELAIAVLAILGNVLVCWAVWLNSNLQNVTNYFVVSLAA
AA2AR_MOUSE  -------------MGSSVYIMVELAIAVLAILGNVLVCWAVWINSNLQNVTNFFVVSLAA
AA2BR_HUMAN  ---------MLLETQDALYVALELVIAALSVAGNVLVCAAVGTANTLQTPTNYFLVSLAA
AA2BR_MOUSE  ---------MQLETQDALYVALELVIAALAVAGNVLVCAAVGASSALQTPTNYFLVSLAT
AA3R_HUMAN   ----MPNNSTALSLANVTYITMEIFIGLCAIVGNVLVICVVKLNPSLQTTTFYFIVSLAL
AA3R_MOUSE   ---MEADNTTETDWLNITYITMEAAIGLCAVVGNMLVIWVVKLNPTLRTTTVYFIVSLAL
OPSD_BOVIN   SPFEAPQYYLAEPWQFSMLAAYMFLLIMLGFPINFLTLYVTVQHKKLRTPLNYILLNLAV
                 :    ..  *.*.  ..    *:   ..    :::.**

                           TM2                        TM3
AA1R_HUMAN   ADVAVGALVIPLAILINIG--PQTYFHTCLMVACPVLILTQSSILALLAIAVDRYLRVKI
AA1R_MOUSE   ADVAVGALVIPLAILINIG--PQTYFHTCLMVACPVLILTQSSILALLAIAVDRYLRVKI
AA2AR_HUMAN  ADIAVGVLAIPFAITISTG--FCAACHGCLFIACFVLVLTQSSIFSLLAIAIDRYIAIRI
AA2AR_MOUSE  ADIAVGVLAIPFAITISTG--FCAACHGCLFIACFVLVLTQSSIFSLLAIAIDRYIAIRI
AA2BR_HUMAN  ADVAVGLFAIPFAITISLG--FCTDFYGCLFLACFVLVLTQSSIFSLLAVAVDRYLAICV
AA2BR_MOUSE  ADVAVGLFAIPFAITISLG--FCTDFHGCLFLACFVLVLTQSSIFSLLAVAVDRYLAIRV
AA3R_HUMAN   ADIAVGVLVMPLAIVVSLG--ITIHFYSCLFMTCLLLIFTHASIMSLLAIAVDRYLRVKL
AA3R_MOUSE   ADIAVGVLVIPLAIAVSLQ--VKMHFYACLFMSCVLLIFTHASIMSLLAIAVHRYLRVKL
OPSD_BOVIN   ADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGFFATLGGEIALWSLVVLAIERYVVVCK
             **:: :   :.  ::   .      * :     :   .  :: :*:.:*:.**: :

                                   TM4
AA1R_HUMAN   PLRYKMVVTPRRAAVAIAGCWILSFVVGLTPMFGWNN----LSAVERAWAANGSMGEPVI
AA1R_MOUSE   PLRYKTVVTQRRAAVAIAGCWILSLVVGLTPMFGWNN----LSEVEQAWIANGSVGEPVI
AA2AR_HUMAN  PLRYNGLVTGTRAKGIIAICWVLSFAIGLTPMLGWNN----CGQPKEGKNHSQGCGEGQV
AA2AR_MOUSE  PLRYNGLVTGMKAKGIIAICWVLSFAIGLTPMLGWNN----CSQTDE--NSTKTCGEGRV
AA2BR_HUMAN  PLRYKSLVTGTRARGVIAVLWVLAFGIGLTPFLGWNSKDSATNNCTEPWDGTTNESCCLV
AA2BR_MOUSE  PLRYKGLVTGTRARGIIAVLWVLAFGIGLTPFLGWNSKDSATSNCTELGDGIANKSCCPL
AA3R_HUMAN   TVRYKRVTTHRRIWLALGLCWLVSFLVGLTPMFGWNN----------MKLTSEYHRNVTFL
AA3R_MOUSE   TVRYRTVTTQRRIWLFLGLCWLVSFLVGLTPMFGWNN----------RKATLASSQNSSTL
OPSD_BOVIN   PMSNFRFG-ENHAIMGVAFTWVMALACAAPPLVGWSR----------YIPEGMQCSCGID
             .:     :      :    :.   *:::: . .*:.**.

                               TM5
AA1R_HUMAN   KCEFEKVISMEYMVYFNFFVWVLPPLLLMVLIYLEVFYLIRKQLNKKVSASSG--DPQKY
AA1R_MOUSE   KCEFEKVISMEYMVYFNFFVWVLPPLLLMVLIYLEVFYLIRKQLNKKVSASSG--DPQKY
AA2AR_HUMAN  ACLFEDVVPMNYMVYFNFFACVLVPLLLMLGVYLRIFLAARRQLKQMESQPLPGERARST
AA2AR_MOUSE  TCLFEDVVPMNYMVYYNFFAFVLLPLLLMLAIYLRIFLAARRQLKQMESQPLPGERTRST
AA2BR_HUMAN  KCLFENVVPMSYMVYFNFFGCVLPPLLIMLVIYIKIFLVACRQLQRTELMDHS----RTT
AA2BR_MOUSE  TCLFENVVPMSYMVYFNFFGCVLPPLLIMLVIYIKIFMVACKQLQSMELMDHS----RTT
AA3R_HUMAN   SCQFVSVMRMDYMVYFSFLTWIFIPLVVMCAIYLDIFYIIRNKLSLNLSNSK---ETGAF
AA3R_MOUSE   LCHFRSVVSLDYMVFFSFITWILVPLVVMCIIYLDIFYIIRNKLSQNLTGFR---ETRAF
OPSD_BOVIN   YYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTVKEAAAQQQ-------ESATT
             .:*  :   *.   .:  **:::   *   :.

                             TM6                          TM7
AA1R_HUMAN   YGKELKIAKSLALILFLFALSWLPLHILNCITLFCPS--CHKPSILTYIAIFLTHGNSAM
AA1R_MOUSE   YGKELKIAKSLALILFLFALSWLPLHILNCITLFCPT--CQKPSILIYIAIFLTHGNSAM
AA2AR_HUMAN  LQKEVHAAKSLAIIVGLFALCWLPLHIINCFTFFCPD--CSHAPLWLMYLAIVLSHTNSVV
AA2AR_MOUSE  LQKEVHAAKSLAIIVGLFALCWLPLHIINCFTFFCST--CQHAPPWLMYLAIILSHSNSVV
AA2BR_HUMAN  LQREIHAAKSLAMIVGIFALCWLPVHAVNCVTLFQPAQGKNKPKWAMNMAILLSHANSVV
AA2BR_MOUSE  LQREIHAAKSLAMIVGIFALCWLPVHAINCITLFHPALAKDKPKWVMNVAILLSHANSMM
AA3R_HUMAN   YGREFKTAKSLFLVLFLFALSWLPLSIINCIIYFN----GEVPQLVLYMGILLSHANSMM
AA3R_MOUSE   YGREFKTAKSLFLVLFLFALSWLPLSIINFVSYFD----VKIPDVAMCLGILLSHANSMM
OPSD_BOVIN   QKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQG--SDFGPIFMTIPAFFAKTSAVY
               * .  :: :  :::  * :.***    :         :         : .:::  :

AA1R_HUMAN   NPIVYAFRIQKFRVTFLKIWNDHFRCQPAPPIDEDLPEERPDD-------
AA1R_MOUSE   NPIVYAFRIHKFRVTFLKIWNDHFRCQPKPPIEEDIPEEKAED-------
AA2AR_HUMAN  NPFIYAYRIREFRQTFRKIIRSHVLRQQEPFKAAGTSARVLAAHGSDGEQ
AA2AR_MOUSE  NPFIYAYRIREFRQTFRKIIRTHVLRRQEPFRAGGSSAWALAAHSTEGEQ
AA2BR_HUMAN  NPIVYAYRNRDFRYTFHKIISRYLLCQADVKSGNGQAGVQPALGVGL---
AA2BR_MOUSE  NPIVYAYRNRDFRYSFHKIISRYVLCQAETKGGSGQAGAQSTLSLGL---
AA3R_HUMAN   NPIVYAYKIKKFKETYLLILKACVVCHPSDSLDTSIEKNSE---------
AA3R_MOUSE   NPIVYACKIKKFKETYFLILRAVRLCQTSDSLDSNMEQTTE---------
OPSD_BOVIN   NPVIYIMMNKQFRNCMVTTLCCGKNPLGDDEASTTVSKTETSQVAPA---
             **.:*    :.:*
```
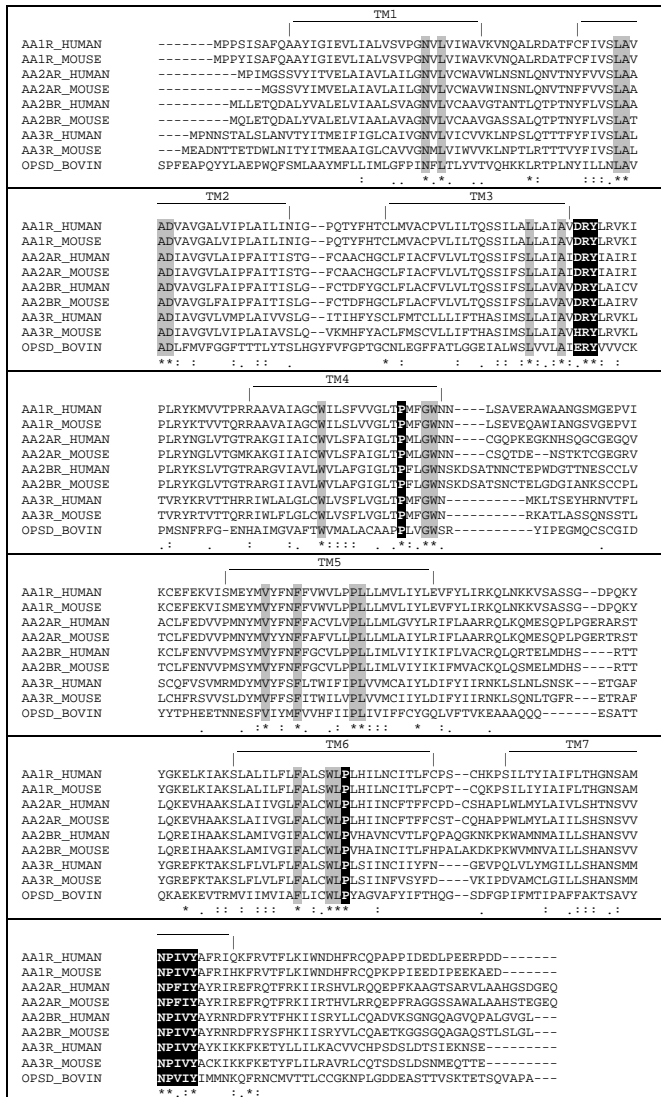
Figure 1. Alignment of the adenosine receptors and bovine rhodopsin (OPSD_BOVIN) amino acid sequences. The highly conserved patterns of the D/ERY motif (D/E3.49, R3.50, and Y3.51), P4.50 and P6.50 and the NPXXY motif (N7.49, P7.50, and Y7.53) are marked with black. The other identical residues are indicated with "*" and marked in grey in the TMs, while the conservatively replaceable residues are indicated with ":" and ".".

The A_1AR was subjected to a preliminary minimization and 200 ps of MD, after which the final structure was minimized. When MD simulations are carried out in the gas phases, skipping the explicit environment requires the use of a set of restraints, to replace the natural stabilizing effects of the membrane bilayer on the TM domains. Accordingly, restraints with a force constant of 10 Kcal/mol•Å^2 where applied to backbone for the first 100 ps, and for the remaining 100 ps, these restraints were reduced to 1 Kcal/mol•Å^2.

The A_1AR model obtained by means of these calculations was complexed with a high affinity ligand, and the complex was optimized.

CPA was chosen for this purpose, since it is commonly used in binding experiments as a radioligand;[5] it was docked into the receptor, bearing in mind the known mutagenesis data.

The CPA geometry was optimized at the AM1 level, and the atomic charges were calculated using the RESP[31] method with the 6-31G* wave function.

The CPA was then manually introduced into the binding site in such a manner as to give H bond interactions with T3.36(91), S3.39(94), T7.42(277), and H7.43(248) and a lipophilic interaction (through the cyclopentyl moiety) with L3.33(88), in accordance with the main mutagenesis data (see Table 1).

The complex modelling was performed by means of a total of 800 ps of molecular dynamics. In order to consider the stabilizing presence of the membrane around the receptor, all the alpha carbons of the protein and the intra-helix H bonds were blocked during modelling by means of decreasing force constants; moreover also the main ligand-receptor interactions (shown in Table 2) were restrained. More specifically, an initial restraint with a force constant of 10 Kcal/mol•Å^2 was applied on all the alpha carbons, this force constant decreased during the whole MD, and in the last 200 ps, its value was 0.1 Kcal/mol•Å^2. As regards the intra-helix H bonds and the main ligand-receptor interactions, a restraint of 10 Kcal/mol•Å^2 and 50 Kcal/mol•Å^2 was applied.

Every 200 ps of MD simulation, the conformation with respect to which the alpha-carbon restraints were defined was updated; this fact, together with the intra-helix H bond constraints, allowed us to take into account the effects of the non-conserved prolines on the helix conformation.

Table 1. Mutational analysis for A$_1$ and A$_{2a}$AR agonist interaction.

	Gen. Numb.	A$_1$AR	Mutational results	A$_2$AR	Mutational results
TM1	1.37	G14	T: increased affinity[32]	T11	
	1.39	E16	A: reduced 4- to 40-fold[33]	E13	Q: slight reduction[34]
	1.48	P25	L: modest reduction[32]	L22	
	1.54	I31	C: no variation[32]	C28	1.54
TM2	2.41	C46	A/S: no variation[35]	Y43	
	2.45	S50	A: no variation[32]	S47	
	2.50	D55	A: increased affinity[33]	D52	
	2.60	L65	F: no variation[32]	F62	
TM3	3.25	C80	A/S: no detectable binding[35]	C77	
	3.27	M82	F: no variation[35]	F79	
	3.30	C85	A/S: reduced 4- to 13- fold[35]	C82	
	3.31	P86	F: reduction of affinity[32]	F83	
	3.32	V87	A: no variation[32]	V84	L: marginal variation[36]
	3.33	L88	A: reduction of affinity[32*]	L85	
	3.36	T91	A: reduction of affinity[32]	T88	A: reduction of affinity[37]
	3.37	Q92	A: reduction of affinity[32]	Q89	A: increased affinity[37]
	3.38	S93	A: no variation[33]	S90	A: marginal variation[38]
	3.39	S94	A: no detectable binding[33]	S91	A: marginal variation[37]
TM4	4.49	C131	A: no variation[38]	C128	
	4.53	S135	A: no variation[38]	S132	
	4.59	T141	A: no variation[38]	T138	
	4.62	F144	L: no variation[38]	L141	
L4-5		R154		E151	A: loss of affinity[27]
		C169	A: no detectable binding[35]	C166	
		E172		E169	A: loss of affinity[27]
		K173		D170	K: no variation[27]
		S176		P173	R: no variation[27]
TM5	5.41	F184		F180	A: marginal variation[40]
	5.42	N185		N181	S: reduction of affinity[40]
	5.43	F186		F182	A: loss of affinity[40]
TM6	6.52	H251	A: no variation[39]	H250	L: loss of affinity[39]
	6.55	N254		N253	A: loss of affinity[40]
	6.56	C255	A: no variation[35]	C254	A: marginal variation[40]
	6.59	L258		F257	A: loss of affinity[40]
L6-7		C260	A: no variation[35]	C259	
		C263	A: no variation[35]	C262	G: no variation[35]
TM7	7.39	I274		I274	A: loss of affinity[40]
	7.42	T277	A: reduction of affinity[41]	S277	A: reduction of affinity[40]
	7.43	H278	L: loss of affinity[42]	H278	A: loss of affinity[40]
	7.46	S281		S281	A: loss of affinity[40]

* When L3.33(88) was converted to Alanine, the binding of ligands with N6 substituents was reduced to a greater extent than ligands without N6 substituents.

After the MD, three steps of minimization were applied to the structure obtained as the average of the last 100 ps. During all these steps, a restraint of 0.1 Kcal/mol•Å^2 was applied to the alpha carbons, and the restraint on the intra-helix H bonds was removed, while as regards the main ligand-receptor interaction in the first two steps, a restraint of

30 and 10 Kcal/mol•Å^2 was applied, respectively, and in the last step the restraint was removed. The same procedure was applied to several different starting interaction geometries, with the aim of exploring other binding possibilities, but at the end of the modelling procedure, only the above-described one shown in Figure 3a maintained all the interactions considered important by mutagenesis studies (see Table 1).

With the aim of validating this model, 300 ps of MD simulation were performed, in which the backbone of the receptor was fixed, but all the ligand-receptor restraints were removed; a sampling of the last 200 ps conformation showed that during the simulation, none of the main ligand-receptor interactions were missed (the variation of the interatomic distances corresponding to these interactions was always less than 20%). Furthermore, the average of the last 200 ps was minimized without constraints, and the superimposition between the initial and the final CPA conformation showed an RMSD of 0.96 Å. The backbone conformation was also evaluated by inspection of the Psi/Phi Ramachandran plot obtained from PROCHECK analysis.[43]

As shown in the Ramachandran plot in Figure 2, the distribution of the Psi/Phi angles of the model is within the allowed regions and no residues have disallowed conformations.
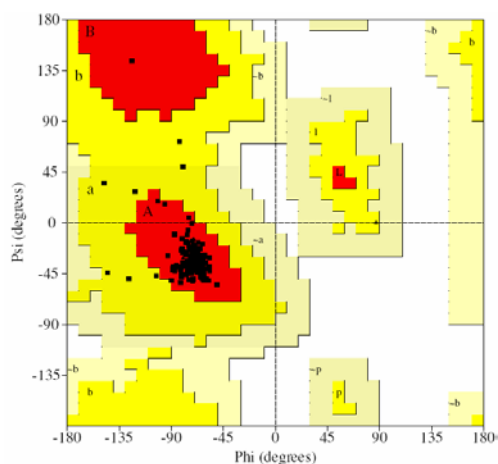
Figure 2. Ramachandran plot of the A₁AR. The most favoured regions are coloured red, additional allowed, generously allowed and disallowed regions are indicated as yellow, light yellow and white fields, respectively.

In order to obtain the A$_{2a}$AR model we applied the same procedure described above, using two starting structures as templates: bovine rhodopsin and the A$_1$AR model obtained after the first 200 ps of MD. As the two results differ very slightly, we preferred to use the second model, because in this way, the manual rotation of the TM3 was identical in both receptors.

The obtained A$_{2a}$AR model was complexed with a high affinity ligand, and the complex was optimized. The complex modelling was carried out by using CGS-21680,[8] a selective agonist, considering the interactions with N6.55(253), S7.42(277), H7.43(278) and S7.46(281), suggested by mutagenesis data (see Table 1).

Also for CGS-21680 some other starting binding positions were considered but, like CPA with the A$_1$AR, at the end of the modelling procedure, none of them maintained all the interactions considered important by mutagenesis data.

At this point the docking of the A$_1$AR selective agonists CPA, RPIA, CADO, the non-selective agonist NECA and the A$_{2a}$AR selective agonists CGS-21680 was performed manually in both receptors. All the compounds tested present the adenine group as their central core, and the initial docking position of the ligands was obtained by superimposing this group on those of the final structure of CPA and CGS-21680 in the A$_1$ and A$_{2a}$AR, respectively. In this position, all ligands exhibit the interaction suggested by mutagenesis data (see Table 2).

The ligand geometry was optimized at the AM1 level, and the atomic charges were calculated using the RESP[31] method with the 6-31G* wave function. To model the various ligand-receptor complexes, 800 ps of molecular dynamics were applied in the same conditions described above, using a constraint scheme analogous to the one used for A$_1$ and A$_{2a}$AR complex modelling.

Table 2. Ligands used to perform the docking and their main interactions with the two receptors.
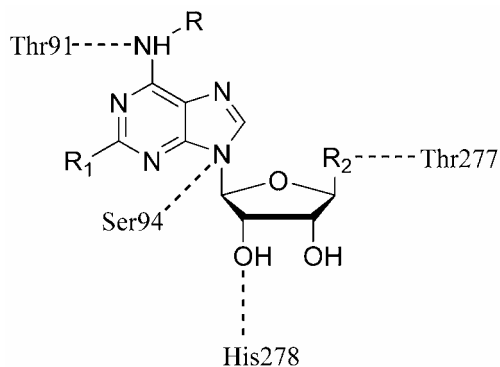
	R	R1	R2
CPA		H	CH$_2$-OH
RPIA		H	CH$_2$-OH
CADO	H	Cl	CH$_2$-OH
NECA	H	H	
CGS21680[a]	H		

[a]CGS21680 presents a further interaction at the level of the carboxylic group with Asn181

1.1.3 Results and discussion

In general, docking of agonists to GPC receptors is subject to greater uncertainty than antagonists, as rhodopsin is crystallized in its inactive state. Until now, there is only a rough picture of the conformational changes that occur during receptor activation. Recent studies suggest that receptor activation could be due to a different rearrangement of TM3 and TM6.[44] Furthermore, on the basis of UV absorption analysis, it has been suggested that when rhodopsin is activated, the χ1 rotamer of the high conserved residue

W6.48(265) shifts from gauche+ to trans;[45] and recently, a 3D model for meta-II rhodopsin was published, featuring a similar change to the conformation of W6.48.[46] Interestingly, this rotamer switch was also confirmed in the present study: during the MD simulation of both the A_1 and the $A_{2a}AR$ complexed with all the agonists, the χ1 rotamer of W6.48 spontaneously shifted from gauche+ to trans. Therefore, for our purposes, although we do not wish to neglect other dynamic features of the GPCR structure, we prefer, in the absence of a crystal structure of a representative activated GPCR, not to change the template on the basis of hypothesized structures that may turn out to be inaccurate.

Table 3 shows the residues involved in the first and second spheres of the binding sites for the ligand-receptor A_1 and A_{2a} complexes studied. For both receptors, the binding site is positioned between TM3, TM6 and TM7, and all the residues considered important by mutagenesis studies are in the first sphere of the binding site.

Table 3. Residues involved in binding sites of A_1 and A_{2a} receptors. Residues in the first sphere are involved in the binding site with a distance lower than 4 Å, the second sphere includes the residues involved in the binding site with a distance between 4 and 6 Å. Residues that appear from mutagenesis to be of crucial importance are indicated in bold, while non-conserved residues are highlighted in grey.

Receptor	Residues first sphere	Residues second sphere
A_1AR	**L3.33(88)**, L3.35(90), **T3.36(91)**, **Q3.37(92)**, **S3.39(94)**, I3.40(95), A3.42(97), L3.43(98), F5.43(186), F6.44(243), S6.47(246), W6.48(247), L6.51(250), N6.55(254), I7.39(274), **T7.42(277)**, **H7.43(278)**, S7.46(281)	L2.46(51), V3.32(87), V4.56(138), G4.57(139), V5.39(182), Y5.40(183), N5.42(185), W5.46(189), V5.47(190), H6.52(251), L6.59(258), A7.47(282)
$A_{2a}AR$	A2.49(51), V2.53(55), V3.32(84), L3.35(87), **T3.36(88)**, S3.39(91), V5.39(178), **N5.42(181)**, **F5.43(182)**, L6.51(249), **H6.52(250)**, **N6.55(253)**, **F6.59(257)**, **I7.39(274)**, **S7.42(277)**, **H7.43(278)**, S7.46(281)	L2.46(48), D2.50(52), L3.33(85), Q3.37(89), I3.40(92), L3.43(95), P4.60(139), W6.48(246), I6.54(252), C6.56(254), T6.58(256), M7.35(270), N7.45(280)

Table 4 shows the principal interactions of the different compounds inside the two receptors.

Table 4. Principal interactions between the ligands and the binding site residues of the two receptors. The distances (Å) of hydrogen bond (HB) and lipophilic (LIPO) interactions are reported together with the group of the ligand giving the interaction (Aden = adenine ring; NH = N substituent of the adenine ring; R[X] = X group of the R substituent; see Table 3). Distances greater than 5 A for HB and 7A for LIPO were not considered.

Residues	CPA		CADO		RPIA		NECA		CGS-21680	
	A₁AR	A₂AR	A₁AR	A₂AR	A₁AR	A₂AR	A₁AR	A₂AR	A₁AR	A₂AR
HB										
T3.36	NH 2.94	NH 3.00	NH₂ 2.93	NH₂ 2.95	NH 3.29	NH 3.03	NH₂ 3.38	NH₂ 2.87	NH₂ 2.93	NH₂ 3.07
S3.39	Ribose 2.83	Aden 2.90	Aden 3.10	Aden 2.90	Aden 3.33	Ribose 2.85	Ribose 2.95	Aden 2.93	R[NH] 2.93	Ribose 3.52
N5.42	----	----	----	----	----	----	----	----	R₁[COO]2.9	R₁[COO]2.8
N6.55	----	Aden 4.14	----	Aden 4.06	----	NH 4.08	----	NH₂ 3.13	R₁[NH] 3.04	R₁[CO] 2.93
T/S7.42	R₃[OH] 3.03	R₃[OH] 2.79	R₃[OH] 2.84	R₃[OH] 2.86	R₃[OH] 2.78	R₃[OH] 2.85	R₃[CO] 2.80	R₃[CO] 2.82	Ribose 3.22	Ribose 3.16
H7.43	Ribose 3.04	Ribose 2.86	Ribose 2.89	Ribose 3.05	Ribose 3.04	Ribose 3.01	Ribose 3.05	Ribose 3.09	Ribose 2.95	Ribose 3.26
S7.46	Ribose 3.37	Ribose 4.03	----	----	Ribose 4.10	R₃[OH] 3.78	Ribose 4.09	R₃[NH] 3.38	R₃[CO] 2.74	R₃CO 2.77
LIPO										
L3.33	R[cicle] 3.58	R[cicle] 6.48	----	----	R[Phe] 4.20	R[Phe] 6.06	----	----	----	R₁[Phe] 4.55
F5.43	R[cicle] 5.59	----	Aden 5.17	----	R[Phe] 3.97	----	Aden 6.74	----	R₁[Phe] 4.84	R₁[Phe] 4.23
F6.44	Ribose 3.64	----	Aden 3.12	----	Ribose 3.73	----	Aden 4.81	R₃[CH₃] 3.87	Ribose 4.19	----
W6.48	Aden 3.89	----	Aden 3.35	----	Aden 3.71	----	----	----	R₁[Phe] 4.84	R₁[Phe] 6.90
L6.51	Aden 3.57	Aden 4.49	Aden 4.36	Aden 4.43	Aden 3.60	Aden 5.16	Aden 4.54	Aden 3.98	Ribose 4.27	Aden 5.70
H6.52	----	Aden 6.33	----	Aden 6.20	----	----	----	Aden 6.45	R₁[Phe] 5.72	R₁[Phe] 3.40
L/F6.59	R[cicle] 6.04	R[cicle] 5.83	----	----	R[Phe] 3.89	R[CH₃] 4.04	----	----	R₁[Phe] 5.08	R₁[CH₃]3.95
I7.39	Aden 4.86	Aden 5.24	Ribose 4.68	Ribose 5.63	Ribose 5.09	Ribose 4.57	Aden 4.34	Ribose 5.21	Aden 3.75	Aden 4.31

17

Figure 3a shows the docking, in the A$_1$ and A$_{2a}$AR binding sites, of CPA, the most A$_1$AR selective ligand among those tested. In the A$_1$ subtype, according to the mutagenesis data, CPA gives H bonds with T3.36(91), S3.39(94), T7.42(277), H7.43(248) and lipophilic interactions with L3.33(88) (through the cyclopentyl moiety) and with F5.43(186), F6.44(243), W6.48(247) and L6.51(250) (through the adenosinic group). In the A$_{2a}$ subtype, the H bonds are maintained, but the lipophilic interactions with L3.33(85), F6.44(242), W6.48(246) and L6.51(249) are absent, because of the different position assumed by CPA in this receptor.

The other two A$_1$AR selective agonists, RPIA and CADO, interact with both receptors in the same manner as CPA. They present in both receptors the H bonds with T3.36(91), S3.39(94), T7.42(277), and H7.43(248) and show only in the A$_1$ subtype lipophilic interactions with F5.43(186), F6.44(243), W6.48(247). NECA, instead, which has a similar activity in the two receptors, possesses, besides the H bonds showed by the other ligands, a further H bond in the A$_{2a}$AR with N6.55(253), which is suggested by mutagenesis studies to be of great importance for this receptor subtype. Furthermore, this agonist shows a lipophilic interaction with L6.51 and F6.44 in both the A$_1$AR and the A$_{2a}$AR.

Figure 3b illustrates CGS-21680 docked into both sites of the A$_1$AR (on the left) and the A$_{2a}$AR (on the right); it also shows that the ligand interacts with T3.36, N5.42, N6.55, S7.42, H7.43 and S7.46 through H bonds. In the A$_{2a}$AR, the aromatic substituent is stabilized by F5.43(182) and H6.52(250) residues (at 4.2Å and 3.4Å from the aromatic ring), whereas in the A$_1$AR, these residues are far away (4.8Å and 5.7Å from the aromatic ring), and are unable to interact with CGS-21680.
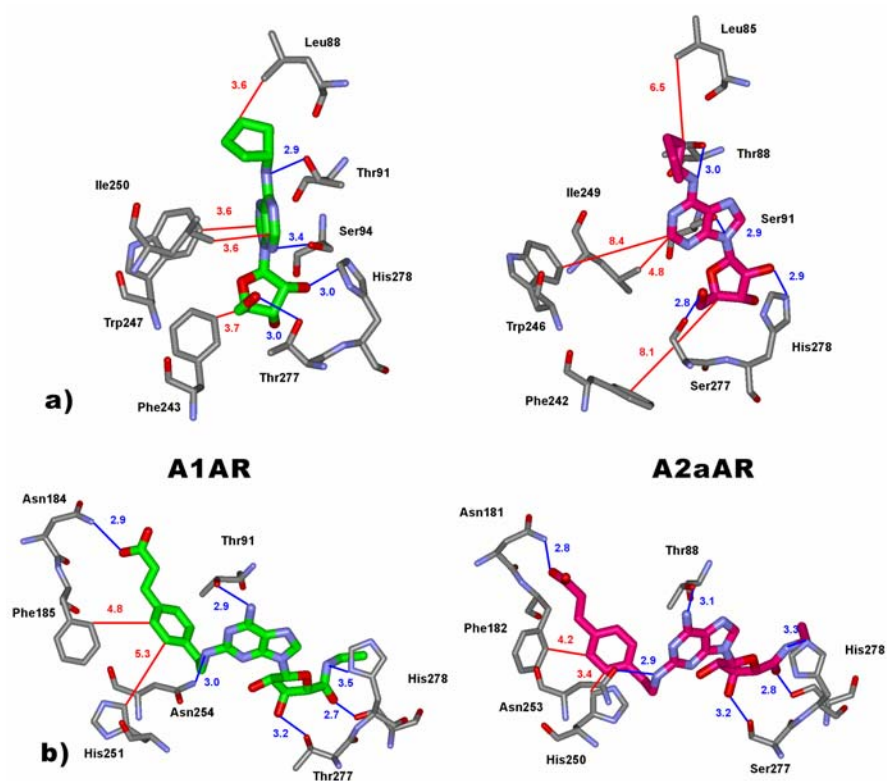
Figure 3. a. CPA docked into the A1AR (left) and A2aAR (right) binding sites. **b.** CGS-21680 docked into the A$_1$AR (left) and A$_{2a}$AR (right) binding sites. Interatomic distances between H-bonded atoms are indicated in blue; carbon-carbon distances showing lipophilic interactions are indicated in red. All distances are in Angstroms.

Figure 4a shows the volumes of the cavities between CPA and the two ARs, thus indicating the different dimensions of the two binding sites, which could be important in determining the selectivity of the derivatives considered.

The A$_1$AR binding site cavity is small and allows close interaction with the ligand, while in the A$_{2a}$AR , less bulky ligands like CPA cannot occupy the larger site so efficiently.

As regards the binding of CGS-21680, Figure 4b confirms that in the A$_{2a}$AR, the aromatic substituent is effectively stabilized, whereas in the A$_1$AR, the large dimensions of the pocket where the substituent is inserted do not allow a strong interaction.
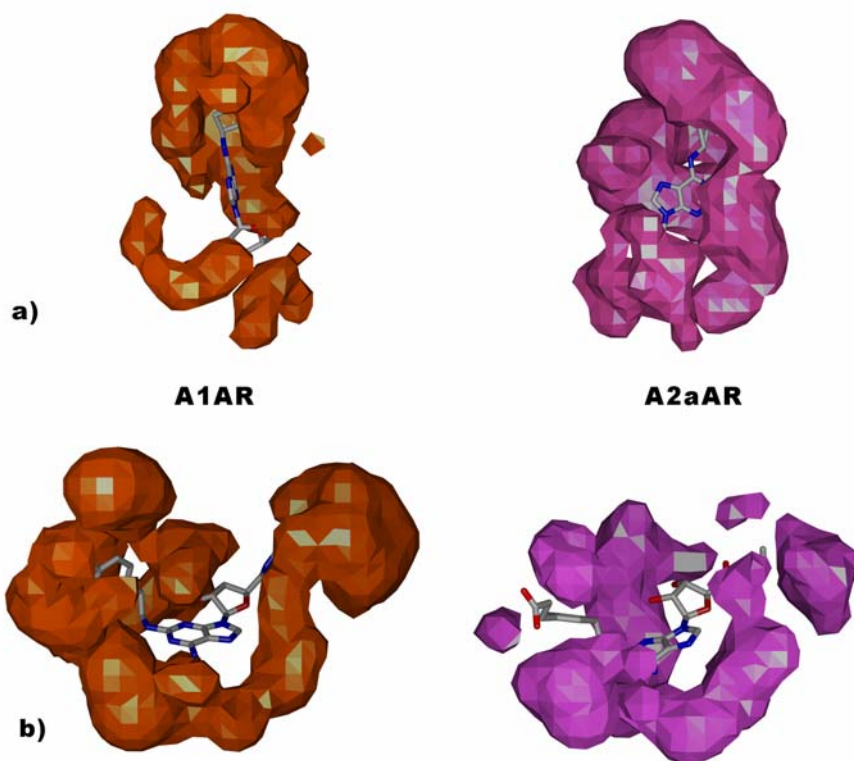
Figure 4. a. CPA docked into the A₁AR (left) and A₂ₐAR (right) binding sites. The volumes of the cavities between CPA and the two receptors are shown. **b.** CGS-21680 docked into the A₁AR (left) and A₂ₐAR (right) binding sites. The volumes of the cavities between CGS-21680 and the two receptors are shown.

All these observations are confirmed by the energy calculations of the ligand-receptor interactions obtained by means of the Batchmin program[20] on the final structures of the complexes (see Table 5).

The interaction energies between the receptor and the ligand were calculated by subtracting the energy of the separate ligand and receptor from the energy of the receptor-ligand complex. These energies are not rigorous thermodynamic quantities, but can only be used to compare the relative stability of the complexes of the same ligand in different receptors. Consequently, these interaction energy values cannot be used to predict binding affinities, since changes in entropy and solvation effects are not taken into account.

Table 5. Binding affinity (Ki values) and interaction energy (kcal/mol) of selective agonists for the A$_1$AR and the A$_{2a}$AR.[38]

Ligand	A$_1$ receptor				A$_{2a}$ receptor			
	Ki (nM)	Interaction Energy (Kcal/mol)			Ki (nM)	Interaction Energy (Kcal/mol)		
		VdW	Elect.	Total		VdW	Elect.	Total
CPA	2.3	-29.1	-2.65	-31.75	790	-22.63	-4.12	-26.76
CADO	1.39)	-28.45	-3.59	-32.04	180	-20.38	-3.44	-23.82
RPIA	2.0)	-31.66	-2.45	-34.11	860	-29.50	-2.51	-32.01
NECA	14	-23.45	-2.72	-26.17	20	-25.33	-2.56	-27.90
CGS-21680	290	-43.10	-6.93	-50.03	27	-46.48	-7.16	-53.64

These calculations underline the fact that A$_1$/A$_{2a}$ selectivity is mainly influenced by the different ability of the two receptors to give lipophilic interactions, instead of giving different H bonds. This would confirm the hypothesis that A$_1$AR-selective ligands exhibit a lower level of interaction in the A$_{2a}$AR due to the larger dimensions of the binding site cavity, which does not allow a strong lipophilic interaction. Analogously, in the case of CGS-21680, the low activity shown in the A$_1$AR could be due to the low stabilization of the aromatic substituent inside the lipophilic pocket.

As shown in Table 3, there are only few non-conserved residues among the A$_1$ and A$_{2a}$ binding sites, and they cannot justify the different binding cavity size. However, on comparing the two receptors, we observe a different disposition of the TMs: this could be due to the fact that in the A$_1$AR, the residues P1.48(25), P3.31(86) and P5.49(92) act as flexible molecular hinges, changing the folding of the helices of TM1, TM3 and TM5; in the A$_{2a}$AR, these residues are not conserved, thus determining a different rearrangement of the helices, with the formation of a large binding site cavity. Site-directed mutagenesis shows that in the A$_1$ receptor, the substitution of the P1.48(25) and P3.31(86) with Leucine and Phenylalanine respectively, causes a reduction in agonist affinity (see Table 1), suggesting their structural role.

The different disposition of the helices is confirmed by an analysis of the H bonds that spontaneously form during the MD of the complexes and are present at the end of the modelling procedure in all the ligand-receptor complexes (see Table 6). This

rearrangement in the A$_{2a}$AR concerns only TM2, TM3 and TM5, leaving the binding site free.

Table 6. Inter-helix H-bonds in the A$_1$and A$_{2a}$AR found after molecular modelling optimization.

Inter-helix bonds in the A$_1$ receptor	Inter-helix bonds in the A$_{2a}$ receptor
S2.45(50)-W4.50(132)	E1.39(13)-Y7.36(271)
I3.34(89)-S4.53(135)	F2.42(47)-A3.45(97)
I3.46(101)-Y7.53(288)	Q3.37(89)-N5.42(181)
Y5.40(183)- H6.52(251)	A3.47(99)-Y5.58(197)
Y5.58(201)-L6.41(240)	
I6.40(239)-N7.45(280)	

The S2.45(50)-W4.50(132) H bond is lacking in the CADO and CGS21680-A$_1$AR complex, while the A3.47(99)-Y5.58(197) in the CGS21680-A$_{2a}$AR complex is replaced by Y5.58(197)-V6.41(39).

In the A$_1$AR, instead, H bonds are generated, involving and re-arranging all the TM domains, and turning the side chains into the inter-helix space. In particular the interactions of L6.41(240) with Y5.58(201) and of H6.52(251) with Y5.40(183), drag F6.44(243), W6.48(247) and L6.51(250) towards the binding cavity, allowing an effective stabilization of the adenosinic group of the agonists.

1.1.4 Conclusions.

We have provided 3D models of the A$_1$ and A$_{2a}$ adenosine receptors, based on the highest resolution structure of bovine rhodopsin. A model of the agonist-receptor complexes was constructed and validated by means of docking studies. The structural effects of ligand binding have been examined on the basis of hydrogen bonds, lipophilic interactions and binding energies in the final complexes obtained from manual docking. Results show that A$_1$/A$_{2a}$ selectivity is not mainly influenced by a different H bond network between ligand and receptor, as the selective A$_1$ agonists present the same H interactions in both receptors. What appears to be decisive is the lipophilic factor: in the A$_1$ subtype, the compounds with a high affinity present lipophilic interactions with

L3.33(88), F5.43(186), F6.44(243), W6.48(247) and L6.51(250). Among them, the only residue that has been tested by mutagenesis in the A_1AR is L3.33(88). In our model, this aminoacid shows a strong interaction with the N6-substituted ligands (CPA and RPIA), and this is in agreement with the mutagenesis studies that highlight an important interaction with these compounds.[32]

As regards the A_{2a} affinity, our studies have confirmed that the interaction with N6.55(253) is crucial, moreover, the docking of CGS21680 shows that the selectivity of this compound could be due to the presence of the R1 substituent, which is able to interact with the lipophilic residues of TM5 and TM6.

The different lipophilic interaction in the two receptors seems to be mediated by the different dimensions of the inter-helix channel, due to the structural diversity of the two receptors and to the consequent diversity of the inter-helix H-bond network. In particular, three non-conserved Prolines, P1.48(25), P3.31(86) and P5.49(92), allow a different interaction in the A_1 receptor between the TMs, thus determining a smaller binding site cavity. These differences allow ligands with large substituents, like CGS-21680, to interact well with the $A_{2a}AR$; in the A_1AR, instead, the small dimensions of the binding site allows a good interaction only with small ligands, like adenosine.

1.2 MOLECULAR MODELING STUDY OF THE INTERACTION OF THE 1,8-NAPHTHYRIDINE COMPOUNDS WITH THE BOVINE AND HUMAN A_1 ADENOSINE RECEPTORS.

Ferrarini PL[*], Betti L, Cavallini T, Giannaccini G, Lucacchini A, Manera C, Martinelli A, Ortore G, Saccomanni G, Tuccinardi T. J. Med. Chem. 47 (2004) 3019-3031

1.2.1 Introduction.

Adenosine receptors from different species show a good amino acid sequence homology (82-93 %), the only exception is the A3 subtype, which only exhibits 74 % primary sequence homology between rat and human or sheep.[47-49]

Although there is only little difference in the A_1 receptor sequence of different species,[50] some species differences in agonist binding have been reported.[51]

The bovine A_1 receptor has an affinity for agonist and antagonist ligands that is 10-fold higher than that of rat and human receptors; in the bovine receptor, the typical A_1 receptor rank order of potency R-PIA > NECA > S-PIA is partially altered in that it has a specifically reduced binding affinity for the 5'-substituted adenosine analogues compared with rat and human receptors.[52]

Furthermore, N^6-substituted adenosine derivatives, such as R-PIA, are more potent at bovine than at human or rat A_1 receptors. This phenomenon has been called the "phenyl effect", and is the strongest at the bovine A1 receptor.[51]

A new series of 1,8-naphthyridine derivatives (see Table 1) bearing various substituents in different positions of the heterocyclic nucleus were synthesized in order to analyze the effects produced on the affinity towards the bovine and human adenosine receptors. The results indicate that all the 1,8-naphthyridine compounds generally possess a higher affinity towards the bovine A_1 receptor compared with the human A_1 receptor.

A molecular modeling study of the docking of the 1,8-naphthyridine compounds with both the bovine and the human A_1 adenosine receptors was carried out with the aim of explaining the marked decrease in the affinity towards human A_1 adenosine receptors in comparison with bovine A_1 adenosine receptors.

1.2.2 Results and Discussion.

The model of the hA1AR previously obtained (see Chapter 1.1) was optimized so as to interact suitably with the specific ligand DPCPX, on the basis of the available site directed mutagenesis data,[28,32,33,35,53] which suggested that residues Ser94 and His251 were fundamental for the affinity of the antagonists. Therefore, the starting geometries of the complexes between the hA$_1$AR and the the antagonist DPCPX were arranged in such a manner that the ligand could favorably interact with these two residues.

The computational procedures are the same fully described in the previous chapter.

The model of the bA$_1$AR was constructed on the basis of the hA$_1$AR model, following a similar procedure. Figure 1 illustrates DPCPX docking into both the site of the hA$_1$AR (on the left) and the bA$_1$AR (on the right) and shows that DPCPX interacts at a similar distance with Ser94 and His251 through its hydroxylic functions by means of H-bonds which appear to be shorter and therefore stronger in the hA$_1$AR (d=2.9 and 3.0 Å) than the bA$_1$AR (d=3.1 Å). However, a series of lipophilic interactions due to Leu90, Ile95, Leu250, Ala273 and Ile274 are able to better stabilize DPCPX in the bA$_1$AR: the cyclophentyl group interacts with Ala273 in hA1AR (d = 4.1 Å) and Ile274 (d=4.4 Å), while it interacts with Ala273 (d = 4.0 Å) and Leu90 (d = 4.0 Å) in bA$_1$AR; moreover, in bA1AR the two n-propyl chains are able to interact with Ile95 (d = 3.9 Å) and Leu250 (d = 4.0 Å), while in hA1AR there is only the interaction with Ile95 (d = 4.3 Å).
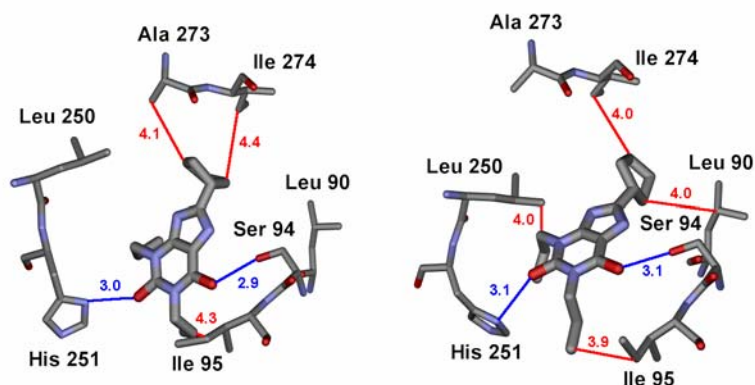
Figure 1. DPCPX docked into the hA$_1$AR (left) and bA$_1$AR (right) binding site. Interatomic distances between H-bonded atoms are reported in blue; lipophilic interactions are reported in red (.Angstroms).

Table 1. Affinity of 1,8-Naphthyridine Derivatives in Radioligand Binding Assays at Bovine Brain A_1, A_{2A} and Human Brain A_1 and A_{2A} Receptors. [a,b]

compd	R₁	R₂	R₃	K_i (nM)					
				bA₁	bA₂A	bA₂A/bA₁	hA₁	hA₂A	hA₂A/hA₁
1	Ph	OH	CH₃	5.3 ± 0.6	460 ± 38	87	430 ± 35	550 ± 47	1.3
2			CH₃	450 ± 92	> 10000	>22	>10000	>10000	
3	Ph	SH	CH₃	80 ± 7	1250 ± 230	16	2000 ± 150	230 ± 20	0.12
4	Ph	Cl	CH₃	1000 ± 110	>10000	>10	>10000	>10000	
5	Ph	OCH₃	CH₃	6000 ± 600	>10000	>2	>10000	>10000	
6	Ph	OPh	CH₃	810 ± 87	>10000	>12	8800 ± 850	>10000	>1.1
7	Ph	N₃	CH₃	>10000	>10000		>10000	>10000	
8	Ph	NH₂	CH₃	17 ± 4	420 ± 29	25	2000 ± 170	210 ± 15	0.10
9	Ph	N(CH₃)₂	CH₃	550 ± 57	7000 ± 630	13	>10000	>10000	
10	Ph	cyclohexylNH	CH₃	28 ± 3	830 ± 79	30	560 ± 52	490 ± 45	0.88
11	Ph	Cep [c]	CH₃	1200 ± 130	>10000	>8	>10000	>10000	
12	Ph	Pipz [d]	CH₃	8700 ± 750	>10000	>1	>10000	>10000	
13	p-FPh	OH	CH₃	5.3 ± 0.4	630 ± 65	119	3900 ± 400	440 ± 45	0.11
14	o-FPh	OH	CH₃	11 ± 4	470 ± 42	43	490 ± 45	40 ± 5	0.081
15	Ph	OH	Br	0.70 ± 0.05	66 ± 5	94	360 ± 35	100 ± 9	0.28
16	Ph	OH	Cl	0.15 ± 0.01	100 ± 15	670	300 ± 27	450 ± 42	1.5
17	Ph	OH	F	4.1 ± 0.6	170 ± 35	41	1400 ± 150	2700 ± 25	1.9
18	Ph	OH	OPh	16 ± 4	>10000	>625	1300 ± 120	2000 ± 180	1.5
19	Ph	OH	OPh	26 ± 4	1900 ± 170	73	>10000	4800 ± 450	< 0.48
20	Ph	OH	OEt	5.2 ± 0.7	>10000	>1920	2600 ± 270	1500 ± 130	0.58
21	Ph	OH	OCH₃	1.6 ± 0.2	1400 ± 140	875	2000 ± 180	490 ± 50	0.25
22	Ph	OH	OH	>10000	>10000		>10000	>10000	
23				1300 ± 280	>10000	>8	>10000	>10000	
24				4900 ± 340	>10000	>2	>10000	6500 ± 620	< 0.65
25	Ph		CH₃	6900 ± 700	>10000	>1	>10000	>10000	
26	p-NO₂Ph	OH	CH₃	9.9 ± 0.8	460 ± 37	46	3100 ± 280	640 ± 57	0.21
27	Ph	NHNH₂	CH₃	100 ± 11	1800 ± 150	18	>10000	4100 ± 400	< 0.41
28	Ph	OH	NH₂	5.3 ± 0.5	5900 ± 420	1110	3400 ± 350	9000 ± 870	2.6
29	Ph	OH	N(CH₃)₂	0.56 ± 0.03	2300 ± 210	4110	980 ± 90	>10000	>1.3
30	Ph	OH	Pip [e]	7.2 ± 0.7	>10000	>1390	7500 ± 720	>10000	>1.3
31	Ph	OH	cyclohexylNH	70 ± 6	2500 ± 230	36	2200 ± 200	1700 ± 160	0.77
32	Ph	OH	Morph [f]	160 ± 15	>10000	>63	>10000	>10000	
33	Ph	OH	Cep [c]	>10000	>10000		>10000	>10000	
34	Ph	OH	Pipz [d]	>10000	>10000		>10000	>10000	
35	Ph	OEt	Br	50 ± 4	2900 ± 250	58	1000 ± 110	970 ± 90	0.97
36				4700 ± 450	7200 ± 650	1.5	>10000	3800 ± 370	< 0.38
37			OEt	48 ± 5	>10000	208	930 ± 90	1400 ± 150	1.5
38	Ph	OEt	OEt	290 ± 30	>10000	>34	>10000	>10000	
39	Ph	OH	CH₃CONH	340 ± 35	>10000	>29	>10000	>10000	
40	Ph	SCH₃	CH₃	380 ± 35	>10000	>26	>10000	210 ± 20	< 0.021
41	p-NH₂C₆H₅	OH	CH₃	6.8 ± 0.5	500 ± 47	73	910 ± 90	110 ± 10	0.12
42	p-AcNHC₆H₅	OH	CH₃	1.0 ± 0.2	2300 ± 220	2300	2000 ± 200	1600 ± 150	0.8
43	m-NO₂C₆H₅	OH	CH₃	15 ± 4	580 ± 60	39	2200 ± 200	3400 ± 350	1.5
44	m-NH₂C₆H₅	OH	CH₃	23 ± 3	230 ± 20	10	4000 ± 380	1300 ± 110	0.32
46	CH₂Ph	OH	CH₃	1900 ± 200	>10000	>5.3	>10000	>10000	
47	H	OH	CH₃	7000 ± 710	>10000	>1.4	>10000	1800 ± 170	< 0.18
48	H	NH₂	CH₃	>10000	>10000		>10000	>10000	
49	CH₃	OH	OH	>10000	>10000		>10000	>10000	
50	CH₃	OH	NH₂	>10000	>10000		>10000	7500 ± 730	< 0.75
51	CH₂CH₂CH₃	OH	CH₃	48 ± 6	900 ± 85	19	1500 ± 140	1000 ± 110	0.66
52	CH₂CH₂CH₃	OH	NH₂	210 ± 20	>10000	>48	4100 ± 400	>10000	>2.4
DPCPX				0.25 ± 0.02	200 ± 17	800	6.7 ± 0.5	144	21.5
SCH58261				357 ± 35	2.0 ± 0.1	0.0056	463 ± 45	3.25 ± 0.2	0.0070

[a] Inhibition of specific [³H]CHA binding to bovine and human brain cortical membranes expressed as K_i ± SEM (n = 3) in nM.

[b] Inhibition of specific [³H]CGS21680 binding to bovine and human striatal membranes expressed as K_i ± SEM (n = 3) in nM.

[c] Cep = ethylcarbethoxypiperazinyl; [d] Pipz = piperazinyl; [e] Pip = piperidinyl; [f] Morph = morpholinyl.

Figure 2 shows that the fairly limited structural differences that exist in the transmembranal regions of the two receptors hA₁AR and bA₁AR (only seven residues) are, however, able to induce a clear difference in the 3D arrangement of the seven helices in the two models (the RMSD calculated on the backbone is 3.79 Å). An important point in determining the conformational differences in the models of the hA₁AR and the bA₁AR could be the replacement of the Met82 residue of the hA₁AR with Lys82 in the bA₁AR, which induces a different arrangement of the interhelix H-bonds and therefore helices TM2 and TM3 are closer in the case of the hA₁AR. The RMSD between the alpha carbons of helix 2 and the corresponding alpha carbons of helix 3 is 9.8 Å in the case of the hA₁AR and 10.5 Å in the case of the bA₁AR.

The different arrangement of these two helices also modifies the rest of the structure and, as a consequence, the binding pocket appears to be narrower in the bA₁AR than in the hA₁AR and this fact could be the reason of the interaction differences shown in Figure 1.
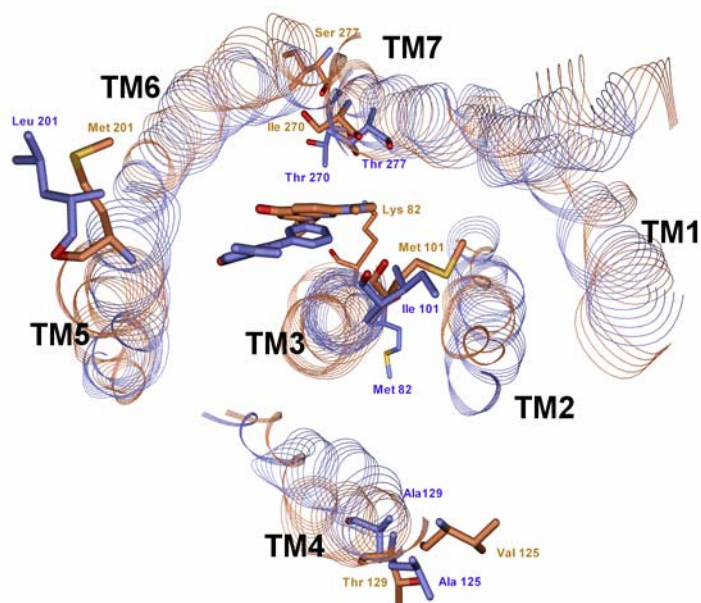
Figure 2. Superimposition of the complexes of DPCPX with hA₁AR (blue) and bA₁AR (orange); the seven not conserved residues of the seven transmembranal helices are shown.

Some 1,8-naphthyridine derivatives were then docked into the two receptor models; the compounds selected for this purpose were **16, 22, 28, 29, 50, 51** and **52**, which possessed different K_i values towards the hA_1AR and the bA_1AR.

The docking procedure was carried out by taking into account the site directed mutagenesis data, as for DPCPX. Therefore the selected compounds were initially placed in the receptor sites so that they could interact favourably with Ser94 and His251.

The groups of compounds considered capable of interacting at the same time with these two residues were the substituent in position 4 and the nitrogen in position 8, which possess suitable chemical characteristics and a suitable spatial arrangement. Therefore the two interaction geometries A and B, shown in Figure 3, are to be taken into consideration: in the first one, A, the substituent in position 4 gives an H bond with Ser94 and the nitrogen in position 8 gives an H-bond with His251; in the other one, B, the same groups of the selected compounds give H-bonds with the same residues, but in an inverse manner.
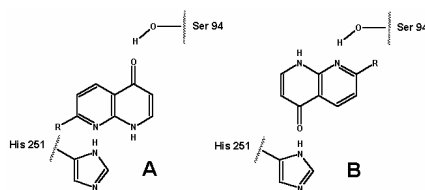
Figure 3. The two possible geometries for the interaction of 1,8-naphthyridine derivatives with A₁ adenosine receptors.

The complexes thus obtained (two for each compound and for each receptor model) were then optimized by means of molecular dynamic simulations followed by energy minimization with the AMBER force field.

The antagonist-receptor interaction energies were calculated as the sum of the non-bonded terms of the molecular mechanics steric energy referring to the interaction between the atoms of the receptor model and the atoms of the ligand.

All active compounds prefer arrangement B when they are fitted into both the hA_1AR and the bA_1AR. The differences in the interaction energy between the two

arrangements, A and B, range from 2 to 6 kcal/mol.

The relative values of the interaction energy are reported in Table 2 and indicate that all active ligands have a more favorable interaction with the model of the bA$_1$AR and this is in agreement with the greater affinity of the considered ligands towards the bA$_1$AR with respect to the hA$_1$AR; a correlation between these values and the selectivity ratio can also be observed.

Table 2. Interaction energies (Kcal/mol) of hA$_1$AR and bA$_1$AR models and selected ligands.

Compd	A$_1$bAR[a]	conf.[b]	A$_1$hAR[a]	conf.[b]	select.[c]	ΔE[d]
DPCPX	-30.2	-	-26.7	-	27	3.5
16	-29.7	B	-24.0	B	2000	5.7
22	-27.4	B	-25.2	B	-	-
28	-27.3	B	-22.2	B	640	5.1
29	-30.7	B	-26.3	B	1750	4.4
50	-21.1	B	-24.1	B	-	-
51	-22.6	A	-21.0	B	31	1.6
52	-23.0	B	-21.9	B	20	1.1

[a]Sum of the non-bonded terms of the interaction between the atoms of the receptor model and the atoms of the ligand; [b]arrangement energetically preferred by the ligand for interaction with the receptor model (see Fig. 4); [c]selectivity of the ligand computed as the ratio between the K$_i$ for hA$_1$AR and K$_i$ for bA$_1$AR; the selectivity is not reported in the case of inactive compounds; [d]energy difference between the interaction energy of the ligand with hA$_1$AR minus the interaction energy with bA$_1$AR; the energy difference is not reported in the case of inactive compounds.

In Figure 4 the complex between compound **16**, the most selective one, and the two models of the hA$_1$AR and bA$_1$AR receptors is shown. Figure 4 also reports the volumes of the cavities between the ligand and the receptors, and is thus able to visualize the difference in the dimensions of the two binding sites, which could be quite important in determining the selectivity of the 1,8-naphthyridine derivatives considered.

The compounds considered here are less bulky than classic A$_1$AR antagonists, and therefore they can interact more strongly with the sterically restricted site in the bA$_1$AR, while they cannot occupy so efficiently the larger site of the hA$_1$AR, as already indicated by the interaction energy values reported in Table 2.

In Figure 5 the details of the interaction of compounds **16** and **29** with the human and bovine A$_1$AR are presented. In the case of the complex with the bA$_1$AR, these antagonists are able to give H-bonds with Ser94 and His251 and at the same time their phenyl substituent in position 2 is able to give a hydrophobic interaction with Ile270

and other residues that make up a pocket (Tyr271, Ala273, Ile274). The pocket in the bA$_1$AR is able to optimally accept the phenyl ring linked in position 2 of the naphthyridine system, and therefore this structural feature of the bA$_1$AR binding site could explain the higher activity found for compounds that possess only an unsubstituted phenyl ring in position 2 of the naphthyridine system like, for example, **16** and **29**. In fact, the presence of substituents on this phenyl ring, as, for example, in **13, 26** and **43** and the replacement of this phenyl ring with a less bulky group, as in **50, 51** and **52**, induces a decrease in the affinity. Moreover, the insertion of a spacer between the naphthyridine system and this phenyl ring, as in **46**, induces the almost complete loss of affinity.

In the case of the hA$_1$AR, the H-bonds and the hydrophobic interaction are weaker due to the larger dimensions of the site; moreover in this receptor, Ile270 is substituted by the less hydrophobic Thr270.

In particular, the lipophilic pocket in the hA$_1$AR is larger and the presence of a phenyl ring linked in position 2 is less important for the affinity.
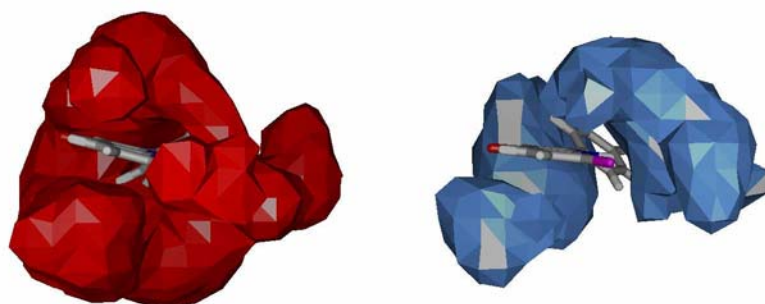
Figure 4. Compound **16** docked into the hA$_1$AR (left) and bA$_1$AR (right) binding sites. The volumes of the cavities between **16** and the two receptors are shown.
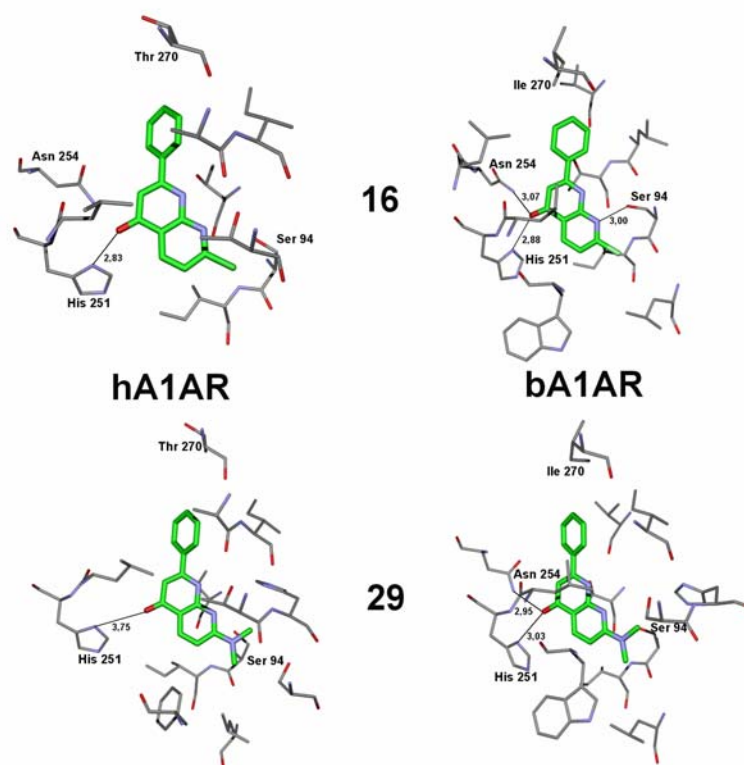
Figure 5. Compounds **16** (up) and **29** (down) docked into the hA₁AR (left) and bA₁AR (right) binding sites.

1.2.3 Conclusions.

In order to verify if the great A₁ species selectivity showed by a series of 1,8-naphthyridine derivatives could be explained in terms of a difference between the binding site of the native hA₁AR and the bA₁AR, these two sites were modelled, taking into account the available mutagenesis data. The interaction energy values of the 1,8-naphthyridine derivatives with the two receptor models were in agreement with their affinities, and therefore with the observed species selectivity; the better interaction with the bA₁AR seems to be due to the smaller size of the binding site of this receptor, which allows a particularly good interaction with these 1,8-naphthyridine derivatives. In fact, the molecules of these ligands have a lower hindrance with respect to the classic A₁AR

antagonists like DPCPX, and therefore they can better occupy this site. The Figure 6 shows that the volume of DPCPX is larger than that of **16** and that it is mainly due to the precence of the freely rotatable *n*-propyl chains in the structure of DPCPX.
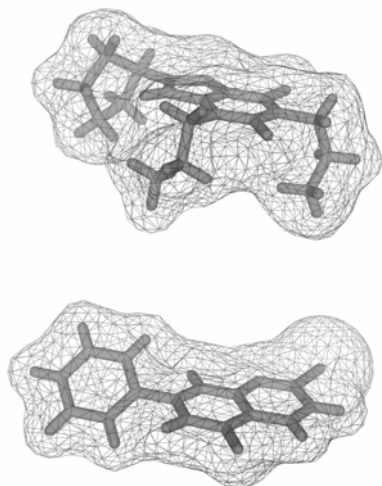
Figure 6. Molecular volumes of DPCPX (up) and **16** (down).

Other differences in the structure of the hA$_1$AR and the bA$_1$AR, and in particular in the loop regions, could provide further reasons for the species selectivity, but the results of the molecular modeling study suggest that in the case of the 1,8-naphthyridine derivatives, their interaction with the intrahelical binding site of the two receptors should be surely responsible for this selectivity.

1.2.4 Experimental section.

1.2.4.1 Computational details. All the molecular mechanics and molecular Dynamics calculations were performed through the Macromodel[20] program by using the AMBER forcefield. The electrostatic charges were those included in the forcefield and a distance-dependent dielectric constant of 4.0 was used. In molecular mechanics minimizations (MM) the minimized value was the Conjugated Gradient until a

convergence value of 0.1 Kcal/A·mol; in molecular dynamics simulations (MD) the temperature was set at 300 °K and the time step was 1 femtosecond.

All graphic manipulations and visualizations were performed by means of the InsightII[54] and WebLabViewer[21] programs.

1.2.4.2 Modeling of the bA$_1$AR. The procedures for the modeling of the bA$_1$AR model were the same already used for the hA$_1$AR (see Chapter 1.1), but this time the template was the hA$_1$AR model and not bovine rhodopsin. In this case, due to the high homology between the hA$_1$AR and the bA$_1$AR, neither translation nor rotation of helices was required.

1.3 PHARMACOPHORE BASED RECEPTOR MODELING: THE CASE OF ADENOSINE A$_3$ RECEPTOR ANTAGONISTS. AN APPROACH TO THE OPTIMIZATION OF PROTEIN MODELS.

1.3.1 Introduction.

A$_3$ receptor activation results in general hypotension and in mast-cells degranulation,[55,56] selective antagonists of these receptors may be used in clinical practice as anti-inflammatory[57] as well as cerebroprotective[58,59] and antiasthmatic agents.[60]

In the last decade, due to the importance of this new candidate biological target, a great effort has been made to design and synthesize new potent and selective agonists and antagonists of the human A$_3$ adenosine receptor (hA$_3$AR). Due to the lack of experimental 3D structural data about hA$_3$AR binding site, the rational design of hA$_3$AR antagonists has been commonly pursued through the construction of receptor models or, alternatively, by QSAR/3D-QSAR approaches - such as CoMFA - using experimental data obtained by a previous generation of ligands. As a result of all this reseach, several classes of compounds, including pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives,[61-63] triazoloquinazoline derivatives,[64] isoquinoline and quinazoline analogues,[65-67] 3,5-Diacyl-2,4-dialkylpyridine,[68-70] and pyridine derivatives,[71,72] have been synthesized and tested as hA$_3$AR antagonists.

The wish to synthesize a new generation of potent and selective antagonists, with improved ADME profile with respect to previous analogues (increased water solubility), drove us to engage in the development of a novel modeling approach. In our research we actually merged the capabilities of two generally mutually exclusive drug design computational methodologies, that is pharmacophore[73] and homology modeling,[74] to drive the building of a predictive three-dimensional model of the human A$_3$ receptor. Differently from the standard homology model building so far reported in literature,[63,68,75] where the raw 3D structure is usually refined by means of molecular mechanics calculations on its complex with a high affinity ligand, for the aim of our work we indeed used a pharmacophoric model to drive the refinement of the receptor model. The idea beyond this approach was to gain information from a set of ligands, by

means of a 3D-QSAR methodology, in order to increase the probability of characterizing the most important features for ligand recognition.

With this idea in mind, we generated several common feature hypotheses for ligands of the human A$_3$ receptor, from which the three most promising pharmacophores were selected: they were all characterized by good statistical parameters and prediction ability towards a test set. Moreover they were very similar in terms of features composition and disposition in three-dimensional space. At the same time, a raw model of hA$_3$AR was generated by homology modeling, using bovine rhodopsin as a template[19] and molecular interaction fields (MIF) were calculated for each transmembrane helix, with the aim to localize the minimum-energy interaction points.

Each pharmacophore was then manually docked into the raw binding site, which was adjusted as well, to get the highest correspondence between the centroid of each pharmacophoric element and the corresponding local minimum of the MIFs. One peculiar alignment was obtained for each pharmacophore through such a procedure so that, after having chosen four reference compounds, twelve ligand-protein complexes could be used to refine the receptor model. Each one of the twelve complexes was in fact relaxed - according to a protocol fully described in this report - and evaluated in terms of energies and "features correspondence", to gain both the final optimum hA$_3$AR model and the best performing pharmacophore, which was used again to export the relative alignment of four more compounds for a qualitative validation step. The predictive power of the final receptor model was probed still further by the finding of a quantitative correlation between experimental free energies of binding and theoretical values, calculated by the application of a scoring function.

Finally, some novel putative A$_3$ antagonists were designed, synthesized and biologically evaluated, to test the reliability of our combined modeling strategy.

1.3.2 Results and Discussion.

1.3.2.1 Pharmacophore generation. hA$_3$AR antagonists are characterized by great structural diversity, making difficult a common chemical pattern to be found.[61-72] Nevertheless, certain common electronic and steric features have been already reported in the literature based on a combination of *ab-initio* calculations, electrostatic potential map comparisons and steric and electrostatic alignment (SEAL) analyses.[76] Starting

from the 55 A$_3$ antagonists showed in chart 1, Prof. Botta and co-workers developed two pharmacoforic models HYPO1, HYPO2 (see Figure 1) using the Catalyst 4.6 software package.[77]

Chart 1. hA$_3$AR Antagonists Considered During the Pharmacophore Generation and Validation

| cmpd	R1	R2	R3	R4	R5	X
1a	Ethyl	NHCONH-(4-SO₃H-Ph)				
1b	Ethyl	NH₂				
1c	Methyl	NHCONH-(3-Cl-Ph)				
1d	Ethyl	NHCONH-4-CH₃-Ph				
1e	Ethyl	NHCONH-4-F-Ph				
1f	Ethyl	NHCONH-(2-OCH₃-Ph)				
1g	Ethyl	NHCONH-(2-Cl-Ph)				
1h	Ethyl	NHCONH-(3-Cl-Ph)				
1i	Methyl	NHCONH-Ph				
1j	Ethyl	NHCONH-(4-OCH₃-Ph)				
1k	Methyl	NHCONH-4-Pyridyl				
1l	Propyl	NH₂				
1m	Propyl	NHCONH-(4-OCH₃-Ph)				
1n	H	NH₂				
1o	Methyl	NHCONH-(4-OCH₃-Ph)				
1p	Ethyl	NHCONH-(4-NO₂-Ph)				
1q	Propyl	NHCONH-Ph				
1r	Propyl	NHCONH-(4-SO₃H-Ph)				
2a	NH₂					
2b	NHCO-Ph					
2c	NHCOCH₂-Ph					
2d	NHCOCH₂CH₃					
2e	NHCOCH₃					
3a	NHCO-(4-OCH₃-Ph)	2-Pyridyl				CH
3b	NHCO-Ph	2-Pyridyl				CH
3c	NHCO-(4-CH₃-Ph)	2-Pyridyl				CH
3d	NHCO-(3,4-CH₃-Ph)	2-Pyridyl				CH
3e	NHCO-(3,4-OCH₃-Ph)	2-Pyridyl				CH
3f	NHCO-3Cl-Ph	2-Pyridyl				CH
3g	NHCO-3OCH₃-Ph	2-Pyridyl				CH
3h	NHCONH-Ph	2-Pyridyl				CH
3i	NHCONH-Ph	H				N
3j	NHCONH-Ph	2-Pyridyl				N
3k	NHCONH-Ph	3-Pyridyl				N
3l	NHCONH-Ph	3-CH₃-2-Pyridyl				N
3m	NHCONH-Ph	N,N-Diethylamino				N
3n	NHCONH-Ph	1-Pyrrolidinyl				N
3o	NHCO-(4-Cl-Ph)	2-Pyridyl				CH
3p	NHCO-(2,4-CH₃-Ph)	2-Pyridyl				CH
3q	NHCO-(3-CH₃-Ph)	2-Pyridyl				CH
4a		Methyl	O-Ethyl	Methyl	COO-Ethyl	
4b		Methyl	O-Propyl	Methyl	COO-Ethyl	
4c		Methyl	O-Ethyl	Ethyl	COO-Ethyl	
4d		Methyl	O-Ethyl	Ethyl	COO-Ethyl	
4e		Ethyl	O-Ethyl	Ethyl	COO-Ethyl	
4f		Ethyl	O-Ethyl	Ethyl	COO-	
4g		Ethyl	O-Ethyl	n-Propyl	COO-	
4h		Ethyl	O-Propyl	Ethyl	COO-	
4i		Methyl	S-Ethyl	Methyl	COO-Ethyl	
4j		Propyl	S-Ethyl	Ethyl	COO-Ethyl	
4k		Methyl	S-Ethyl	Propyl	COO-Ethyl	
4l		Ethyl	S-Ethyl	Ethyl	COOCH₂C	
4m		Ethyl	S-Ethyl	Ethyl	COO-Ethyl	
5a	Methyl	Methyl	O-Ethyl	Ethyl	COO-Ethyl	
5b	Methyl	Ethyl	S-Ethyl	Ethyl	COO-Ethyl	
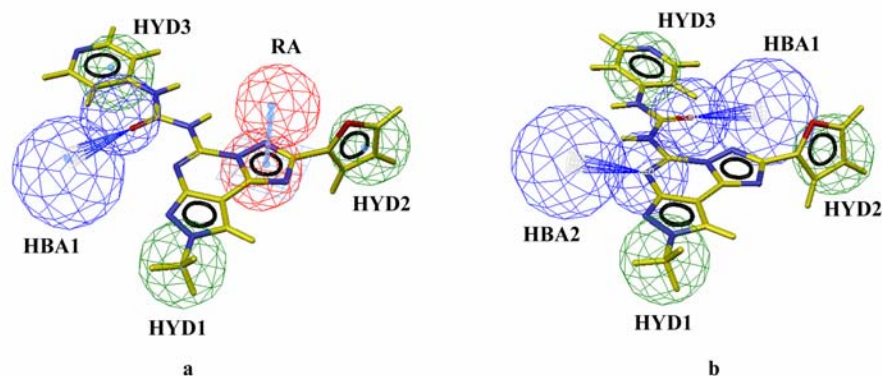
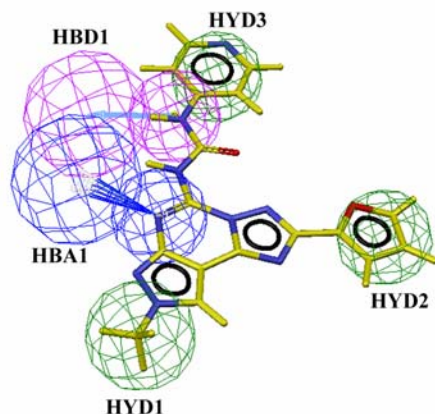**Figure 1.** HYPO1 (a) and HYPO2 (b) superposed to compound 1k. Pharmacophore features are color-coded: green for hydrophobic (HYD), blue for hydrogen bond acceptor (HBA) and red for aromatic ring (RA).

In Figure 1 the whole set of features of both pharmacophores is displayed superposed to compound **1k**, the most active compound among those tested. Both hypotheses were characterized by five features and shared a common scheme consisting of three hydrophobics at the vertices of a triangle (HYD1, HYD2 and HYD3). They differed for a hydrogen bond acceptor (HBA1) pointing towards opposite directions, and for a hydrogen bond acceptor (HBA2), which replaced, in HYPO2, the ring aromatic (RA) found in HYPO1.

On the basis of what already reported by some of us about the importance of an hydrogen bonding donor (HBD) interaction between the NHCONH group present on great part of the compounds considered in this study and hA₃AR,[61,63] Catalyst was pushed to generate new hypotheses with at least one HBD feature. As a result of this effort HYPO3 was chosen for further investigation according to the aforementioned selection rules.

In Figure 2 the whole set of features of HYPO3 is displayed superposed to compound **1k**. HYPO3 presented once more the recurring scheme of three hydrophobic interactions laying at the vertices of a triangle. No more RA features were present in the model, while two hydrophilic interactions (one HBA and one HBD) appeared by one

side of the triangle. HYPO1, HYPO2 and HYPO3 were then exported and manually docked into the receptor model as described afterward in this article.

Figure 2. HYPO3 superposed to compound **1k**. Pharmacophore features are color-coded: green for hydrophobic (HYD), blue for hydrogen bond acceptor (HBA) and pink for hydrogen bond donor (HBD).

1.3.2.2 Receptor modeling and structure optimization. All the information regarding the primary structure of the human A$_3$ receptor, and the subdivision into transmembrane, cytoplasmatic and extracellular domains were obtained from the GPCR Data Bank.[25] A raw structure of hA$_3$AR was obtained through molecular modeling, using bovine rhodopsin as a template.[19] The receptor-template superposition was carried out maintaining the maximum analogy between them, and choosing the regions with a conserved or semi-conserved sequence. The alignment was studied on several adenosine receptors by means of the ClustalW programme[29] and was guided by the highly conserved amino acid residues (see the Figure 1 in Chapter 1.1), including the D/ERY motif (D/E3.49, R3.50, and Y3.51), the two Pro residues P4.50 and P6.50 and the NPXXY motif in the TM7 (N7.49, P7.50, and Y7.53).[30]

The raw hA$_3$AR model was then disassembled, and molecular interaction fields (MIF) were calculated for every single transmembrane helix, by means of GRID program,[78] with the aim of investigating the best regions of protein-protein and ligand-protein interactions. Several probes were used at this stage (see Table 1 for details), such as to

resemble both helix-helix interactions and the features present in the hypotheses proposed by Catalyst and previously described.

Table 1. Details of the probes used for MIF calculations: the complementary Catalyst features are reported in the last column

Probe	Brief description	Complementary Catalyst feature
H	Hydrogen	HYD
DRY	The hydrophobic probe	HYD
C1=	sp2 CH aromatic or vinyl	RA
N:=	sp2 N with lone pair	HBA
N1=	sp2 Amine NH cation	PI, HBD
NH=	sp2 NH with lone pair	HBA, HBD
N1+	sp3 amine NH cation	PI, HBD
N1:	sp3 NH with lone pair	HBA, HBD
N2	Neutral flat NH2 e.g amide	HBD
N2=	sp2 Amine NH2 cation	PI, HBD
N2+	sp3 amine NH2 cation	PI, HBD
N2:	sp3 NH2 with lone pair	HBA, HBD
NM3	Trimethyl ammonium cation	PI
O1	Alkyl hydroxyl OH group	HBA, HBD
OC2	Ether or furan oxygen	HBA
O	sp2 carbonyl oxygen	HBA
F-	Fluorine anion	

After having reassembled the seven helices, the points of local minimum energy of each MIF were visualized superposed to them. At this point, Catalyst hypotheses were manually docked singularly, and for each of them a manual rearrangement of the relative positions and orientations of the helices was performed in order to: a) superpose each pharmacophoric feature found by Catalyst to a minima of a complementary MIF (i.e. the MIF able to favourably interact with such a feature, as shown in Table 1); b) retain the crucial interactions between the pharmacophore features and some important

residues, as highlighted by site directed mutagenesis studies (see Table 2 for the mutagenesis details).[79-82]

Table 2. Mutational analysis for the human A₃ receptor antagonists interaction.

Region	A₃AR	Mutational results
TM3	H95	A: reduction of affinity[a]
L 4-5	K152	A: reduction of affinity[a]
TM6	W243	A: reduction of affinity[a]
	L244	A: modest variation[a]
	S247	A: modest reduction of affinity[a]
	N250	A: lost of binding[a]
TM7	H272	E: reduction of affinity[a]
	Y282	F: reduction of affinity[b]

[a]See reference 79. [b]See reference 80

Having started from three pharmacophoric models (HYPO1, HYPO2 and HYPO3), three different receptor models (MODEL1, MODEL2 and MODEL3) were obtained by this procedure. In all the three cases, the homology model directly obtained through bovine rhodopsin as template were not able to take into account all the mutagenesis data (see Table 2). In particular, in contrast with these data, Hys95 and Hys272 did not point towards the intrahelical channel. Therefore. the TM3 and TM7 were to be rotated respectively of 60° clockwise and 90° counter-clockwise (extra cellular point of view), to let them turn towards the intra-helical channel therefore allowing the interaction of Hys95 and Hys272 with the ligands. On the other hand, in agreement with a finding of Gouldson et al.,[83] rotations and translations of the TM domains are important steps in ligand-receptor interaction process in different GPCRs.

Due to the antagonist profile of the considered ligands, possible rearrangements of the receptor in an activate form able to interact with the agonists was not take into account.

Figure 3 shows the three receptor models and the alignment between the three different pharmacoforic hypotheses (HYPO1, HYPO2, and HYPO3) and the MIFs calculated as above described. As reference, compound **1k** is also displayed.
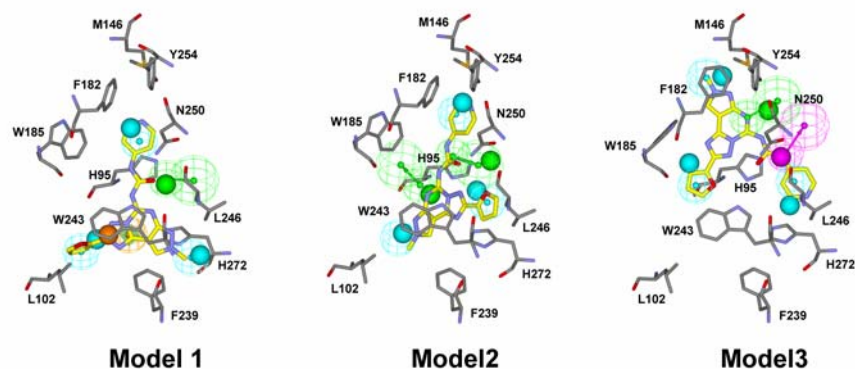
Figure 3. The three receptor models complexed with **1k** and the alignment between the three pharmacoforic hypotheses and the MIFs (displayed as spheres) calculated on the transmembrane helices. In blue are represented the MIFs obtained through the DRY probes, in green the one calculated with OC2 and N:= probes, in orange the MIF derived by the calculation with C1 probe, while in magenta is reported the MIF obtained through the NH= probe.

In MODEL 1 the aromatic feature RA (orange in the figure), matched by the core of **1k**, highlights a π-π interaction with Trp243. This residue has been widely reported to play an important role in antagonists recognition.[79-81] HBA1 (green in the figure) correctly maps the hydrogen bond between the C=O of the ligand and the side chain NH of Asn250. This residue, if mutated, causes loss of affinity for both agonists and antagonists.[79-81] Finally, HYD1 (blue in the lower right part of the figure), HYD2 (in the lower left part of the figure) and HYD3 (in the upper part of the figure) correspond to three receptor hydrophobic clefts (H1, H2 and H3), delimited by: i) Phe239, Trp243, Hys272 (H1); ii) Leu102, Phe239, Trp243 (H2) and iii) Hys95, Phe182, Trp185, Tyr254 (H3).

In MODEL 2 the hydrogen bond interaction between the C=O of the ligand and the side chain NH of Asn250 (HBA1) and the correspondence between HYD3 and the lipophilic cleft constituted by Hys95, Phe182, Trp185, Tyr254 were maintained. Differently from MODEL 1, HYD1 (blue in the lower left part of the figure) corresponded to the lipophilic cleft constituted by Leu102, Phe239, Trp243 and HYD2 (in the lower right part of the figure) corresponded to a lipophilic interaction with Leu246. Regarding the

HBA2 feature, not present in HYPO 1, it matched with a H bond interaction between the nitrogen of the central nucleus of the ligand and His95.

As shown in Figure 3 the features of HYPO3 determined a completely different disposition of the pharmacoforic model inside MODEL 3: HYD1 (blue in upper part of figure) corresponded to a lipophilic cleft delimited by Met146, Phe182 and Trp185, HYD2 (in the lower left part of the figure) corresponded to an interaction with Trp185, while HYD3 (in the lower right part of the figure) matched to a lipophilic interaction with Leu246 and His272. Regarding the electrostatic interaction HBA1 (green in figure), it corresponded to a H bond interaction with Tyr254 while HBD1 (coloured magenta in figure) matched to the H bond between the C=O of Asn250 and the NH of the ligand.

In order to determine which of the three models was the most reliable, derivatives **1j**, **1k**, **2c**, **3a**, among the ones used in the 3D-QSAR studies, were chosen as reference compounds and manually docked into each one of the three receptor models, using both alignment rule and conformer of the corresponding pharmacophoric hypothesis. The twelve generated complexes were then subjected to a relaxation protocol (see the Experimental Section).

Regarding the agreement with the mutagenesis data, the analysis of the results highlighted that for MODEL 1 all the four ligands interacted in the A$_3$ receptor with His95, Trp243, His272 and formed an H bond with Asn250.

In MODEL 2 there are the lipophilic interactions with His95, Trp243, His272 for all the four ligands, while only three ligands possessed the electrostatic interaction with Asn250, but for the A$_3$-**3a** complex it was absent. Furthermore in all the final complexes it was not possible to find the H bond interaction predicted by HYPO2 on the nitrogen of the aromatic system (HBA2 feature).

In MODEL 3 all the ligands interacted with His95, Trp243, His272 and Asn250, however, as a result of the optimization procedure (see Experimental Section), in the A3-**2c** and A3-**3a** complexes the ligands interacted with Asn250 as H bond acceptor, in contrast to the pharmacophoric hypothesis (HYPO3) which predicted an HBD feature (HBD1).

Therefore, although the mutagenesis data were generally respected for all the three models, only MODEL 1 was able to respect for all the four tested compounds both the mutagenesis data and the pharmacophoric features predicted by Catalyst.

As second test for verifying the reliability of the models, the ligand-receptor interaction energy was calculated for all the twelve complexes. Generally it was not possible to find a quantitative correlation between the calculated energy of the complexes and the activity of the compounds, mainly because of the lack of the solvation and entropic terms. However, in this case, only the interaction of the same ligand with different conformations of the same receptor was to be evaluated and therefore these terms could be considered approximately constant.

As shown in Table 3, MODEL 1 showed the highest interaction energy for all four ligands. Although the difference in energies between the three models is small and probably likely within the error of the method, the fact that for all ligands the best interaction energy was found with MODEL 1, suggested to select this model.

Table 3. Interaction energies (KJ/mol) of compounds **1i**, **1k**, **2c**, **3a** in the three considered alignment models: MODEL 1, MODEL 2 and MODEL 3.

cmpd	MODEL 1	MODEL 2	MODEL 3
1j	-55.9	-52.1	-53.5
1k	-61.4	-53.4	-56.8
2c	-58.5	-53.5	-55.7
3a	-55.2	-47.0	-47.6

Finally, compounds **2d**, **3i**, **4a** and **4f** were docked inside the three models using the same protocol applied for compounds **1j**, **1k**, **2c**, and **3a**.

To probe which of the three models were the most reliable, the free energies of binding of the complexes with compounds **1k**, **1j**, **2c**, **2d**, **3a**, **3i**, **4a** and **4h** were calculated by means of the AutoDock 3.0 scoring function.[84] This docking application proved in fact to be reliable in many works present in literature,[85,86] since its free energy function, based on the principles of QSAR, has been parameterized using a large number of

protein-inhibitor complexes for which both structure and inhibition constants were known.

As shown in Table 4 and Figure 4, a good correlation between the experimental and calculated values of free energy of binding for all the studied compounds was obtained only for MODEL 1 as the quadratic correlation (r^2) has a value of 0.69 while for MODEL 2 and MODEL 3 r^2 has the value of 0.56 and 0.57 respectively.
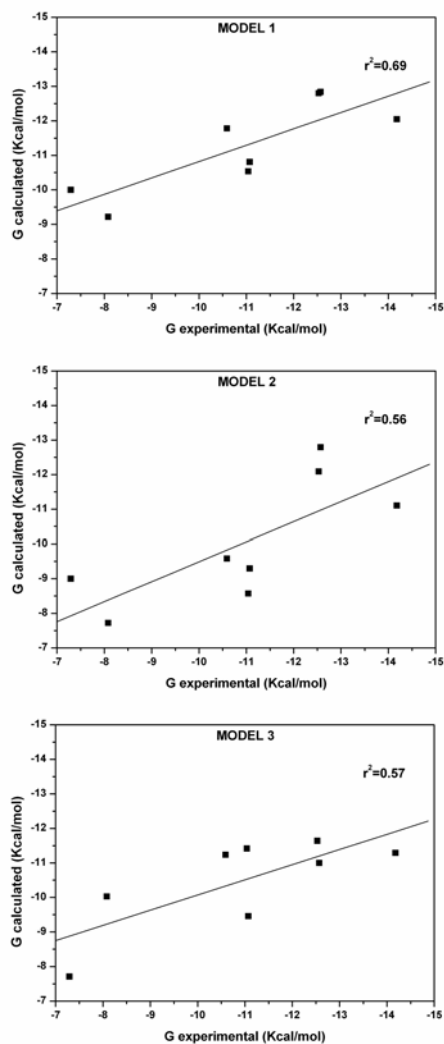

Figure 4. Experimental versus calculated (AutoDock) binding energy of compounds **1k**, **1j**, **2c**, **2d**, **3a**, **3i**, **4a** and **4h** (indicated with ■)

Table 4. K_i (nM) of the compounds analyzed, Experimental Binding Energy (G_{Exp}, Kcal/mol), and AutoDock Binding Energy (G_{Calc}, Kcal/mol) for all the A_3 receptor complexes.

			MOD 1	MOD 2	MOD 3
cmpd	K_i	G_{exp}	G_{calc}	G_{calc}	G_{calc}
1k	0.04	-14.18	-12.05	-11.10	-11.29
1j	0.60	-12.57	-12.84	-12.79	-11.00
2c	0.65	-12.53	-12.80	-12.09	-11.64
2d	7.66	-11.07	-10.81	-9.30	-9.46
3a	17	-10.59	-11.78	-9.58	-11.24
3i	1180	-8.08	-9.22	-7.72	-10.03
4a	4470	-7.29	-10.00	-9.00	-7.71
4h	7.94	-11.04	-10.54	-8.57	-11.42

All these results suggested MODEL 1 was the best model and for this reason it was chosen for further analyses; consequently also the HYPO 1 hypothesis should be considered the best one.

With the aim to achieve a qualitative knowledge of the features of hA₃AR, the eight relaxed complexes of compounds **1k**, **1j**, **2c**, **2d**, **3a**, **3i**, **4a** and **4h** were compared each other through the superposition of the seven helices. The outcome of such a procedure is shown in Figure 5. The most evident result was that the receptor model was able to vary the width and shape of its three hydrophobic clefts in order to accommodate different ligands. In particular, Trp185 could act as a gate, making the H2 and H3 clefts wider or narrower. In the case of compounds showing a Ki<1 nM (left part of Figure 5), H2 was very small, thus allowing a perfect matching with the ligand and the wider H3 was able to accommodate long chains such as phenylcarbamoyl moieties. On the other hand, for compounds showing Ki>1 nM (right part of Figure 5), Trp185 was in a sort of "open" conformation so that weaker contacts were possible with the ligands. Therefore, the model seemed to be able to discriminate between very active compounds (Ki<1 nM) and less active ones (Ki>1 nM).
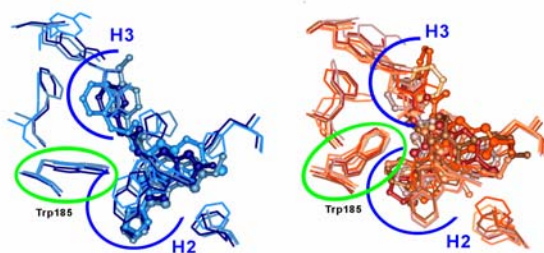
Figure 5. Superposition of the complexes of hA₃AR with the most active ligands, **1j, 1k** and **2c** (left) and with the less active ones, **2d**, **3a**, **3i**, **4a** and **4f** (right); for sake of clarity, only residues under consideration are reported

Regarding the interactions between **1k** and the hA₃ receptor, beyond the interactions mentioned above, already suggested by site directed mutagenesis studies, the residue Tyr254 assume in our A₃ model an important role as it formed a H bond with the pyridyl ring of the ligand (see Figure 6). This interaction could justify the improved activity of compounds like **1k**, presenting a proper hydrogen bonding acceptor group in that region (4-pyridyl) and the irreversible inhibition of p-fluorosulfonyl-pyrazolo[4,3-e]1,2,4-triazolo[1,5-c]pyrimidine derivatives.[87] The interaction with Tyr254, residue non-conserved in other adenosine receptors, could still more explain the high selectivity of **1k**.
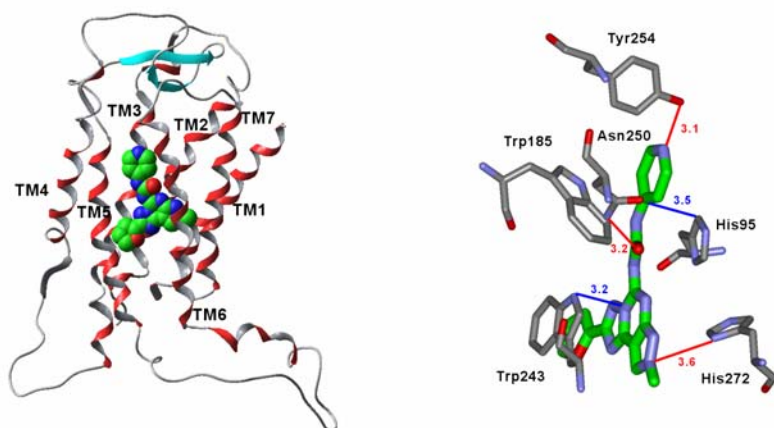
Figure 6. Side view of MODEL1, complexing **1k** into its intrahelical region and details of the important interactions.

1.3.2.3 Design and synthesis of hA$_3$AR antagonists. The analysis of the recognition geometry of compound **1k** into MODEL1 suggested some structural changes to be made on the pyrazolotriazolopyrimidines, with the aim to improve both their water solubility and ADME properties together with the structural features which appear to be fundamental for the activity and selectivity.

Particularly, a small hydrophilic pocket, bordered by two serine residues (Ser242 and Ser275), lay empty nearby the H1 cleft of the receptor. Hence, the introduction of a short alkyl chain at N^8 of **1k**, terminating with a hydrophilic group, was hypothesized, to verify the possibility to give rise to binding interactions within this pocket.

On the contrary it appears that the interactions with His98, Trp243, Asn250, Tyr254 and His272 (see Figure 6) are to be conserved; therefore the tricyclic system of **1k** together with the pyridine-urea moiety should be maintained.

Following these observations compounds **16**, **17** and **18** (see Figure 7) were designed, virtually evaluated *in silico* (applying the whole modeling procedure described above), synthesized and tested for their affinity towards human A_1, A_{2A}, A_{2B} and A_3 adenosine receptors (see the Experimental Section).
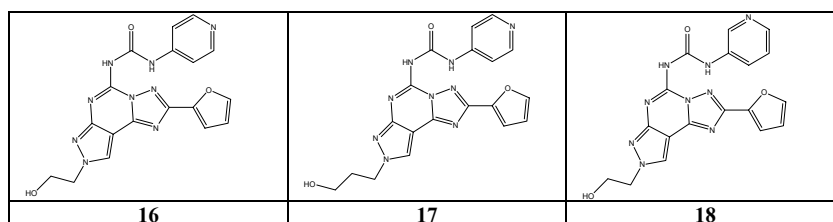
Figure 7. Compound synthesized and tested.

The affinity values measured are shown in Table 5, which also shows the theoretical values, predicted for the A_3 receptor by the AutoDock scoring function (notably, the scoring function hypothesized an activity in the range between 0.5 and 2 nM for these ligands). Compound **17** resulted the most promising selective A_3 antagonist. The docking of this derivative into MODEL 1 is shown in Figure 8: the expected hydrogen bonding interaction between the hydroxyl substituent at N^8 and Ser275 is highlighted. Even if the affinity of **17** for the A_3 receptor dropped of about two orders of magnitude with respect to compound **1k**, nevertheless, its receptor subtypes selectivity was

47

completely conserved and this could confirm the importance of the pyridine ring for the A$_3$ selectivity. As regards the affinity drop, this could be due to a partial damage of the π-π stacking interaction with Trp243 due to the formation of the new hydrogen bond (see Figure 8) as well as to an entropy penalty associated with the high number of degrees of freedom of the alkyl chain, in any case there is a quite good affinity prediction by our hA$_3$AR model.

Table 8. Affinity Values of the Compounds Synthesized During this Work (See the Parmacology Section)

cmpd	K$_i$ (nM)				
	hA$_1$[a]	hA$_{2A}$[b]	hA$_{2B}$[c]	hA$_3$[d]	Predicted hA$_3$[e]
16	>1000	>1000	>1000	5.1 (4.1-6.5)	1.78
17	350±30	>1000	>1000	2.0 (1.7-2.4)	0.78
18	>1000	>1000	>1000	34 (28-40)	1.58

[a]Displacement of specific [³H]DPCPX binding at human A$_1$ receptors expressed in CHO cells. [b]Displacement of specific [³H]ZM 241385 binding at human A$_{2A}$ receptors expressed in CHO cells. [c]Displacement of specific [³H]MRE2029F20 binding at human A$_{2B}$ receptors expressed in CHO cells. [d]Displacement of specific [³H]MRE3008F20 binding at human A$_3$ receptors expressed in CHO cells. [e]Predicted A$_3$ Ki affinity (using our A$_3$ model and AUTODOCK scoring function).
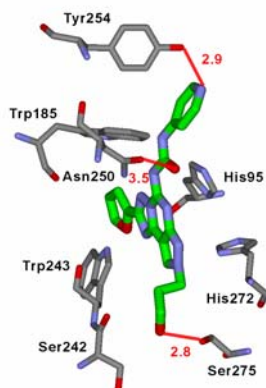

Figure 8. Compound **17** docked into the putative binding site. An hydrophilic moiety at N^8 may be accounted for an interaction with the Ser275.

48

1.3.3 Conclusions.

The construction of a human G-protein coupled receptor model through a homology procedure solely based on the bovine rhodopsin structure is a quite unreliable task, because of usually low homology percentages and the high degree of mobility of the helices. The use of mutagenesis data is nowadays an important improvement of the procedure because it allows to take into account residues experimentally found to be necessary for interaction. The relative positions of these residues, however, remain unknown and can be only hypothesized through the docking into the receptor of a ligand, which retains all these interactions: generally the alignment for this docking is to be manually performed. In the procedure described herein, differently, the use of a pharmacophoric model, representing the activity data of a lot of ligands, allowed the building of a receptor model in which the relative positions and distances between important residues was determined by the distances between pharmacophoric features.

As a whole, our combined modeling strategy, slightly differs from canonical procedures. The building of 3D-QSAR models in fact, helped us to exploit a set of active molecules to highlight statistically relevant features in ligand-receptor interactions and such an information resulted to be crucial to generate a reliable model of hA₃AR (MODEL1). The information produced through docking studies, in like manner, allowed to select the best performing pharmacophore model (HYPO1), among a set of plausible hypotheses.

MODEL1 seemed to be able to explain different modes of binding of very active compounds with respect to less active ones and also to reproduce with a good approximation the free energies of binding. The model was also able to explain the selectivity of **1k** towards the hA₃AR due to the presence of the non-conserved residue Tyr254 and therefore suggested that a mutagenesis study on this residue could be of great importance for the knowledge of the molecular features determining the selectivity at the AR subtypes.

The reliability of our synergistic approach was tested by the rational design and synthesis of a series of novel compounds. Biological assays evidenced that reasonably good activity values were reached and, at the same time, water solubility was enhanced with respect to previously synthesized compounds.

1.3.4 Experimental section.

Molecular Mechanics (MM) and Molecular Dynamics (MD) calculations were performed using the AMBER force field as implemented in the MacroModel software package,[20] using a "distance-dependent" dielectric constant of 4.0. Electrostatic charges for the set of ligands were calculated with RHF/AM1 semi-empirical calculation and RESP program.[31]

All MM minimizations were performed with either Polak-Ribier conjugate gradient or steepest descent as minimizers and a threshold value of 0.05 kJ/Å•mol as the convergence criterion. The temperature was set at 300 °K and the time step was 1.0 fs in MD simulations.

All graphical manipulations and visualizations were performed by means of the InsightII,[54] UCSF-CHIMERA[88] and WebLab Viewer[21] programs.

The alignment of several adenosine receptors was studied with the ClustalW program[29] using the blosum algorithm, with a gap open penalty of 10 and a gap extention penalty of 0.05. From the ClustalW alignement, the structure of the seven TM helices of hA$_3$AR and the first intracellular loop were constructed directly from the coordinates of the corresponding amino acids in rhodopsin by means of Modeller program.[89] Through Maestro interface the TM3 and TM7 were rotated respectively of 60° clockwise and 90° counter-clockwise (extra cellular point of view), to let Hys95 and Hys272 turn towards the intra-helical channel. As the amino acid length differs from the template and for the rotation of TM3 and TM7 the other loops were constructed by means of the "Loop optimization method" of Modeller, applying the "very_slow" loop refinement method. The model was subjected to a preliminary minimization and to 400 ps of MD (after 50 ps of equilibration), the final structure was then minimized. When MD simulations were carried out in the gas phase, all the alpha carbons of the TM of the protein were blocked by means of decreasing force constants, to simulate the stabilizing presence of the membrane around the receptor. For the first 200 ps, restraints with a force constant of 10 Kcal/mol•Å^2 were applied to Cα, and for the remaining 200 ps these restraints were gradually reduced to 1 Kcal/mol•Å^2.

The refinement of the ligand-protein complexes was initially performed by means of a total of 400 ps of MD. All the alpha carbons of the TM, and the main ligand-receptor

interactions were constrained during the trajectory by means of decreasing force constants. More in detail, an initial restraint with a force constant of 10 Kcal/mol•Å^2 was applied on the alpha carbons, this force constant decreased during the whole MD and in the last 200 ps a value of 0.1 Kcal/mol•Å^2 was applied. As regard the H bond ligand-receptor interactions, suggested by the HBA1, HBA2 and HBD1 features found by Catalyst, a restraint of 50 Kcal/mol•Å^2 was applied in order to stabilize a ligand-receptor complex structure maintaining all these interactions. After the end of the MD simulation, three steps of minimization were applied on the average structure obtained during the last 100 ps of the MD run. During these three steps a restraint of 0.1 Kcal/mol•Å^2 was applied on the alpha carbons, while as regard the main ligand-receptor interaction, in the first two steps a restraint of respectively 25 and 10 Kcal/mol•Å^2 was applied and in the last one the restraints were removed.

The quantitative evaluation of the free energy of binding of the twelve complexes was performed by means of the AutoDock scoring function,[84] using the Lamarckian Genetic Algorithm. The region of interest used by AutoDock was defined considering the **1k** docked in hA$_3$AR as center group; in particular a grid of 40, 54, and 50 points in *x*, *y*, and *z* direction was build centered on the center of mass of **1k**. A grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the energetic maps calculation.

1.4. REFERENCES.

1. Fredholm, B.B.; Abbracchio, M.P.; Burnstock, G.; Dubyak, G.R.; Harden, T.K.; Jacobson, K.A.; Schwabe, U.; Williams, M. Trends Pharmacol. Sci. 18 (1997) 79-82.

2. Libert, F.; Schiffmann, S.N.; Lefort, A.; Parmentier, M.; Gerard, C.; Dumont, J.E.; Vanderhaeghen, J.J.; Vassart, G. EMBO J. 10 (1991) 1677-1682.

3. Libert, F.; Van Sande, J.; Lefort, A.; Czernilofsky, A.; Dumont, J.E.; Vassart, G.; Ensinger, H.A.; Mendla, K.D. Biochem. Biophys. Res. Commun. 187 (1992) 919-926.

4. Townsend-Nicholson, A.; Shine, J. Brain Res. Mol. Brain Res. 16 (1992) 365-370.

5. Ralevic, V.; Burnstock, G. Pharmacol. Reviews 50 (1998) 413-492.

6. Fink, J.S.; Weaver, D.R.; Rivkees, S.A.; Peterfreund, R.A.; Pollack, A.E.; Adler, E.M.; Reppert, S.M. Brain Res. Mol. Brain Res. 14 (1992) 186-195.

7. Bruns, R.F. Ann. N. Y. Acad. Sci. 603 (1990) 211-225.

8. Jarvis, M.F.; Schulz, R.; Hutchison, A.J.; Do, U.H.; Sills, M.A.; Williams, M. J. Pharmacol. Exp. Ther. 251 (1989) 888-893.

9. Feoktistov, I.; Polosa, R.; Holgate, S.T.; Biaggioni, I. Trends Pharmacol. Sci. 19 *(*1998), 148-153.

10. Müller, C.E.; Stein, B. Curr. Pharm. Des. *2* (1996) 501-530.

11. Poulsen, S-A.; Quinn, R.J. Bioorg. Med. Chem. 6 (1998) 619-641.

12. Muller, C.E. Expert Opin. Ther. Pat. 7 (1997) 419–440.

13. Hess, S. Expert Opin. Ther. Pat. 11 (2001) 1533–1561.

14. Stone, T.W.; Collis, M.G.; Williams, M.; Miller, L.P.; Karasawa, A.; Hillaire-Buys, D. In Pharmacological Sciences: Perspectives for Research and Therapy in the Late 1990s.

15. Cuello, A. C.; Collier, B. Eds.; Birkhauser Verlag: Basel, Switzerland, (1995) 303-309.

16. Richardson, P.J.; Kase, H.; Jenner, P.G. Trends Pharmacol. Sci. 18 (1997) 338-344.

17. Ledent, C.; Vaugeois, J.M.; Schiffmann, S.N.; Pedrazzini, T.; El Yacoubi, M.; Vanderhaeghen, J.J.; Costentin, J.; Heath, J.K.; Vassart, G.; Parmentier, M. Nature 388 (1997) 674-678.

18. Baraldi, P.G.; Cacciari, B.; Spalluto, G.; Borioni, A.; Viziano, M.; Dionisotti, S.; Ongini, E. Curr. Med. Chem. 2 (1995) 707-722.

19. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C.A.; Motoshima, H.; Fox, B.A.; Le Trong, I.; Teller, D.C.; Okada, T.; Stenkamp, R.E.; Yamamoto, M.; Miyano, M. Science 289 (2000) 739-745.

20. Macromodel ver. 7.0, Schrodinger Inc., 1999.

21. WebLab Viewer Pro 3.7 (Accelrys Inc., San Diego, CA)

22. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng , E.C.; Ferrin, T.E. J. Comput. Chem. 25 (2004) 1605-1612.

23. Gaussian 98, a.5. M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B.G. Johnson, W. Chen, M.W. Wong, J.L. Andres, M. Head-Gordon, E. S. Replogle, J.A. Pople, Gaussian, Inc., Pittsburgh PA, 1998.

24. Ballesteros, J.A.; Weinstein, H.W. Methods in Neurosci. 25 (1995) 366-428.

25. GPCRDB: Information system for G protein-coupled receptors (GPCRs), www.gpcr.org.

26. Olah, M.E.; Jacobson, K.A.; Stiles, G.L. J. Biol. Chem. 269 (1994) 24692-24698.

27. Kim, J.; Jiang, Q.; Glashofer, M.; Yehle, S. J.; Wess, Jacobson, K.A. Mol. Pharm. 49 (1996) 683-691.

28. Rivkees, S.A.; Lasbury, M.E.; Barbhaiya, H. J. Biol. Chem. 270 (1995) 20485-20490.

29. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. Nucl. Acids. Res. 22 (1994) 4673-4680.

30. Fredriksson, R.; Lagerström, M.C.; Lundin, L.G.; Schiöth, H.B. Mol. Pharm. 63 (2003) 1256-1272

31. Bayly, C.I.; Cieplak, P.; Cornell, W.D.; Kollman, P.A. J. Phys. Chem. 97 (1993) 10269-10280.

32. Rivkees, S.A.; Barbhaiya, H.; IJzerman, A.P. J. Biol. Chem. 274 (1999) 3617-3621.

33. Barbhaiya, H.; McClain, R.; Ijzerman, A.; Rivkees, S.A. Mol. Pharmacol. 50 (1996) 1635-1642.

34. Gao, Z.; Jiang, Q.; Jacobson, K.A.; Ijzerman, A.P. Biochem. Pharmacol. 60 (2000) 661-668.

35. Scholl, D.J.; Wells, J.N. Bioc. Pharm. 60 (2000) 1647-1654.

36. Jiang, Q.; Lee, B.X.; Glashofer, M.; van Rhee, A.M.; Jacobson, K.A. J. Med. Chem. 40 (1997) 2588-2595

37. Jiang, Q.; Van Rhee, A.M.; Kim, J.; Yehle, S.; Wess, J.; Jacobson, K.A. Mol. Pharmacol. 50 (1996) 512-521.

38. Fredholm, B.B.; IJzerman, A.P.; Jacobson, K.A.; Klotz, K.N.; Linden, J. Pharm. Rev. 53 (2001) 527-552.

39. Gao, Z.G.; Chen, A.; Barak, D.; Kim, S.K.; Müller, C.E.; Jacobson, K.A. J. Biol. Chem. 277 (2002) 19056-19063.

40. Kim, J.; Wess, J.; van Rhee, A.M.; Schoneberg, T.; Jacobson, K.A. J. Biol. Chem. 270 (1995) 13987-13997.

41. Olah, M.E.; Stiles, G.L. Pharm. Ther. 85 (2000) 55-75.

42. Dawson, E.S.; Wells, J.N. Mol. Pharm. 59 (2001) 1187-1195.

43. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thorton, J.M. J. Appl. Crystallogr. 26 (1993) 283–291.

44. Ballesteros, J.A.; Jensen, A.D.; Liapakis, G.; Rasmussen, S.G.F.; Shi, L.; Gether, U.; Javitch, J.A. J. Biol. Chem. 276 (2001) 29171-29177.

45. Croce, R.; Breton, J.; Bassi, R. Biochem. 35 (1996) 11149-11159.

46. Nikiforovich, G.V.; Marshall, G.R. Biochem. 42 (2003) 9110-9120.

47. Salvatore, C. A.; Jacobson, M. A.; Taylor, H. E.; Linden, J.; Johnson, R. G. Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 10365-10369.

48. Linden, J.; Taylor, H. E.; Robeva, A. S.; Tucker, A. L.; Stehle, J. H.; Rivkees, S.A.; Fink, J. S.; Peppert, S.M. Mol. Pharmacol. 44 (1993) 524-532.

49. Ji, X.-D.; von Lubitz, D.; Olah, M.; Stiles, G. L.; Jacobson, K. A. Drug Dev. Res. 33 (1994) 51-59.

50. Müller, C. E.. Expert Opin. Ther. Pat. 5 (1997) 419-440.

51. Müller, C. E; Scior, T. Pharm. Act. Helv. 68 (1993) 77-111.

52. Olah, M.; Ren, H.; Ostrowski, J.; Jacobson, K. A.; Stiles, G. L. J. Biol. Chem. 267 (1992) 10764-10770.

53. Townsend-Nicholson, A.; Schofield, P. R. J. Biol. Chem. 269 (1994) 2373-2376.

54. Insight II (98) Molecular Modeling System; MSI, 1998

55. Schaick E.A.v., Jacobson K.A., Kim H.O., IJzerman A.P., Danhof M. Eur. J. Pharmacol. 308 (1996) 311-314.

56. Hannon J.P., Pfannkuche, H.J., Fozard J.R. Br. J. Pharmacol. 115 (1995) 945-952.

57. Ramkumar V., Stiles G.L., Beaven M.A., Ali H. J. Biol. Chem. 268 (1993) 16887-16890.

58. Lubitz D.K.J.E.v., Carter M.F., Deutsch S.I., Lin R.C.S., Mastropaolo J., Meshulam Y., Jacobson K.A. Eur. J. Pharmacol. 275 (1995) 23-29.

59. Lubitz D.K.J.E.v., Lin R.C.S., Popik P., Carter M.F., Jacobson K.A. J. Med. Chem. 35 (1992) 407-422.

60. Beaven M.A., Ramkumar V., Ali H. Trends. Pharmacol. Sci.15 (1994) 13-14.

61. Baraldi P.G., Cacciari B., Romagnoli R., Spalluto G., Klotz K.N., Leung E., Varani K., Gessi S., Merighi S., Borea P.A. J. Med. Chem. 42 (1999) 4473-4478.

62. Baraldi P.G., Cacciari B., Romagnoli R., Spalluto G., Moro S., Klotz K.N., Leung E., Varani K., Gessi S., Merighi S., Borea P.A. J. Med. Chem. 43 (2000) 4768-4780.

63. Baraldi P.G., Cacciari B., Moro S., Spalluto G., Pastorin G., Da Ros T., Klotz K.N., Varani K., Gessi S., Borea P.A. J. Med. Chem. 45 (2002) 770-780.

64. Kim Y.C., Ji X.D., Jacobson K.A. J. Med.Chem. 39 (1996) 4142-4148.

65. Van Muijlwijk-Koezen, J.E.; Timmerman, H.; Link, R.; van der Goot, H.; Ijzerman, A.P. J. Med. Chem. 41 (1998) 3987-3993.

66. Van Muijlwijk-Koezen J.E., Timmerman H., Link R., van der Goot H., IJzerman A.P. Med. Chem. 41 (1998) 3994-4000.

67. Van Muijlwijk-Koezen J.E., Timmerman H., van der Goot H., Menge W.M.P.B., von Drabbe Künzel J.F., de Groote M.,Ijzerman A.P. J. Med Chem. 43 (2000) 2227-2238.

68. Li A.H., Moro S., Melman N., Ji X.D. Jacobson K.A. J. Med. Chem. 41 (1998) 3186-3201.

69. Li A.H., Moro S., Forsyth N., Melman N., Ji X.D., Jacobson K.A. J. Med. Chem. 42 (1999) 706-721.

70. Xie R., Li A.H., Ji X.D., Melman N., Olah M.E., Stiles G.L., Jacobson K.A. J. Med. Chem. 42 (1999) 4232-4238.

71. Van Rhee A.M., Jiang J.L., Melman N., Olah M.E., Stiles G.L., Jacobson K.A. J. Med. Chem. 39 (1996) 2980-2989.

72. Jiang J.L., Van Rhee A.M., Chang L., Patchornik A., Ji X.D., Evans P., Melman N., Jacobson K.A. J. Med. Chem.40 (1997) 2596-2608.

73. Di Santo R., Tafi A., Costi R., Botta M., Artico M., Corelli F., Forte M., Caporuscio F., Angiolella L., Palamara A. T. Antifungal Agents. J. Med. Chem. 48 (2005) 3019-3031.

74. Chapter 1.1

75. Jiang J.L, Li A.H., Jang S.Y., Chang L., Melman N., Moro S., Ji X.D., Lobkovsky E.B., Clardy J.C., Jacobson K.A. J. Med. Chem. 42 (1999) 3055-3065.

76. Moro S., Li A.H., Jacobson K.A. J. Chem. Inf. Comput. Sci. 38 (1998) 1239-1248.

77. Catalyst 4.6, Accelrys, Inc.: 9685 Scranton Road, San Diego, CA 92121 USA.

78. Cruciani G, Goodford P. Copyright Molecular Discovery Ltd. 2001-2003 GREATER graphical interface for GRID version, 1.1.7 GRIB, UPF/IMIM, Barcelona (Spain) http://www.moldiscovery.com.

79. Chen A., Gao Z.G., Barak D., Liang B.T., Jacobson K.A. Biochem Biophys Res Commun. 284 (2001) 596-601.

80. Gao Z.G., Chen A., Barak D., Kim S.K., Muller C.E., Jacobson K.A. J. Biol. Chem. 277 (2002) 19056-19063.

81. Gao Z.G., Kim S.K., Biadatti T., Chen W., Lee K., Barak D., Kim S.G., Johnson C.R., Jacobson K.A. J. Med. Chem. 45 (2002) 4471-4484.

82. Jacobson K.A., Gao Z.G., Chen A., Barak D., Kim S.A., Lee K., Link A., Van Rompaey P., Van Calenberg S., Liang B.T. J. Med. Chem. 44 (2001) 4125-4136.

83. Gouldson P.R., Snell C.R., Reynolds C.A. J. Med. Chem. 40 (1997) 3871-3886.

84. Morris G. M., Goodsell D. S., Halliday R.S., Huey R., Hart W. E., Belew R. K. and Olson A. J. J. Comput. Chem. 19 (1998) 1639-1662.

85. Costantino G., Macchiarlo A., Campioni E., Pellicciari R. J. Med. Chem. 44 (2001) 3786-3794.

86. Tanczos A. C., Palmer R. A., Potter B. S., Saldanha J. W., Howlin B. J. Comp. Biol. and Chem. 28 (2004) 375-385.

87. Baraldi P.G., Cacciari B., Moro S., Romagnoli R., Ji X., Jacobson K.A., Gessi S., Borea P.A., Spalluto G. *J. Med. Chem.* 44 (2001) 2735-2742.

88. Huang C.C., Couch G.S., Pettersen E.F., and Ferrin T.E. Pacific Symposium on biocomputing 1:724 (1996), http://www.cgl.ucsf.edu/chimera

89. Fiser A., Do R.K., Sali A. Science 9 (2000) 1753-1773.

CHAPTER 2

MATRIX METALLOPROTEINASES

2.1 AMBER FORCE FIELD IMPLEMENTATION, MOLECULAR MODELING AND DESIGN OF A NEW MMP-2/MMP-1 SELECTIVE INHIBITOR.

Tuccinardi T, Martinelli A[*], Nuti E, Carelli P, Balzano F, Uccello-Barretta G, Murphy G, Rossello A[*] Bioorg. Med. Chem In press (DOI: 10.1016/ j.bmc.2006.01.056)

2.1.1 Introduction.

Metalloproteinases are one of the five catalytic classes of proteinases found in man. All metalloproteinases use a Zn^{2+} ion, linked stably to their catalytic site, to effect amide bond hydrolysis on their natural peptide substrates.[2] Several types of zinc enzymes are known and are presently the object of intensive studies, *e.g.* carboxypeptidase A (CPA), histone deacetylases (HDACs), tumour necrosis factor α-convertase (TACE), and matrix metalloproteinases (MMPs).[3] The human MMP family is known to include at least 23 enzymes, divided into four sub-groups on the basis of their substrate specificity: collagenases (MMP-1, -8, -13), gelatinases (MMP-2, -9), stromelysins (MMP-3, -10, -11), and membrane-type MMPs (MT-MMPs).[4] MMP activity is highly regulated at many levels and, under normal conditions it is controlled by endogenous inhibitors known as tissue inhibitors of MMPs (TIMPs).[5] Uncontrolled over-expression of MMPs can promote a variety of diseases including arthritis, tumour metastasis, multiple sclerosis and periodontal degradation. On this basis a large number of MMP inhibitors (MMPi) have been developed as potential therapeutic agents.[3,4-12] Many of the known MMPi contain a zinc-binding group (ZBG) linked to different scaffolds to ensure strong interactions within the co-factor-binding regions of these enzymes; among them the more developed and potent class of MMPi utilizes a hydroxamate as the ZBG.[13-21] Among the MMPs, the two gelatinases (MMP-2 and MMP-9), matrilysin (MMP-7), and MT1-MMP (MMP-14) play a significant role in certain key functions of tumour cells, facilitating metastatic tumour dispersion and angiogenesis, resistance to apoptosis and activation of EGF receptors.[22-28] In the past, some potent "broad spectrum" MMPi have been proposed and tested against tumours, but at present none of them are on the market.[29,30] In fact, many of these new molecules have determined a severe musculoskeletal syndrome, with fibroproliferative effects in the joint capsule of the

knee.[31-33] These effects are thought to be linked to an impairment of normal tissue remodelling governed by MMP-1 and/or by sheddases such as TNF- -convertase.[34] For these reasons, a lack of activity with respect to MMP-1 is considered to be an important factor in reducing some of the side effects found for "non-selective" MMPi.[35] The recent development of synthetic MMPi possessing a good potency and selectivity towards the two gelatinases, together with the discovery that some of these molecules active on MMP-2 show important pro-apoptotic effects on tumour cell cultures, confirm the validity of their use as potential anti-tumour agents.[3,7-9,13-21,36-38] Nowadays new compounds possessing a selective inhibitory activity on MMPs that are over-expressed in tumours are in great demand, with particular reference to viability control and invasiveness of the cancer cell, and many groups and pharmaceutical companies all over the world are involved in improved studies on these new synthetic MMPi, not only to develop new drugs but also to discover new agents to be used as diagnostics in cancer.[39,40] As reported above, it has been demonstrated that the catalytic zinc ion in the active pocket of MMPs is directly involved in the degradation of extracellular matrix components. In the latent form of gelatinases, the active site presents the zinc bound to three histidine residues and blocked by the sulphydrylic group of a cysteine. During activation, the detachment of the propeptide also involves the dissociation of this residue, and the active site becomes accessible for the substrate.[41] A mechanism has been proposed for the catalytic activity, on the basis of structural information. In the first step, the zinc ion is tetrahedrally coordinated to three histidines and a water molecule; during proteolysis, the water donates a proton to Glu219 (1HFC[42] sequence number), which transfers it to the nitrogen of the scissile amide bond that is stabilized by Ala182; this is followed by the generation of a salt bridge between Glu219 and the free amine of the cleaved substrate.[43] As the zinc ion plays an important role in the substrate degradation, one of the possible strategies to develop new MMPi is based on the search for the best ZBG able to compete with the substrate in zinc coordination. Common ZBGs of MMPi include hydroxamates, carboxylates, thiolates, and phosphonates.[44-46] Among these the hydroxamate group has proved to be the most potent (100-2000 times compared with their carboxylate analogue).[45] It is known from crystallographic data that two zinc atoms are present in MMPs: a structural tetra-

coordinated atom, which is linked with three histidines and an aspartate with a tetrahedral structure, and a catalytic atom, which, in the presence of hydroxamate inhibitors, is linked to three histidines and the ligand with a penta-coordination in a bipyramidal trigonal geometry, considering that the hydroxamate group behaves like a bidentate ligand.[47-49] Some new N-arylsulfonyl-substituted alkoxyaminoaceto hydroxamic acid derivatives have recently been synthesized and tested, and proved to possess a good selectivity for MMP-2 over MMP-1;[50] these may be considered as promising leads for the development of new selective inhibitors of metalloproteinases. A rational planning of these inhibitors requires the availability of adequate molecular models, to make it possible to study the mechanism of the enzyme-ligand interaction. Resolution of the crystal structure of MMP-2 gave researchers the opportunity to develop new compounds by means of a structure-based approach; so far, however, no X-ray structures of the MMP-2-inhibitor complexes have been reported, and there is a real need for theoretical studies on the binding mode of MMP-2 with its inhibitors, since they can provide insight into the interaction occurring in the active site. In order to perform the docking of ligands with metalloproteinases, the use of computational methods has proved to be problematic, because of the restrictions imposed by the presence of the zinc ion. These applications are developed on the basis of an empirical force field, which cannot be used directly in the design of ZBGs of MMPs[51] and other metalloproteinases, given the lack of parameters relating to the metallic ions. The use of generalized parameters in the case of these ions is highly unadvisable, considering that the coordination bonds formed by these ions largely depend on the nature of the ligands. There are two ways to model the force field of this zinc ion: the bonded model and the non-bonded model. In the bonded model, the coordinates between zinc and ligand/MMP are described by the commonly used bonded terms, including bond stretching.[52-57] On the contrary, in the non-bonded approach, Van der Waals and non-bonded electrostatic terms are used to model the zinc-ligand/MMP interactions. The non-bonded method is very sensitive to the choice of electrostatic model, and can suffer from an inability to retain a low coordination number. Furthermore, with the AMBER force field, the non-bonded approach generally fails to give the correct coordination number, even when long-range electrostatic interactions are correctly accounted for using an infinite

cutoff.[55,58] For all these reasons, we adopted the bonded approach for the zinc ion representation, and therefore specific parameters for the zinc ion were to be added to the AMBER force field.[59] There are many works in the literature regarding the theoretical study of the zinc ion: in 1996 Ryde studied the coordination of the catalytic zinc ion in alcohol dehydrogenase through combined quantum-chemical and molecular mechanical calculations,[60] while in 1999 he wrote a theoretical study regarding carboxylate binding modes.[61] In 2002 Torrent et al. performed an ONIOM study of Methane Monooxygenase and Ribonucleotide Reductase,[62] and in 2002 Olsen et al. published a quantum mechanical study concerning the influence of a hydrogen-bonding network on β-Lactamase.[63] Regarding MMPs, Donini[55] and recently Rizzo[64] carried out respectively MM-PBSA and MM-GBSA calculations on certain ligands complexed with MMP-1, MMP-2 and MMP-3, using the zinc non-bonded model. As regards the zinc bonded model, Toba et al[58] reported molecular dynamics and free energy perturbation studies carried out on enzyme-inhibitor complexes of human fibroblast collagenases, using AMBER force field parameters coming from human carbonic anhydrase II studies.[65] These parameters were taken from a tetrahedral model, in which the tetra-coordinated zinc atom was linked with three histidine residues and one water molecule or hydroxide ion; in other cases, the crystal structure of hydroxamate inhibitors complexed with MMPs revealed that the catalytic zinc was penta-coordinated with three histidine nitrogens in MMPs and two hydroxamate oxygens in inhibitors. Consequently, specific force field parameters to study MMPs and their complexes with hydroxamate inhibitors needed to be calculated by quantum mechanical *ab initio* methods. In this work, we calculated the force field parameters for the structural zinc and the catalytic zinc complexed with hydroxamate inhibitors, and we subsequently tested their efficacy. Furthermore, we studied the docking of some known inhibitors with MMP-1 and -2 and from an analysis of the results, we designed a new MMP-2 active and selective inhibitor.

2.1.2 Results and Discussion

2.1.2.1 Parameters calculations and testing. Starting from the MMP-1 structure derived from experimental crystallographic studies (2TCL66), we calculated the geometry, the partial charges and the force constant parameters concerning for the two

zincs of the metalloproteinases, using the methods reported in the Experimental section. Table 1 shows all the parameters obtained using this procedure.

In order to test the parameters obtained and to verify whether they could be used for molecular dynamics (MD) simulations, we carried out 1 ns of MD on the X-ray structure of MMP-1 (1HFC42) complexed with HAP (see Experimental section for details) using our parameters. As shown in Figure 1A, after 140 ps of MD, the system reached an equilibrium, since the total energy for the last 900 ps remained constant. Analysing the RMS deviation from the X-ray structure of all the heavy atoms of the complex, we observed that after an initial increase, in the last 700 ps the RMSD remained approximately constant around the value of 1 Å (see Figure 1B), suggesting that our MD procedure was correct. As regards the geometry of the sites surrounding the two zincs, during MD we analyzed the RMS deviation from the X-ray structure of the position of the two zincs and the heavy atoms of the groups that were bound to the zincs (three imidazole and the hydroxamic fragment for the catalytic zinc and three imidazoles and the carboxylic fragment for the structural zinc).

As shown in Figure 1C and 1D, after the first 50 ps of MD, both systems were stabilized with a good RMSD: the average of the RMSD was 0.54 Å for the catalytic zinc binding site and 0.53 Å for the structural one. In order to verify whether using our parameters, the zinc bonded model gave a better geometry than the zinc non-bonded model, we carried out a second simulation on the MMP-1 X-ray structure, applying the same procedure seen above, but considering the two zincs as non-bonded. Figures 1C and 1D showed that considering the two zinc atoms as non-bonded, the RMSD became worse, exceeding the value of 1 Å in some steps, and in both sites the RMSD average of MD was 0.84 Å. Thus these results confirmed that the use of the zinc non-bonded model determined a worse geometry of the zinc binding site compared with the one obtained using the bonded model.

Table 1. Parameters calculated for the structural and catalytic zinc binding sites

| Structural Zn binding site			Catalytic Zn binding site		
Interaction	**Bond length**	**Force**	**Interaction**	**Bond length**	**Force**
Z1–N*	2.0766	99.00	Z2-N2	2.1051	80.00
Z1-O	2.0192	132.00	Z2-NT	2.1619	70.00
			Z2-O1	2.1070	95.00
			Z2-O2	2.1111	103.50
			O1-N2	1.4061	520.00
Interaction	**Angle**	**Force**	**Interaction**	**Angle**	**Force**
N*-Z1-N*	111.0000	10.00	N2-Z2-N2	111.8620	8.00
C-N*-Z1	126.0000	27.00	N2-Z2-NT	97.9730	8.00
C-O-Z1	108.3030	39.00	N2-Z2-O2	123.3505	8.75
O-Z1-N2	112.0000	11.55	N2-Z2-O1	90.2945	9.50
O-Z1-NU	95.3530	11.55	NT-Z2-O2	86.7380	8.25
			NT-Z2-O1	165.1020	12.85
			O1-Z2-O2	78.3650	9.90
			C-N*-Z2	126.0000	26.50
			N2-O1-Z2	111.8140	67.00
			Z2-O2-C	110.5900	33.00
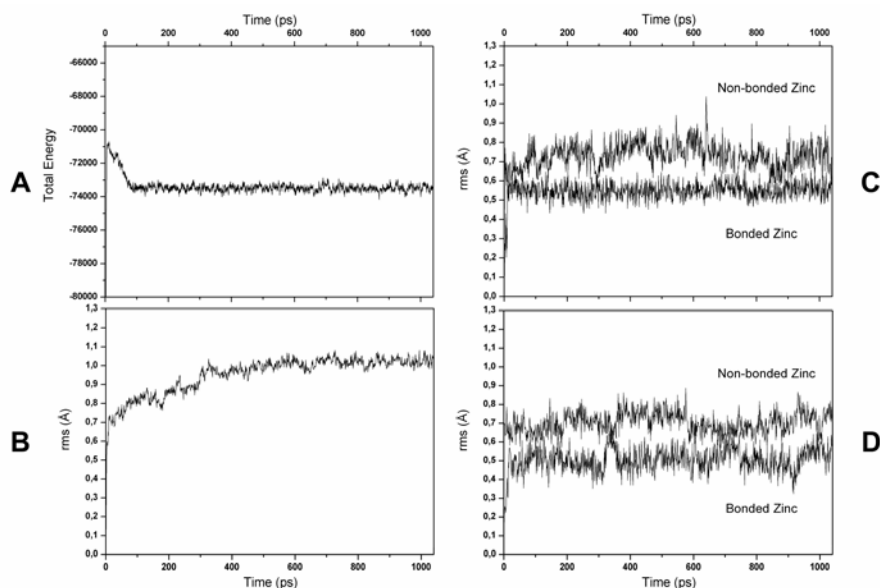
Figure 1. Analysis of the MD simulation of HAP complexed with MMP-1. (a) Total energy (Kcal/mol) of the system plotted vs time, (b) root-mean-square deviation (rmsd) in angstroms (Å) between the system and the X-ray starting structure for all the heavy atoms, (c) rmsd between all the heavy atoms of the catalytic zinc binding site and the X-ray starting structure for the bonded zinc system and non-bonded zinc system (d)) rmsd between all the heavy atoms of the structural zinc binding site and the X-ray starting structure for the bonded zinc system and non-bonded zinc system

From the X-ray structure of the complex, we observed that the ligand formed six H bonds with the MMP-1. As shown in Figure 2, HAP interacts with Gly179, Leu181, Ala182, Glu219, Pro238 and Tyr240 (1HFC[42] sequence number). The complex obtained by minimizing the average structure of the last 900 ps of MD showed that all the six H bonds were still present and that the ligand position was very similar to the X-ray one, since the RMSD between the heavy atoms of the ligand was 1.25 Å. Moreover, this value was principally determined by the different position of the phenyl substituent: the RMSD of the heavy atoms of the ligand, excluding those of the phenyl ring, was 0.36 Å (see Figure 3).
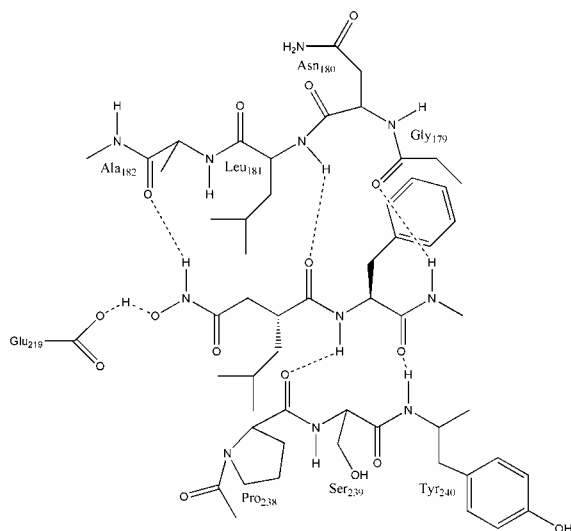
Figure 2. Interaction scheme of HAP complexed with MMP-1.
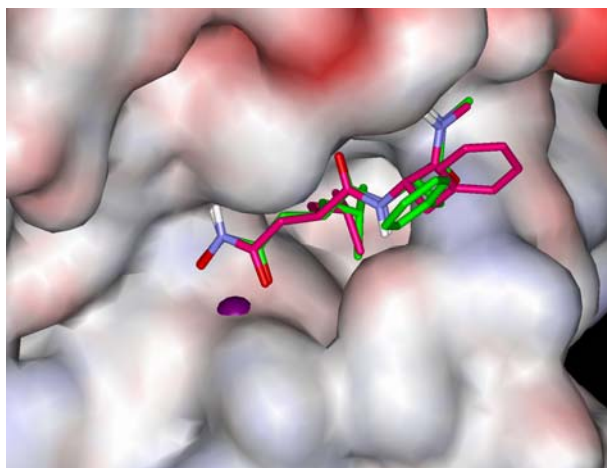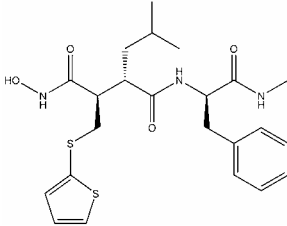
Figure 3. Binding position of HAP into MMP-1 after MD compared with the starting X-ray one.

All these analyses proved that the new parameters calculated and our dynamics procedure were reliable and could be used for further docking studies.

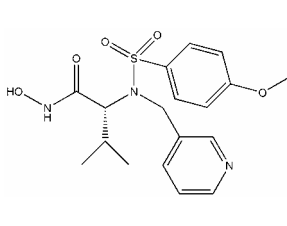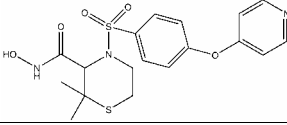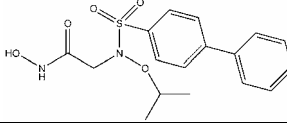**2.1.1.2 Docking studies.** We docked Batimastat, CGS 27023A, Prinomastat and compound **a**[50] (see Table 2) into the MMP-1 and MMP-2 using the same procedure described above.

Table 2. Ligands used for the docking studies and their MMP-1, -2 inhibitory activity and MMP-1/MMP-2 selectivity.

Batimastat		CGS 27023A	
IC$_{50}$(nM)[7]		IC$_{50}$(nM)[7]	
MMP-1	**MMP-2**	**MMP-1**	**MMP-2**
10	4	33	20
Ratio	2.5	**Ratio**	1.7
Prinomastat		Compound a	
IC50(nM)[7]		IC50(nM)[7]	
MMP1	**MMP2**	**MMP1**	**MMP2**
8	0.08	12460	12.4
Ratio	100	**Ratio**	1004

As regards the initial placement of the ligands, the position of Batimastat was very similar to that of HAP, while the other three ligands were introduced inside the binding site in such a manner that they could interact with Leu181, Ala182, Glu219 and with the S1' pocket, in agreement with the experimental data relative to structurally correlated MMP-ligand complexes (see Experimental section).

2.1.2.3 MMP-1 complexes. At the end of the simulation, Batimastat presented all the interactions shown by HAP, with the formation of H bonds with Gly179, Leu181, Ala182, Glu219, Pro238 and Tyr240 (see Figure 5A). The IC$_{50}$ values of the two inhibitors confirmed our data as they had a very similar inhibition potency (in MMP-1 HAP showed an IC$_{50}$ of 7 nM, and Batimastat 10 nM). As regards CGS 27023A, the pyridine substituent interacted in the S2' site while the isopropyl group was directed towards the His183 of the structural zinc binding site. The hydroxamate formed H

bonds with Ala182 and Glu219, while one of the oxygen atoms of the sulfonamido group interacted with Leu181 and Ala182, forming two H bonds, and the *p*-methoxyphenyl group was inserted into the S1' pocket (see Figure 5B). As is already known, the MMP-1 S1' pocket is small and closed by the non-conserved residue Arg214 (leucine in MMP-2). Although this residue shows conformational flexibility, it is known that the low inhibition potency of many inhibitors against MMP-1 is often due to the presence of large P1' groups that would require large conformational changes in order to interact with the S1' pocket.[67] Also in our simulation, the presence of the p-methoxyphenyl group in the S1' pocket determined a movement of the Arg214: in the initial structure, this residue was stabilized by four H bonds, three of them with Leu235 and Thr241 (two H bonds), and a fourth intramolecular H bond between the carboxy and amino groups of the same Arg214 (see Figure 4A); during the simulation, we observed the rotation of the δ torsional of the arginine with the disappearance of the intramolecular H bond and the H bond with Leu235, thus determining an increase in the pocket depth (see Figure 4)
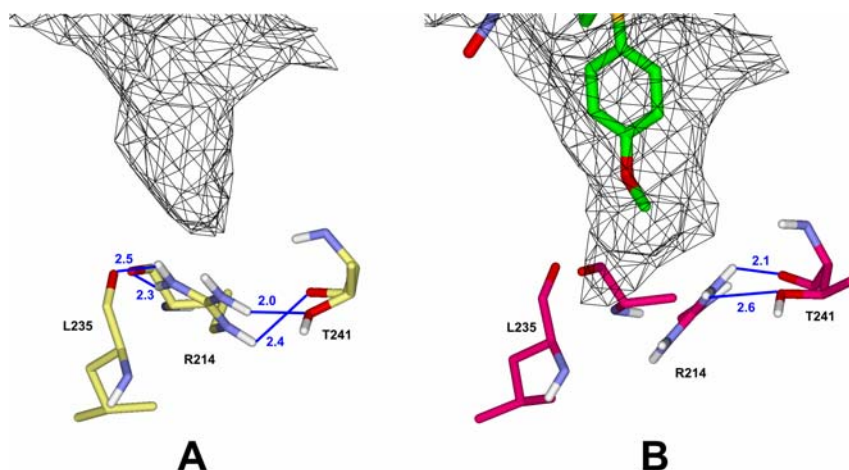

Figure 4. Analysis of the S1' cavity in the X-ray structure (A) and after modelling with CGS 27023A (B). The presence of the *p*-methoxyphenyl group determines the movement of Arg214 and an increase in cavity depth.

This movement of Arg214 was not observed during the simulation of MMP-1 complexed with Batimastat and HAP, confirming that the conformational adjustment

was due to the interaction of the *p*-methoxyphenyl substituent. The position of Prinomastat in MMP-1 was similar to that of CGS 27023A but there were some differences: the hydroxamate formed an H bond with Glu219, losing the interaction with Ala182, and one of the oxygen atoms of the sulfonamido group interacted only with the nitrogen of Leu181, losing the interaction with Ala182 (see Figure 5C). As regards the S1' pocket, the movement of Arg214 was very similar to that seen in the CGS 27023A-MMP-1 simulation. We observed the rotation of the δ torsional of the arginine and the loss of all the H bonds except those with Thr241, but in this complex, the nitrogen of the P1' group of Prinomastat formed an H bond with Thr241. As regards compound **a**, it was unable to penetrate with the P1' substituent inside the S1' pocket completely, thus determining a major solvent exposure of the ligand and low interactions with the MMP-1 binding site, including a bad coordination of the hydroxamate with the zinc atom. As shown in Figure 4D, after about 280 ps of MD, the ligand lost the two H bonds of the oxygen atoms of the sulfonamido chain with Leu181 and Ala182, and only the interaction of the hydroxamate with Glu219 and Ala182 remained. These simulations indicated that Batimastat showed a good inhibition potency, since it presented six H bonds with the protein. CGS 27023A showed four H bonds, but in order to interact in the S1' pocket, the P1' group determined a conformational adjustment of Arg214, with a consequent decrease in inhibition potency. In the MMP-1-Prinomastat complex, we observed the same movement of the arginine, with the formation of a new H bond in the S1' pocket, however the large dimensions of the P1' substituent caused its incomplete penetration into the S1' pocket, determining as a consequence the loss of the important interactions with Ala181. Compound **a** proved to be practically inactive in the MMP-1; this was probably due to the presence of the biphenyl group, since this substituent was not able to interact in the S1' pocket, and this fact determined the loss of the interaction of the oxygen atoms of the sulfonamido group with Ala181 and Leu182.
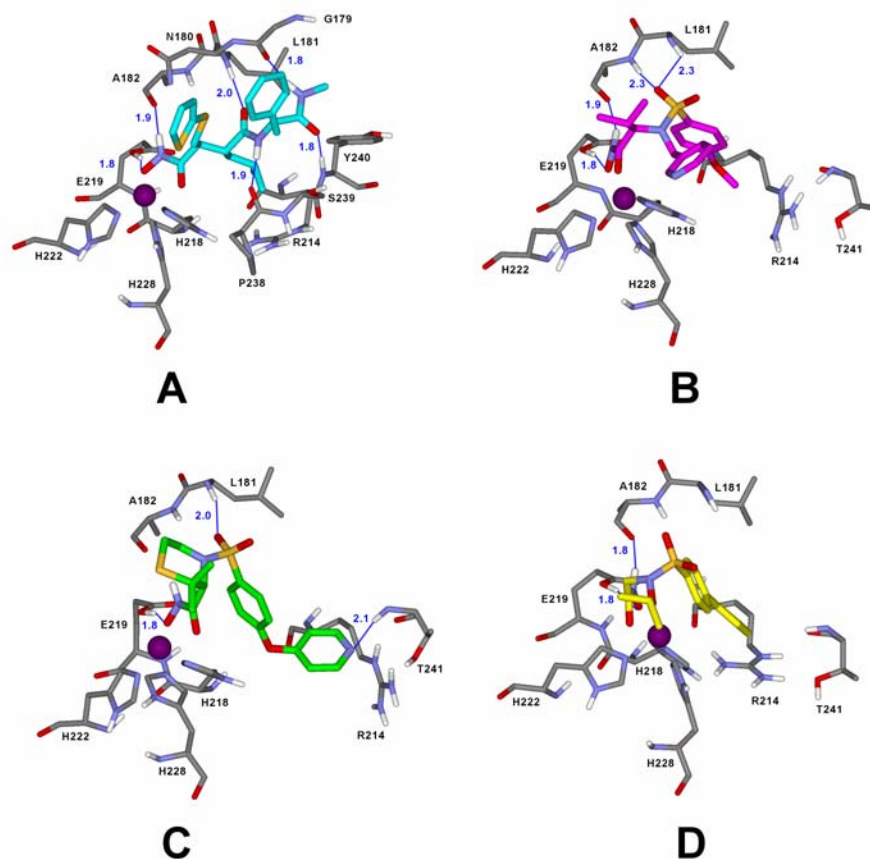
Figure 5. Docking of Batimastat (A), CGS 27023A (B), Prinomastat (C) and compound **a** (D) into MMP-1. The main interatomic distances are reported in blue, all distances are in Angstroms.

2.1.2.4 MMP2 complexes. The docking of Batimastat into MMP-2 showed all the same H bonds that were also present in the MMP-1 complex, thus confirming the low selectivity of this ligand. As in the MMP-1, CGS 27023A showed the pyridine substituent in the S2' site, while the isopropyl group was directed towards the His166 (1QIB[68] sequence number) of the structural zinc binding site. The hydroxamate formed H bonds with Glu202 and Ala165, while the sulfonamido group interacted with Leu164 and Ala165, forming two H bonds, and the *p*-methoxyphenyl group interacted in the S1' pocket. Prinomastat presented a disposition very similar to that of CGS 27023A: with the hydroxamate it formed two H bonds with Glu202 and Ala165, while the

sulfonamido group interacted with Leu164 and Ala165. As regards the P1' substituent, it was inserted into the S1' pocket and formed an H bond with Thr227. Similarly, compound **a** showed the four H bonds with Glu202, Ala165 and Leu164, and the biphenyl system was inserted into the S1' pocket.

2.1.2.5 Design of new compounds. An analysis of the eight complexes suggested that Batimastat showed a good inhibition potency because of the presence of several H bonds; however, these interactions were maintained in both proteins, and for this reason, this ligand did not show any selectivity. On the contrary CGS 27023A showed a low level of selectivity, which was probably due to the conformational change that occurred in the MMP-1 pocket to allow the penetration of the *p*-methoxyphenyl group. Prinomastat presented a large P1'substituent, which, although forming two H bonds, determined a conformational adjustment of the ligand in the MMP-1, with the loss of the two crucial interactions with Leu181 and Ala182; this fact could explain its good MMP-2/MMP-1 selectivity. Compound **a** was found to be the most selective ligand of the four analyzed, and this fact could be due to the presence of a large P1' group, rigid and unable to form H bonds, or to interact with the MMP-1-S1' pocket. This analysis suggested that in order to maintaining the MMP-2/MMP-1 selectivity, the presence of the biphenyl group as a P1' substituent could be very useful. In order to improve the MMP-2-inhibitor potency of compound a, maintaining its MMP-2/MMP-1 selectivity, an analysis of the binding site characteristics could be decisive. For this reason, by means of the GRID program,[69] we examined the molecular interactions fields (MIFs) obtained for different probes for the MMP-2. In particular, the examination of the DRY, C1= and C3 probes revealed that besides the lipophilic pocket, corresponding to the S1' pocket, there was a second pocket in which a lipophilic interaction was favourable, and this pocket corresponded to the zone occupied by the CGS 27023A isopropyl group. For this reason, we hypothesized that the design of an analogue of compound **a** with an isopropyl substituent on the alpha carbon atom might be able to increase the MMP-2 inhibition (see Figure 6), while maintaining a low MMP-1 inhibitory potency, due to the biphenyl group interaction. The substitution of the C alpha determined the formation of a chiral carbon; Figure 7 shows the docking into MMP-2 of the two enantiomers of the designed ligand. In both complexes, the inhibitors formed all the H bonds shown by

compound **a**, but unlike (*R*)-*N*-hydroxy-2-(*N*-isopropoxybiphenyl-4-ylsulfonamido)-3-methylbutanamide (compound (*R*)-**b**), the *S* enantiomer (compound (*S*)-**b**) presented the isopropyl group directed towards the core of MMP-2, and with this disposition the substituent was unable to interact in the region defined by GRID, leading to worse interactions with the protein.
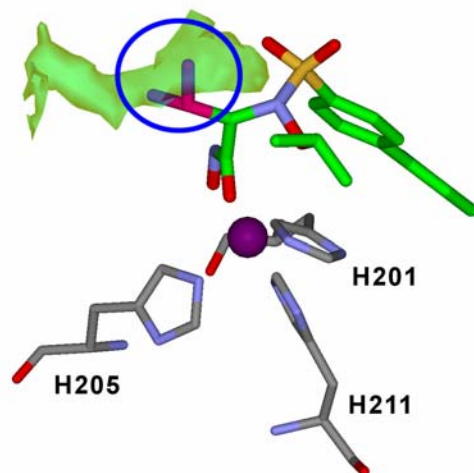
Figure 6. GRID analysis of the MMP-2/compound (*R*)-**b** complex. The green surface indicates the lipophilic area highlighted by using the DRY probe; the isopropyl substituent which could interact in this zone is coloured magenta.
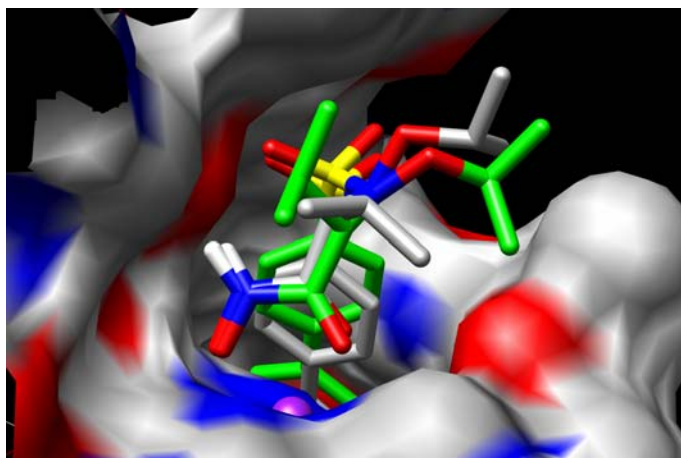

Figure 7. Superimposition of (*R*)-**b** (grey) and (*S*)-**b** (green) docked into MMP-2. Unlike (*R*)-**b**, the isopropyl substituent of (*S*)-**b** is directed towards the core of the metalloprotein.

2.1.2.6 Biological results. Table 3 shows the inhibitory indices (IC_{50}) towards MMP-1 and MMP-2 of the *N*-hydroxy-2-(*N*-isopropoxybiphenyl-4-ylsulfonamido)-3-methylbutanamides (*(R/S)*-b, (*R*)-b and *(S)*-b), compared with those of the previously described *N*-hydroxy-2-(*N*-isopropoxybiphenyl-4-ylsulfonamido)acetamide (compound a). Selectivity indices for MMP-2 over MMP-1, reported in parentheses, are expressed as ratios of their inhibitory indices. On MMP-2, the substitution on the P1 site with an *i*-Pr group on the carbon atom alpha to the hydroxamate improves the inhibitory potency about five times, passing from compound a (IC_{50} = 12.4 nM) to racemic (*R,S*)-b (IC_{50} = 2.44 nM). The (*R*)-b enantiomer shows an improvement of the potency on MMP-2 of threefold (IC_{50} = 0.81 nM) compared with its racemate (*R,S*)-b and about nineteen times, compared with its dystomer (*S*)-b (IC_{50} = 15.2 nM). An analysis of the selectivity indices for the *N*-hydroxy-2-(*N*-isopropoxybiphenyl-4-ylsulfonamido)-3-methylbutanamides (*(R/S)*-b, *(R)-b* and *(S)*-b) and the alpha un-substituted compound *N*-hydroxy-2-(*N*-isopropoxybiphenyl-4-ylsulfonamido)acetamide (compound a) indicates that the newly synthesised P1 substituted (*i*-Pr group) compounds maintain a good selectivity profile. As can be seen, the MMP-1/MMP-2 ratios are range from 241.8 for the racemic *(R/S)*-b to 600 for the eutomer *(R)*-b respectively. A comparison of these selectivity indices with that of the previous *N*-hydroxy-2-(*N*-isopropoxybiphenyl-4-ylsulfonamido)acetamide (compound a) indicates a reduction of about two times of the selectivity ratio MMP-1/MMP-2 passing from the new, most potent MMP-2 inhibitor, (*R*)-b to compound a. In every case, a comparison between the new in vitro potent antiangiogenic agent MMP-2 inhibitor (*R*)-b[1] and the reference drugs used, indicates that this compound is 24.7 times more potent on MMP-2 than the reference drug CGS 27023A. Moreover (*R*)-b proves to be 352.9 and 6 times more selective than CGS 27023A and Prinomastat respectively.

Table 3. Inhibitory activity of *N*-hydroxy-2-(*N*-isopropoxybiphenyl-4-ylsulfonamido)acetamide (compound **a**) and *N*-hydroxy-2-(*N*-isopropoxybiphenyl-4-ylsulfonamido)-3-methylbutanamides ((*R/S*)-**b**, (*R*)-**b** and (*S*)-**b**) (showed in figure) towards MMP-1 and MMP-2.

Compd	IC$_{50}$ (nM)		MMP-1/MMP-2 ratio
	MMP-1[a]	MMP-2[a]	
a	12460±960	12.4±1.5	1004.9
(*R/S*)-b	590±25	2.44±0.28	241.8
(*R*)-b	486±76	0.81±0.22	600
(*S*)-b	1199 ±52	15.2±1.3	78.9

[a] Mca-Lys-Pro-Leu-Gly-Leu-Dap(Dnp)-Ala-Arg-NH$_2$ and Mca-Pro-Leu-Gly-Leu-Dpa-Ala-Arg-NH$_2$ have been used as substrates for MMP-1 and MMP-2 respectively; [b] selectivity indices are reported as MMP-1/MMP-2 IC$_{50}$ ratios.

2.1.3 Conclusions.

In a rational planning of the new biphenylsulfonamido- based hydroxamate derivatives recently studied by us, which possess a good selectivity for MMP-2 over MMP-1, we obtained the parameters for studying MMPs and their complexes with hydroxamate inhibitors by means of the AMBER force field. This modified force field was able to simulate the geometrical data coming from X-ray structures, thus making it possible to carry out reliable molecular modelling studies on hydroxamate inhibitor-MMP complexes. On the basis of the docking studies, the two enantiomers of N-hydroxy-2-(N-isopropoxybiphenyl-4-ylsulfonamido)-3-methylbutanamide, (R)-b and (S)-b, were designed and synthesised, as more potent MMP-2 inhibitors than our previous compound a. An analysis of the inhibition profile on MMP-1 and MMP-2 of these two inhibitors shows that the eutomer (R)-b is 24.7 times and 15.3 times more potent than CGS 27023A and the parent compound a on MMP-2, maintaining a higher index of MMP-2/MMP-1 selectivity compared with the CGS 27023A and Prinomastat. On this basis, the hydroxamate (R)-b can be considered as a progenitor of a new class of biphenylsulfonamido-based inhibitors that differ from a in the presence of an alkyl side

chain in the P1' position, and show a good potency and selectivity profile on the two MMPs studied.

2.1.4 Experimental section.

2.1.4.1 Parameters Calculations. In order to obtain all the parameters regarding the structural and catalytic zinc, it was necessary to optimize the geometry of the two systems, calculate the partial charges, and then estimate and optimize the missing Force constants.

2.1.4.1.1 Geometry optimization. Starting from the geometry coming from the experimental crystallographic structure of MMP-1,[66] we constructed two models referring to the two zinc ions with their chemical surroundings. The model referring to the structural zinc was made up of three methylimidazoles and one acetate ion, while the one referring to the catalytic zinc was made up of three methylimidazoles and one acetylhydroxamate ion (see figures of Table 1 where R is replaced by methyl). These models were subjected to a full geometry optimization by means of quantum mechanical calculations based on the Gaussian98 programme,[70] using the B3LYP chemical model with a Lanl2DZ basis set, a direct SCF calculation and an SCF convergence criterion to 10-5. The B3LYP chemical model has been shown to be an accurate density functional method,[71] and it gives as good or better geometries and energies as MP2 ab initio methods for first-row transition metal complexes.[72] The B3LYP model is a combination of the Becke three-parameter hybrid functional[73] with the Lee-Yang-Parr correlation functional (which also includes density gradient terms).[74,75] As regards the basis set, we used Lanl2DZ; this means that the zinc atom is described through the Los Alamos non-relativistic electron core potential of Hay and Wadt[76] and an essentially double-zeta basis set including 3d orbitals and 3d diffuse functions for the valence shell. In this basis set, the rest of the atoms are described through the Dunning-Huzinaga full double- ζ basis set.[77] This method has already been tested and has been found to be suitable to deal with systems containing metal atoms, and in particular zinc.[62,78]

2.1.4.1.2 Partial charges. The next step was to determine the charges to be assigned to the zinc, to its chemical surroundings, and, lastly, to the ligands. The method used was an analysis of the RESP[79] (Restrained Electrostatic Potential fit); the basic idea behind

this charge-fitting algorithm is that the charges on non-hydrogen atoms are restrained to an "optimal" value of zero. The RESP charges show less conformational variability than the standard ESP charges.

2.1.4.1.3 Force Constants. A frequency analysis was carried out on the two models, and the diagonal elements of the Hessian matrix, calculated in their internal coordinates, were used as the value of the stretching constants. The bending force constants were approximated to one tenth of the relative stretching constants. The torsional force constants for zinc were not taken into consideration, since the geometry refers to a structure which is quite rigid and devoid of any significant torsional freedom.[65] Given that molecular mechanics not only takes into consideration the binding terms, but also the non-binding ones (Coulombian and van der Waals interactions), unlike quantum mechanics which determines the total energy, it was necessary to test the values of the force constants, obtained by means of frequency calculations, within the AMBER force field, and, where necessary, to adjust these values. For this reason, several conformations were taken into consideration, in which the stretching and bending values were varied around the equilibrium value. Quantum mechanical calculations were carried out on these models by the B3LYP/LanL2DZ method, together with energy calculations based on the AMBER force field with the parameters calculated for the zinc ion. Only small adjustments were required to the force constants obtained from the frequency calculation, in order to obtain the best agreement between the energy values calculated by the two methods. In fact the modified AMBER force field after optimization was able to calculate the conformational energy of the model with a very good agreement with respect to the DFT-B3LYP-LanL2DZ method and, for the different parameters, the RMS of the difference between the relative energies (Kcal/mol) calculated by means of quantum mechanics (B3LYP/LanL2DZ) and those calculated by means of molecular mechanics (AMBER) showed an average value of 0.85 Å.

2.1.4.2 MD Simulations. All simulations were performed using AMBER 8.0.[80] MD simulations were carried out using the modified parm94 force field at 300 K. An explicit solvent model TIP3P water was used and the complexes were solvated with a 15 Å water cap. Sodium ions were added as counterions to neutralize the system. Prior

to MD simulations, two steps of minimization were carried out; in the first stage, we kept the protein fixed with a constraint of 500 Kcal/mol and we just minimized the positions of the water molecules; then in the second stage, we minimized the entire system applying a constraint of 20 Kcal/mol on the α Carbon. The two minimization stages consisted of 5000 steps in which the first 1000 were Steepest Descent (SD) and the last 4000 Conjugate Gradient (CG). Molecular dynamics trajectories were run using the minimized structure as a starting input, and the particle mesh Ewald (PME) algorithm was used for dealing with long-range interactions.[81] The time step of the simulations was 2.0 fs with a cutoff of 12 Å for the non-bonded interaction and SHAKE was employed to keep all bonds involving hydrogen atoms rigid. Constant-volume was carried out for 40 ps, during which the temperature was raised from 0 to 300 K (using the Langevin dynamics method); then 600 ps (1 ns for the X ray complex) of constant-pressure MD were carried out at 300 K. All the α Carbons not within 15 Å of the catalytic zinc were blocked with a harmonic force constant of 20 Kcal/mol•Å. The final structure of the complexes was obtained as the average of the last 500 ps of MD minimized with the CG method until a convergence of 0.05 Kcal/ Å•mol.

2.1.4.3 Docking of the Ligands. The four inhibitors were placed into MMP-1 (1HFC[42]) and MMP-2 (1QIB[68]) using as reference positions those of the MMP complexes determined by X-Ray and NMR methods deposited at the Protein Databank.[82] For Batimastat, we superimposed the binding site of MMP-1 and MMP-2 on the MMP-8-Batimastat complex (1MMB[83]). As regards Prinomastat, we used the position of a diphenyl-ether sulphone-based hydroxamic acid complexed with the MMP-1 (966C[84]) as a template, while for CGS 27023A and compound a we used the NMR study carried out on MMP-2 complexed with a hydroxamic acid inhibitor (1HOV[85]).

2.1.4.4 GRID Calculations. The GRID box dimensions were chosen to encompass all the important parts of the active site. The grid spacing was set to 1 Å, and the molecular interaction fields (MIFs) were calculated for the DRY (hydrophobic probe), C1= (sp2 CH aromatic or vinyl), C3 (methyl CH3 group), N2 (neutral flat NH2, e. g., amide), O (sp2 carbonyl oxygen) and OH (phenol or carboxy OH) probes using the GRID program, version 22.[69]

2.2. REFERENCES.

1. Rossello, A.; Nuti, E.; Carelli, P.; Orlandini, E.; Macchia, M.; Nencetti, S.; Zandomeneghi, M.; Balzano, F.; Uccello-Barretta, G.; Albini, A.; Benelli, R.; Cercignani, G.; Murphy, G.; Balsamo, A. Bioorg. & Med. Chem. Lett. 15 (2005) 1321-1326.

2. http://www.clip.ubc.ca/

3. Smith, H. J.; Simons, C. In Proteinase Peptidase Inhibition: Recent Potential Targets for Drug Development; Taylor & Francis: London & New York, 2002.

4. Folgueras, A. R.; Pendas, A. M.; Sanchez, L. M.; Lopez-Otin, C. Int. J. Dev. Biol. 48 (2004) 411-424.

5. Sternlicht, M. D.; Werb, Z. Ann. Rev. Cell. Dev. Biol. 17 (2001) 463-516.

6. Woessner, J.F. jr.; Nagase, H. Oxford University Press, Oxford. 2000; pp. 1-223.

7. Supuran, C. T.; Scozzafava, A. In Proteinase and Peptidase Inhibition: Recent Potential Targets for Drug Development; Smith, H. J.; Simons, C., Eds.; Taylor & Francis: London & New York, 2002; pp 35-61.

8. Whittaker, M.; Floyd, C. D.; Brown, P.; Gearing, A. J. H. Chem. Rev. 99 (1999) 2735-2776.

9. Supuran, C. T.; Casini, A.; Scozzafava, A. Med. Res Rev. 23 (2003) 535-558.

10. Coussens, L. M.; Fingleton, B.; Matrisian, L. Science. 295 (2002) 2387-2392.

11. Purcell, W. T.; Ettinger, D. S. Current Oncology Reports. 5 (2003) 114-125.

12. Wielockx, B.; Libert, C.; Wilson, C. Cytokine & Growth Factor Reviews. 15 (2004) 111–115.

13. Wada, C. K. Curr Top Med Chem. 4 (2004) 1255-1267.

14. Hanessian, S. Curr Top Med Chem. 4 (2004) 1269-1287.

15. Puerta, D. T.; Cohen, S. M. Curr Top Med Chem. 4 (2004) 1551-1573.

16. Rush, T. S. 3rd; Powers, R. Curr Top Med Chem. 4 (2004) 1311-1327.

17. Skiles, J. W.; Gonnella, N. C.; Jeng, A. Y. Curr Med Chem. 11 (2004) 2911-2977.

18. Puerta, D. T.; Lewis, J. A.; Cohen, S. M. J. Am. Chem. Soc. 126 (2004) 8388-8389.

19. Rao, B. G. Curr Pharm Des. 11 (2005) 295-322.

20. Maskos, K. Biochimie 87 (2005) 249-263.

21. Cuniasse, P.; Devel, L.; Makaritis, A.; Beau, F.; Georgiadis, D.; Matziari, M.; Yiotakis, A.; Dive, V. Biochimie 87 (2005) 393-402.

22. Itoh, Y.; Nagase, H. Essays in Biochem. 38 (2002) 21-36.

23. Sounni, N. E.; Janssen, M.; Foidart, J. M.; Noel, A. Matrix Biol. 22 (2003) 55-61.

24. Sato, H.; Takino, T.; Miyamori, H. Cancer Sci. 96 (2005) 212-217.

25. Cowan, K. N.; Jones, P. L.; Rabinovitch, M. Circ. Res. 84 (1999) 1223-1233.

26. Jones, P. L.; Crack, J.; Rabinovitch, M. J. Cell. Block. (1997) 279-293.

27. Eguchi, P. J.; Dempsey, G. D.; Frank, G. D.; Motley, D.; Inagami, T. J. Biol. Chem. 276 (2001) 7957-7962.

28. Ahonen, M.; Poukkula, M.; Baker, A. H.; Kashiwagi, M.; Nagase, H.; Eriksson, J. E.; Kähäri V. M. Oncogene 22 (2003) 2121-2134.

29. Coussens, L. M.; Fingleton, B.; Matrisian, L. Science 295 (2002) 2387-2392.

30. Borkakoti, N. Biochem. Soc. Trans. 32 (2004) 17-19.

31. Hutchinson, J. W.; Tierney, G. M.; Parson, S. L.; Davis, T. R. C. J. Bone and Joint Surgery 80 (1998) 907-908.

32. Holmbeck, K.; Bianco, P.; Caterina, J.; Yamada, S.; Kromer, M.; Kuznetsov, S. A.; Mankani, M.; Robey, P. G.; Poole, A. R.; Pidoux, I.; Ward, J. M.; Birkedal-Hansen, H. Cell 99 (1999) 81-92.

33. Steward, W. P. Cancer. Chemother. Pharmacol. 43 (1999) 56-60.

34. Dahlberg, L.; Billinghurst, R. C.; Manner, P.; Nelson, F.; Webb, G.; Ionescu, M.; Reiner, A.; Tanzer, M.; Zukor, D.; Chen, J.;Van Wart, H. E.; Poole, A. R. Arthritis & Rheumatism 43 (2000) 673-682.

35. Scatena, R. Exp. Opin. Invest. Drugs 9 (2000) 2159-2165.

36. Wada, C. K.; Holms, J. H.; Curtin, M. L.; Dai, Y.; FlorJancic, A. S.; Garland, R. B.; Guo, Y.; Heyman, H. R.; Stacey, J. R.; Steinman, D. H.; Albert, D. H.; Bouska, J. J.; Elmore, I. N.; Goodfellow, C. L.; Marcotte, P. A.; Tapang, P.; Morgan, D. W.; Michaelides, M. R.; Davidsen, S. K. J. Med. Chem. 45 (2002) 219-232.

37. Zook, S. E.; Dagnino, R. Jr.; Deason, M. E.; Bender, S. L.; Melnick, M. J.WO Patent 97/20824, 1997; 127, 108945. Chem. Abstr. 123 (1995) 2870.

38. Kruger, A.; Arlt, M. J.; Gerg, M.; Kopitz, C.; Bernardo, M. .M.; Chang, M.; Mobashery, S.; Fridman, R. Cancer Res. 65 (2005) 3523-3526.

39. Turpeenniemi-Hujanen T. Biochimie 87 (2005) 287-297.

40. Haubner, R.; Wester, H. J. Curr Pharm Des. 10 (2004) 1439-1455.

41. Morgunova, E.; Tuuttila, A.; Bergmann, U.; Isupov, M.; Lindqvist, Y.; Schneider, G.; Tryggvason, K. Science 284 (1999) 1667-1670.

42. Spurlino, J. C.; Smallwood, A. M.; Carlton, D. D.; Banks, T. M.; Vavra, K. J.; Johnson, J. S.; Cook, E. R., Falvo, J.; Wahl, R. C.; Pulvino, T. A.; Wendoloski, J. J.; Smith, D. L. Proteins 19 (1994) 98-109.

43. Leung, D.; Abbenante, G.; Fairlie, D. P. J. Med. Chem. 72 (2000) 305-341.

44. Whittaker, M.; Floyd, C. D.; Brown, P.; Gearing, A. J. H. Chem. Rev. 99 (1999) 2735-2776.

45. Babine, R. E.; Bender, S. L. Chem. Rev. 97 (1997) 1359-1472.

46. Coussens, L. M.; Fingleton, B.; Matrisian, L. M. Science 295 (2002) 2387–2390.

47. Oreola, A.; Donini, T.; Kollman, P. A. J. Med. Chem. 43 (2000) 4180-4188.

48. Browner, M. F.; Smith, W. W. Biochemistry 34 (1995) 6602-6610.

49. Rowsell, S.; Hawatin, P.; Minshull, C. A.; Jepson, H.; Brockbank, S. M. V.; Barratt, D. G.; Slater, A. M.; McPheat, W. L.; Waterson, D.; Henney A. M.; Pauptit, R. A. J. Mol. Biol. 319 (2002) 173-181.

50. Rossello, A.; Nuti, E.; Orlandini, E.; Carelli, P.; Rapposelli, S.; Macchia, M.; Minutolo, F.; Carbonaro, L.; Albini, A.; Benelli, R.; Cercignani, G.; Murphy, G.; Balsamo, A. Bioorg. Med. Chem. 12 (2004) 2441-2450.

51. Cheng, F.; Zhang, R.; Luo, X.; Shen, J.; Li, X.; Gu, J.; Zhu, W.; Shen, J.; Sagi, I.; Ji, R.; Chen, K.; Jiang H. J. Phys. Chem. B 106 (2002) 4552-4559.

52. Hu, X.; Shelver, W. H. J. Mol. Graph. Model. 22 (2003) 115-126.

53. Hu, X.; Balaz, S.; Shelver, W. H. J. Mol. Graph. Model. 22 (2004) 293-307.

54. Hanessian, S.; Moitessier, N.; Therrien, E. J. Comput. Aided Mol. Des. 15 (2001) 1873-1881.

55. Donini, O. A.; Kollman, P. A. J. Med. Chem. 43 (2000) 4180-4188.

56. Falconi, M.; Altobelli, G.; Iovino, M. C.; Politi, V.; Desideria, A. J. Comput. Aided Mol. Des. 17 (2003) 837-848.

57. Terp, G. E.; Christensen, I. T.; Jorgensen, F. S. J. Biomol. Struct. Dyn. 17 (2000) 933-946.

58. Toba, S.; Damodaran, K. V.; Merz, K. M. Jr. J. Med. Chem. 8 (1999) 1225-1234.

59. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S. J.; Weiner, P. J. Am. Chem. Soc. 106 (1984) 765-784.

60. Ryde, U. J. Comput. Aided Mol. Des. 10 (1996) 10, 153-164.

61. Ryde, U. Biophys J. 77 (1999) 2777-2787.

62. Torrent, M.; Vreven, T.; Musaev, D. G.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B. J. Am. Chem. Soc. 124 (2002) 192-193.

63. Olsen, L.; Antony, J.; Hemmingsen, L. J. Phys. Chem. A 106 (2002) 1046-1053.

64. Rizzo, R.C.; Toba, S.; Kuntz, I. D. J. Med. Chem. 47 (2004) 3065-3074.

65. Merz, K. M. J. Am. Chem. Soc. 113 (1991) 406-411.

66. Borkakoti, N.;Winkler, F. K.; Williams, D. H.; D'Arcy, A.; Broadhurst, M. J.; Brown, P. A.; Johnson, W. H.; Murray, E. J. Nat. Struct. Biol. 1 (1994) 106-110.

67. Lovejoy, B.; Welch, A. R.; Carr, S.; Luong, C.; Broka, C.; Hendricks, R. T.; Campbell, J. A.; Walker, K. A.; Martin, R.; Van Wart, H.; Browner, M. .F. Nat. Struct. Biol. 6 (1999) 217-221.

68. Dhanaraj, V.; Williams, M. G.; Ye, Q. Z.; Molina, F.; Johnson, L. L.; Ortwine, D. F.; Pavlovsky, A.; Rubin, J. R.; Skeean, R. W.; White, A. D.; Humblet, C.; Hupe, D. J.; Blundell, T. L. Croatica Chemica Acta 72 (1999) 575-591.

69. GRID v. 22 Molecular Disovery Ltd. (http://www.moldiscovery.com)

70. Gaussian 98, a.5. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.;

Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian, Inc., Pittsburgh PA, 1998

71. Bauschlicher, C. W. Chem. Phys. Lett. 246 (1995) 40-44.

72. Ricca, A.; Bauschlicher, C. W. Theor. Chim. Acta 92 (1995) 123-131.

73. Becke, D. Physical Rev. A 38 (1988) 3098-3100.

74. Lee, C.; Yang, W.; Parr, R. G. Physical Review B 37 (1988) 785-789.

75. Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Chem. Phys. Lett. 157 (1989) 200-206.

76. Hay, P. J.; Wadt, W. R. J. Chem. Phys. 82 (1988) 299-310.

77. Dunning Jr., T. H.; Hay, P. J. In Modern Theoretical Chemistry; Schaefer, H. F.,III, Ed.; Plenum: New York 1976, Vol. 3, 1-28.

78. Antony, J.; Hansen, B.; Hemmingsen, L.; Bauer, R. J. Phys. Chem. A 104 (2000) 6047-6055.

79. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. J. Phys. Chem. 97 (1993) 10269-10280.

80. Amber version 8.0; University of California at San Francisco: San Francisco, CA.

81. Darden, T.; Darrin, Y.; Pedersen, L. *J. Chem. Phys.* 98 (1993) 10089-10092.

82. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne P. E. Nucl. Acids Res. 28 (2000) 235-242.

83. Grams, F.; Crimmin, M.; Hinnes, L.; Huxley, P.; Pieper, M.; Tschesche, H.; Bode, W. Biochemistry 34 (1995) 14012-14020.

84. Lovejoy, B.; Welch, A. R.; Carr, S.; Luong, C.; Broka, C.; Hendricks, R. T.; Campbell, J. A.; Walker, K. A.; Martin, R.; Van Wart, H.; Browner, M. F. Nat.Struct.Biol. 6 (1999) 217-221.

85. Munie, G.; McDonald, J. J.; Stevens, A. M.; Howard, C. P.; De Crescenzo, G. A.; Welsch, D.; Shieh, H. S.; Stallings, W. C. Biochim.Biophys.Acta 1598 (2002) 10-23.

CHAPTER 3

CANNABINOID RECEPTORS

3.1 CANNABINOID CB2/CB1 SELECTIVITY. RECEPTOR MODELING AND AUTOMATED DOCKING ANALYSIS

Tuccinardi T, Ferrarini PL, Manera C, Ortore G , Saccomanni G, Martinelli A[*] J. Med. Chem. 49 (2006) 984-994

3.1.1 Introduction.

It has been about 41 years since Gaoni and Mechoulam identified Δ^9-tetrahydrocannabinol (THC) as the principal psychoactive molecule present in cannabis.[1,2]

The pharmacological effects of cannabinoids are mediated through at least two receptors, termed CB1 and CB2, although recently, a lot of evidence has been reported about the detection in mouse brain of a third cannabinoid receptor subtype.[3]

As regards their distribution and functionality, CB1 receptors are predominantly located in the central nervous system, and they are probably responsible for most of the overt pharmacological effects of cannabinoids.[4-7] The CB2 receptor is found in peripheral tissues, like spleen, tonsils and immunocytes.[8] This subtype is of particular interest, since it has been identified as a potential target for therapeutic immune treatment, due to its involvement in signal transduction processes in the immune system. Furthermore, a synthetic analogue of THC has recently proved to have significant anti-inflammatory and antitumor effects without any psychoactive effects, and it has been determined that its antineoplastic effect was mediated primarily through actions on CB2 receptors. For all these reasons, at present, research and development of new potent and selective ligands for CB2 is still of great importance.[9]

Both CB1 and CB2 are seven-transmembrane (TM) receptors that belong to the rhodopsin-like family Class A of G protein-coupled receptors (GPCRs), and control a wide variety of multiple intracellular signal transduction pathways. CB1 and CB2 agonists inhibit adenylyl cyclase by activation of a pertussis toxin-sensitive G-protein;[10] moreover CB1 activation inhibits the calcium channels and activates inwardly rectifying potassium channels.[11]

How do GPCRs change their conformations, in response to agonist binding, to activate the associated G-proteins? There are two main hypotheses to account for ligand-mediated G-protein activation, *conformational selection model*[12] and *ligand induction*.[13] According to the *conformational selection model,* there are two conformations of the receptor, the inactive (R) and the active (R*) one; the agonist preferentially binds the receptor in the R* conformation, thus increasing the duration of the period in which the receptor remains in the active state. The *ligand induction model* predicts that transition between the R and R* state is extremely rare in the absence of the agonist, and the free energy of the agonist binding to R is used to facilitate (or induce) the transition to R*.

Bovine rhodopsin, crystallized by Palczewsky et al,[14] provided the first direct visualization of the seven-transmembrane helices of a G-protein-coupled receptor in the inactive state. As regards the activated state, spectroscopic techniques on purified receptor preparations permitted the first direct insight into structural changes that occur during receptor activation, suggesting the conformational differences between the inactive and active states.[13]

A knowledge of the 3D structure of cannabinoid receptors could be of great help in the task of understanding their function and in the rational design of specific ligands. For this purpose, many biochemical, pharmacological and computational studies have been carried out on cannabinoid receptors.

Cannabinoid receptor agonists can be divided into four structurally distinct classes of compounds. These include classic cannabinoids (like Δ^9-THC), non-classical cannabinoids, represented by CP55940, aminoalkylindoles, such as WIN55212-2, and endogenous cannabinoids like arachidonylethanolamide, also called anandamide (AEA).[11]

Aminoalkylindoles derivatives are structurally dissimilar from the other classes, and as site-direct mutagenesis has revealed that the set of aminoacids important for their binding differs significantly from those of the other classes of ligands, the binding site of this type of ligand is probably different from the others.[15]

As regards in the CB1 receptor more specifically, mutation studies have reported that a K3.28(192)A mutation results in a greater loss in affinity for AEA and CP55940, while the WIN55212-2 affinity remains unchanged.[16] In other case, the mutation of

F3.36(191)A, W5.43(279)A and W6.48(356)A in the CB1 receptor determined a loss of affinity only for WIN55212-2, whereas the AEA and CP55940 affinity was unaffected.[17] As regards the CB2 receptor, Song and co-workers reported that the mutation of F5.46(197)V determined a 14-fold decrease in CB2 affinity for WIN55212-2, while the CP55940 and AEA affinity was unchanged.[18]

In the present study, we constructed and refined a three-dimensional model of the CB1 and CB2 receptors in their activated forms. Furthermore, we analysed the docking of WIN55212-2 in both receptor models, with the aim of finding the reasons for its selectivity towards the CB2 receptor.

Finally, the CB2 receptor model thus constructed was used for an automated docking approach on several selective CB2 ligands by means of AUTODOCK 3.0.[19]

3.1.2 Results and Discussion.

3.1.2.1 Molecular modelling. Models of the CB1 and CB2 receptors were generated using the bovine rhodopsin structure determined at 2.8 Å as a template.[14] The length of the transmembrane helices of the two receptors was defined by aligning the rhodopsin primary sequence with both receptor sequences using CLUSTAL W,[20] and verifying the probable presence of the α-helix by means of the Psipred programme (see Figure 1).[21,22]

Following the suggestion of Salo et al,[23] we omitted the highly conserved Proline residue P5.50(215) of rhodopsin, considering the presence of a gap in that position.

It is well-known that rhodopsin was crystallized in its inactive state, and therefore the CB models obtained with the procedure described above represent an inactivated state; according to the activation hypothesis, this state is not suitable for studying the binding of agonists, and the receptor models should be built in their active state.

The activation of the GPC receptors seems to be determined by a different rearrangement of TM3 and TM6, since the disruption of interaction between these two helices produces constitutive receptor activation, and in particular, the extent of constitutive activation is closely correlated with the extent of conformational rearrangement in TM6.[24]

```
                                                        TM1
CB1   VPADQVNITEFYNKSLSSFKENEENIQCGENFMDIECFMVLNPSQQLAIAVLSLTLGTFTVLENLLVLCVILHSR
1F88  -----MNGTEGPNFYVPFSNKTG----VVRSPFEAPQYYLAEPWQFSMLAAYMFLLIMLGFPINFLTLYVTVQHK
CB2   --MEECWVTEIANGSKDGLDSNP----------MKDYMILSGPQKTAVAVLCTLLGLLSALENVAVLYLILSSH

              TM2                                          TM3
CB1   SLRCRPSYHFIGSLAVADLLGSVIFVYSFIDFHVFHRK-DSRNVFLFKLGGVTASFTASVGSLFLTAIDRYISIH
1F88  KLRT-PLNYILLNLAVADLFMVFGGTTTTLYTSLHGYVFGPTGCNLEGFFATLGGEIALWSLVVLAIERYVVVC
CB2   QLRRKPSYLFIGSLAGADFLASVVFACSVFNFHVFHGV-DSKAVFLLKIGSVTMTFTASVGSLLLTAIDRYLCLR

                    TM4                                      TM5
CB1   RPLAYKRIVTRPKAVVAFCLMWTIAIVIAVLPLLGWN---CEKLQSVCSDIF--PH---IDETYLMFWIGVTSVL
1F88  KPMSNFRFG-ENHAIMGVAFTWVMALACAAPPLVGWSRYIPEGMQCSCGIDYYTPHEETNNESFVIYMFVVHFII
CB2   YPPSYKALLTRGRALVTLGIMWVLSALVSYLPLMGWT---CCPRP--CSELF--PL---IPNDYLLSWLLFIAFL

                                                          TM6
CB1   -LLFIVYAYMYILWKAHSHAVRMIQRGTQKSIIIHTSEDGKVQVTRPDQARMDIRLAKTLVLILVVLIICWGPLL
1F88  PLIVIFFCYGQLVFTVKEAAA-----QQQES--------------ATTQKAEKEVTRMVIIMVIAFLICWLPYA
CB2   -FSGIIYTYGHVLWKAHQHVASLSG------------HQDRQVPGMARMRLDVRLAKTLGLVLAVLLICWFPVL

            TM7
CB1   AIMVYDVFGKMNKLIKTVFAFCSMLCLLNSTVNPIIYALRSKDLRHAFRSMFPSCEGTAQPLDNSMGDSDCLHKH
1F88  GVAFYIFTHQGSDFGPIFMTIPAFFAKTSAVYNPVIYIMMNKQFRNCMVTTLCCGKNPLGDDEASTTVSKTETSQ
CB2   ALMAHSLATTLSDQVKKAFAFCSMLCLINSMVNPVIYALRSGEIRSSAHHCLAHWKKCVRGLG-------------

CB1   ANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL
1F88  VAPA---------------------------------
CB2   SEAKEEAPRSSVTETEADGKITPWPDSRDLDLSDC--
```

Figure 1. Alignment of the cannabinoid receptors and bovine rhodopsin (1F88) amino acid sequences. The identical residues are highlighted in black, while the α helix and β sheet prediction carry out by Psipred is respectively marked in grey and underlined

Computational and experimental studies have indicated that conformational switches can be generated in the TM helices as a result of the formation of flexible molecular hinges by a residue of Proline.[25]

During activation in the β_2 adrenergic receptor, P6.50(288) permits the movement of the intracellular end of TM6 away from TM3 and upwards towards the lipid bilayer, suggesting that probably the crucial movements for activation involve flexibility about the hinge formed by the highly conserved Proline in TM6 (P6.50).[26]

Conformational memories calculations of TM6 in the ß2 adrenergic receptor, combined with mutation and SCAM studies, suggested the presence of a rotamer toggle switch able to modulate the TM6 Proline kink; according to this hypothesis C6.47(285)_trans / W6.48(286)_gauche+ / F6.52(290)_gauche+ represent the inactive form of the ß2 adrenergic receptor, while C6.47_gauche+ / W6.48_trans / F6.52_trans represent the active state.[27]

As regards the activation of CB receptors, Singh and co-workers suggested that W6.48(356) / F3.36(200) interaction may act as the toggle switch for CB1 activation, with W6.48_gauche+ / F3.36_trans representing the inactive and W6.48_trans / F3.36_gauche+ the CB1 active form.[28]

Following the indications of these studies, therefore, in order to analyse the agonist binding interaction, we modified the inactive template of our CB1 and CB2 receptors by rotating TM3 and TM6 in a counter-clockwise direction (extracellular point of view) and straightening TM6, using P6.50 as a flexible hinge; finally, we adjusted the conformation of the $\chi 1$ rotamer of W6.48 and F3.36: trans the former and gauche+ the latter.

The receptor models thus obtained were optimized through 400 ps of molecular dynamics (MD), in accordance with the procedure described in the experimental section.

The backbone conformation was evaluated by inspection of the Psi/Phi Ramachandran plot obtained from PROCHECK analysis.[29]

As shown in the Ramachandran plots of Figure 2, the distribution of the Psi/Phi angles of both models are within the allowed regions and no residue has a disallowed conformation.
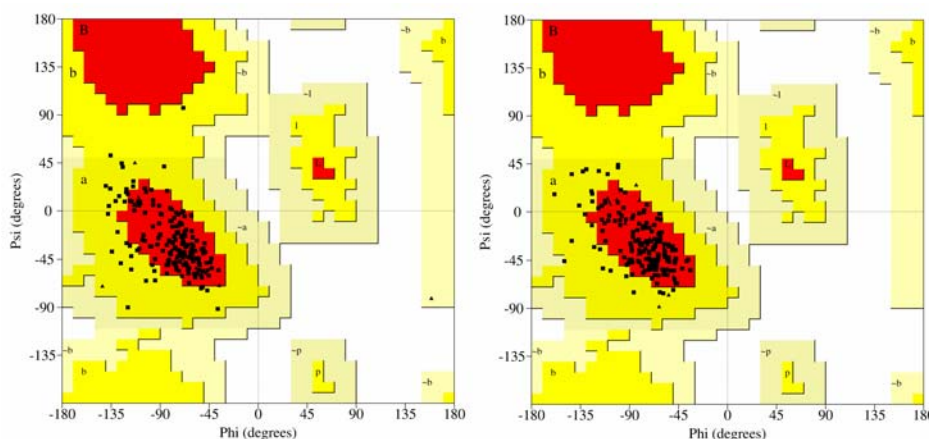
Figure 2. Ramachandran plot of the CB1 (on the left) and CB2 (on the right) receptor. The most favoured regions are coloured red, additional allowed, generously allowed and disallowed regions are indicated as yellow, light yellow and white fields, respectively.

3.1.2.2 Docking of WIN55212-2. The models obtained with these calculations were complexed with a high affinity ligand, and the complexes were optimized.

WIN55212-2 was chosen for this purpose, since it is commonly used in binding experiments and because it shows a high activity in both receptors, in particular CB2 (see Table 3). It was docked bearing in mind the known mutagenesis data.

In the CB1 receptor F3.36(200), W5.43(279) and W6.48(356) might be important for WIN55212-2 binding,[17] whereas CB2 mutagenesis studies suggest the importance of S3.31(112)[30] and F5.46(197)[18] in this subtype. In order to consider these interactions, we inserted the ligand with the morpholinic group positioned between TM3 and TM4, while the naphthyl substituent was directed towards the central core of TM5 and TM6. In this manner, the lipophilic core of the ligand was able to interact with W5.43 and W6.48, and in the CB2 receptor the naphthyl ring could interact with F5.46(197) and the morpholinic group with S3.31(112).

The two complexes were then submitted to 400 ps of MD (see experimental section for details) and Figure 3 shows the docking of the ligand in the two receptors.
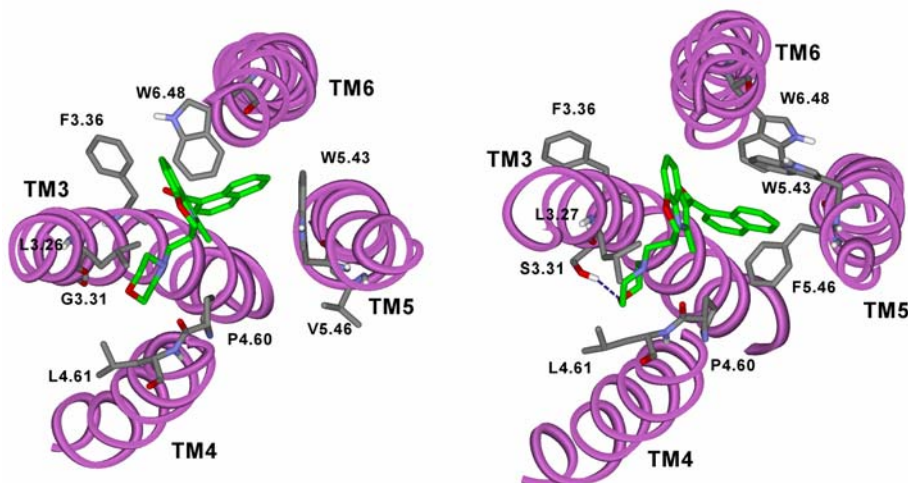
Figure 3. WIN55212-2 docked into the CB1 (left) and CB2 (right) receptors (extracellular point of view).

In the CB1R, the binding site is characterised by a lipophilic pocket delimited by F3.36(200), W5.43(279) and W6.48(356), which principally interact through aromatic stacking with the naphthyl and indole ring system (for distance analysis, see Table 1),

while the morpholinic group is positioned in a secondary lipophilic pocket formed by L3.26(190), P4.60(251) and L4.61(252).

Table 1. Principal WIN55212-2-Receptor interaction in CB1 and CB2. Distances exceeding 7 Å are not reported but indicated as --.

	WIN55212-2					
	Interaction with CB$_1$ (Å)			Interaction with CB$_2$ (Å)		
	Naphthyl	Indole	Morph.	Naphthyl	Indole	Morph.
L3.26	--	5.43	3.52	--	4.21	--
F3.27/L3.27	--	--	6.25	--	3.50	4.06
G3.30/S3.31	--	--	3.81	--	5.68	3.20
F3.36	4.94	5.70	--	--	--	--
P4.60	--	4.01	3.72	--	3.69	3.59
L4.61	--	--	3.90	--	--	3.78
W5.43	4.37	6.70	--	3.54	6.20	--
V5.46/F5.46	--	--	--	4.08	--	--
W6.48	--	5.05	--	5.90	--	--

The CB2 binding site is similar to the CB1 one, with a primary lipophilic pocket delimited by F3.36(117), W5.43(194), W6.48(258), but the WIN55212-2 orientation is slightly different. In the CB2 site, the ligand veers away from F3.36(117), since it feels the effect of a strong interaction with F5.46(197), which is a non-conserved residue (V5.46(282) in the CB1) capable of stabilising the naphthyl ring. As regards the secondary lipophilic pocket in which the morpholinic group is positioned, the substituent interacts with L3.27(108), P4.60(168) and L4.61(169), and the non-conserved S3.31(112) (G3.31(195) in the CB1R) forms a hydrogen bond with the oxygen atom of the morpholinic group. As can be seen in Figure 4, the main interactions of WIN55212-2 with the CB2 receptor were stable: the hydrogen bond with S3.31(112), the aromatic stacking and the π-π interaction of the naphthyl ring with F5.46(197) and W5.43(194), respectively, were maintained during all the last 100 ps of MD.
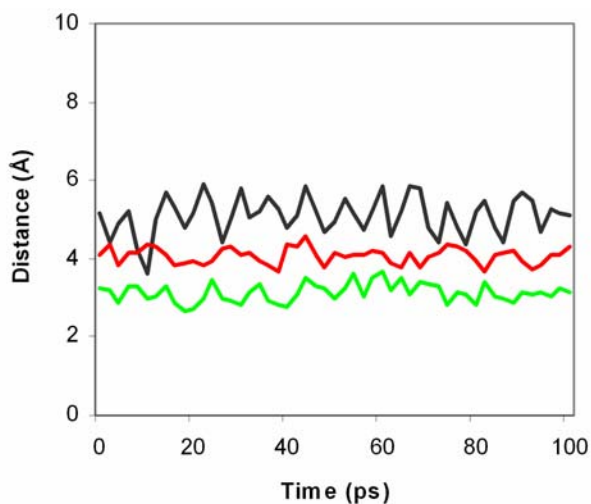
Figure 4. Distances of the main interaction between WIN55212-2 and the CB2 receptor during the last 100 ps of MD simulation. In black, distances between the naphthyl ring of the ligand and the centroid of the aromatic nucleus of F5.46(197), in red, distances between the centroid of the naphthyl ring and the indole of W5.43(194), in green, distances between the O of the morpholinic ring and the OH of the S3.31(112). The distances were updated every 2 ps.

In order to validate our hypothesis and to characterize the structural differences of the two binding sites, we examined the molecular interaction fields (MIFs) obtained by means of the GRID programme[31] for ten different probes (see Table 2).

Table 2. Overview of GRID Probes used in this study

Name	Chemical group
DRY	Hydrophobic probe
C1=	sp^2 CH aromatic or vinyl
C3	Methyl CH_3 group
OC2	Ether oxygen
OH	Phenol or carboxy OH
N:	sp^3 N with lone pair
CL	Organic chlorine atom
O	sp^2 carbonyl oxygen
O1	Alkyl hydroxyl OH group
ArCONHR	Aromatic cis or trans amide

An analysis using C1= and C3 probes showed the presence in both receptors of a large lipophilic pocket corresponding to the space occupied by the naphthyl and indole ring of WIN55212-2, and a secondary one corresponding to the morpholinic position.

Moreover, the OC2 probe showed the presence in the CB2 binding site of a favourable interaction area in the space occupied by the morpholinic group (see Figure 5).

Thus, the observation of the two binding sites and an analysis using different probes encouraged us to carry on a further development of these models.
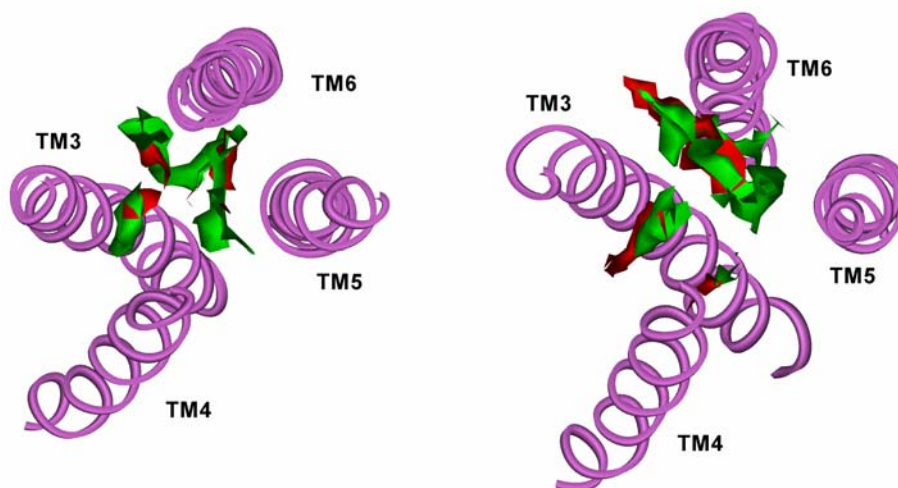
Figure 5. GRID analysis using the C1= (green surface) and OC2 (red surface) probes into the CB1 (left) and CB2 (right) receptors. The C1 probe shows the presence in both receptors of a large lipophilic pocket and a small secondary one, while the OC2 probe shows the presence in the CB2 receptor of a favourable interaction area in the secondary lipophilic pocket

3.1.2.3 Docking of AEA. In order to further test the validity of the models, we docked into both receptors the most well known endogenous ligand AEA, which has a completely different structure with respect to WIN55212-2, through an automated docking procedure (see experimental section for details).

In CB1 AEA adopts an U-shaped molecular conformation and it is placed among TM2-3-6-7 with the aliphatic chain directed towards the intracellular side of the receptor. The amide oxygen atom of the ligand interacts with K3.28(193), in agreement with site directed mutagenesis studies,[16] and the hydroxy group forms an H bond with S7.39(383). The residues that delimited the AEA binding pocket are principally

hydrophobic, including F2.57(170), F3.25(189), L3.29(193), V3.32(196), F3.36(200), F7.35(379), in agreement with the CB1 model proposed by McAllister et al.[17] Anyway, differently from this model, in our study F2.57(170) interacts with the aliphatic chain of AEA through a C-H•••π interaction, whereas F3.25(189) has an interaction with the amide oxygen atom.

In the CB2 receptor AEA does not interact with K3.28(109), but it forms a H-bond with S3.31(112) through the amide oxygen atom, and this is in agreement with mutagenesis studies.[30] Moreover, the hydroxy group interacts with the oxygen backbone of L3.27(108).

The CB2-AEA binding is included among TM3-4-5-6, as for WIN55212-2, and the AEA aliphatic chain interacts principally with W5.43(194) and W6.48(258).

The AEA docking results seem to support the validity of these CB1 and CB2 models since they are in good agreement with the main mutagenesis data available for this ligand.

3.1.2.4 Automated docking. In order to investigate the characteristics of the CB2 model and also its predictivity, we chose from relevant literature 96 ligands (see Table 3) that probably interact in the WIN55212-2 binding site, and using the AUTODOCK 3.0 programme[19] we docked these compounds into the CB2 model. On the basis of their central nucleus and substituents, they can be divided into two classes, indole and naphthyridine derivatives. As concerns the indole derivatives, GTPγS assays indicated that **JWH-151** is a full agonist at CB2, whereas **JWH-120** and **JWH-267** are partial agonists;[32] as for the naphthyridine derivatives, our studies concerning the modulation of mast cell activation[33] highlight that compound **3g** is a full agonist.[34] For these reasons we consider it more reasonable to dock all the ligands tested into the activated form of the CB2 receptor.

The chosen parametrization of AUTODOCK (see experimental section) was tested for its ability to reproduce the binding geometry of WIN55212-2 obtained by means of the molecular dynamics procedure. AUTODOCK easily found the binding geometry corresponding to the one obtained by manual docking, as the rms deviation between the lowest energy docked conformation and the WIN55212-2 manually docked one was 0.62 Å (rms evaluated over all the heavy atoms of the ligand).

Table 3. Structure and binding data of the ligands used in this study. In the last column is indicated which cluster was used in the automated docking. (*ener.*= the cluster with the best average of estimated free energy, *pop.* = the best populated cluster, - = the considered cluster don't belong to none of these two classes). "G sper" and "G calc" indicate, respectively, the experimental and calculated binding free energy (Kcal/mol). "morph" indicated the morpholine ring.

Compd	KiCB₂ (nM)	KiCB₁ (nM)	Ratio CB₂/CB₁	Cluster
WIN-55212-2[30]	1.9±0.1	0.28±0.16	6.8	Ener.

Naphthyridine derivatives

Compd	R	R₁	R₂	n	G sper	G calc	Cluster
2a[41]	H	cyclohexyl	CH₂	0	>-8.18	-8.12	Ener.
2e[41]	H	benzyl	CH₂	0	>-8.18	-7.86	Ener.
3a[41]	CH₂CH₂-N-morph	cyclohexyl	CH₂	0	-9.54	-11.11	Ener.
3b[41]	CH₂CH₂-N-morph	morph	CH₂	0	>-8.18	-10.43	Pop.
3c[41]	CH₂CH₂-N-morph	CH₂cyclohexyl	CH₂	0	-9.45	-11.22	Ener.
3d[41]	CH₂CH₂-N-morph	N-CH₂pipz	CH₂	0	>-8.18	-10.24	-
3e[41]	CH₂CH₂-N-morph	benzyl	CH₂	0	-8.62	-10.63	Ener.
3f[41]	CH₂CH₂-N-morph	4-CH₃-cyclohexyl	CH₂	0	-10.26	-11.43	Ener.
3g[41]	CH₂CH₂-N-morph	cyclopentyl	CH₂	0	-9.95	-10.80	Ener.
3h[41]	CH₂CH₂-N-morph	cycloheptyl	CH₂	0	-10.44	-11.59	Ener.
3i[41]	CH₂CH₂-N-morph	isopentyl	CH₂	0	-9.95	-10.83	Ener.
3l[41]	CH₂CH₂-N-morph	p-Cl-benzyl	CH₂	0	>-8.18	-10.00	-
4a[41]	benzyl	cyclohexyl	CH₂	0	-10.91	-10.57	Ener.
4c[41]	benzyl	benzyl	CH₂	0	>-8.18	-9.68	Ener.
4l[41]	benzyl	p-Cl-benzyl	CH₂	0	>-8.18	-8.43	Ener.
5a[41]	o-F-benzyl	cyclohexyl	CH₂	0	-10.03	-10.26	Ener.
5e[41]	o-F-benzyl	benzyl	CH₂	0	-8.48	-9.91	Ener.
6a[41]	p-F-benzyl	cyclohexyl	CH₂	0	-11.26	-10.65	Ener.
6e[41]	p-F-benzyl	benzyl	CH₂	0	-9.80	-10.05	Ener.
7a[41]	n-hexyl	cyclohexyl	CH₂	0	-11.04	-10.07	Ener.
7e[41]	n-hexyl	benzyl	CH₂	0	-8.85	-9.38	Ener.
8a[41]	n-butyl	cyclohexyl	CH₂	0	-10.58	-9.59	Ener.
8e[41]	n-butyl	benzyl	CH₂	0	>-8.18	-8.56	-
15a[41]	CH₂CH₂-N-morph	benzyl	NH₂	0	>-8.18	-10.93	Ener.
16a[41]	CH₂CH₂-N-morph	cyclohexyl	NH₂	0	>-8.18	-10.81	Ener.
18a[41]	CH₂CH₂-N-morph	cyclohexyl	Cl	0	-10.36	-11.27	Ener.
25a[41]	CH₂CH₂-N-morph	benzyl	CH₂	1	>-8.18	-8.99	-
25b[41]	benzyl	benzyl	CH₂	1	-8.37	-9.28	-
26a[41]	CH₂CH₂-N-morph	cyclohexyl	CH₂	1	>-8.18	-10.38	Ener.
26b[41]	benzyl	cyclohexyl	CH₂	1	-8.56	-9.60	Ener.
26c[41]	n-hexyl	cyclohexyl	CH₂	1	>-8.18	-9.21	-
26d[41]	n-butyl	cyclohexyl	CH₂	1	>-8.18	-8.55	-

Indole derivatives

Compd	R	R₁	R₂	R₃	G sper	G calc	Cluster
JWH-072[31]	C₃H₇	H	H	H	-9.23	-9.70	Pop.
JWH-015[31]	C₃H₇	CH₃	H	H	-10.71	-10.17	Ener.
JWH-018[31]	C₅H₁₁	H	H	H	-11.64	-10.48	Ener.
JWH-007[31]	C₅H₁₁	CH₃	H	H	-11.64	-10.90	Ener.
JWH-120[31]	C₃H₇	H	CH₃	H	-11.17	-10.26	Ener.
JWH-148[31]	C₃H₇	CH₃	CH₃	H	-10.71	-10.64	Ener.
JWH-122[31]	C₅H₁₁	H	CH₃	H	-12.16	-11.06	Ener.
JWH-149[31]	C₅H₁₁	CH₃	CH₃	H	-12.46	-11.36	Ener.
JWH-212[31]	C₅H₁₁	H	C₂H₅	H	-10.91	-10.36	Ener.
JWH-211[31]	C₅H₁₁	CH₃	C₂H₅	H	-10.80	-10.77	Ener.
JWH-210[31]	C₅H₁₁	H	C₂H₅	H	-12.49	-11.24	Ener.
JWH-213[31]	C₅H₁₁	CH₃	C₂H₅	H	-12.78	-11.48	Ener.

Compd	R	R₁	R₂	R₃	G sper	G calc	Cluster
JWH-180[31]	C₅H₁₁	H	C₃H₇	H	-10.93	-10.24	Ener.
JWH-189[31]	C₅H₁₁	CH₃	C₃H₇	H	-10.80	-10.63	Ener.
JWH-182[31]	C₅H₁₁	H	C₃H₇	H	-12.21	-11.13	Ener.
JWH-181[31]	C₅H₁₁	CH₃	C₃H₇	H	-12.55	-11.43	Ener.
JWH-239[31]	C₃H₇	H	C₄H₉	H	-9.93	-10.58	Ener.
JWH-241[31]	C₃H₇	CH₃	C₄H₉	H	-9.97	-10.33	-
JWH-240[31]	C₅H₁₁	H	C₄H₉	H	-11.10	-10.88	Ener.
JWH-242[31]	C₅H₁₁	CH₃	C₄H₉	H	-11.16	-11.20	Ener.
JWH-076[31]	C₃H₇	H	H	CH₃	-9.51	-10.42	Ener.
JWH-046[31]	C₃H₇	CH₃	H	CH₃	-10.63	-10.79	Ener.
JWH-048[31]	C₅H₁₁	CH₃	H	CH₃	-12.69	-11.45	Ener.
JWH-235[31]	C₃H₇	H	H	C₂H₅	-9.42	-10.29	Ener.
JWH-236[31]	C₃H₇	CH₃	H	C₂H₅	-9.03	-9.93	-
JWH-234[31]	C₅H₁₁	H	H	C₂H₅	-11.48	-11.25	Ener.
JWH-262[31]	C₅H₁₁	CH₃	H	C₂H₅	-11.25	-11.22	Ener.

Compd	R	R₁	R₂	R₃	R₄	R₅	G sper	G calc	Cluster
JWH-265[31]	C₃H₇	H	OCH₃	H	H	H	-9.68	-10.11	Ener.
JWH-266[31]	C₃H₇	CH₃	OCH₃	H	H	H	-8.65	-9.89	-
JWH-267[31]	C₅H₁₁	H	OCH₃	H	H	H	-11.10	-10.80	Ener.
JWH-268[31]	C₅H₁₁	CH₃	OCH₃	H	H	H	-10.09	-10.79	Ener.
JWH-079[31]	C₃H₇	H	H	OCH₃	H	H	-10.22	-9.54	-
JWH-094[31]	C₃H₇	CH₃	H	OCH₃	H	H	-9.56	-9.82	Pop.
JWH-081[31]	C₅H₁₁	H	H	OCH₃	H	H	-10.78	-10.84	Ener.
JWH-098[31]	C₅H₁₁	CH₃	H	OCH₃	H	H	-11.89	-10.77	Ener.
JWH-259[31]	C₅H₁₁	H	H	OC₂H₅	H	H	-9.72	-9.87	Ener.
JWH-261[31]	C₃H₇	CH₃	H	OC₂H₅	H	H	-9.07	-9.41	-
JWH-258[31]	C₅H₁₁	H	H	OC₂H₅	H	H	-10.88	-10.32	Ener.
JWH-260[31]	C₃H₇	CH₃	H	OC₂H₅	H	H	-10.36	-11.05	Ener.
JWH-163[31]	C₃H₇	H	H	H	OCH₃	H	-9.35	-9.62	Pop.
JWH-151[31]	C₃H₇	CH₃	H	H	OCH₃	H	-10.26	-9.94	Ener.
JWH-166[31]	C₅H₁₁	H	H	H	OCH₃	H	-11.89	-10.48	Ener.
JWH-153[31]	C₅H₁₁	CH₃	H	H	OCH₃	H	-10.85	-10.69	Ener.
JWH-165[31]	C₃H₇	H	H	H	H	OCH₃	-9.75	-10.03	Ener.
JWH-160[31]	C₃H₇	CH₃	H	H	H	OCH₃	-8.67	-9.78	Ener.
JWH-164[31]	C₅H₁₁	H	H	H	H	OCH₃	-11.13	-10.84	Ener.
JWH-159[31]	C₅H₁₁	CH₃	H	H	H	OCH₃	-10.88	-10.83	Ener.

Compd	R	R₁	R₂	R₃	G sper	G calc	Cluster
1[36]	C(O)-N-morph	OCH₃	4-Cl	C(O)	-8.67	-10.06	Ener.
2[36]	COOMe	OCH₃	2-Cl	C(O)	-8.73	-9.04	Pop.
5[36]	COOMe	OCH₃	4-Cl	C(O)	>-8.18	-9.13	Ener.
6[36]	CH₂-N-morph	OCH₃	4-Cl	C(O)	-9.10	-11.12	Ener.
7[36]	C(O)-N-morph	OCH₃	2-Cl	C(O)	-9.76	-11.46	Ener.
8[36]	C(O)-N-morph	OCH₃	3-Cl	C(O)	-8.80	-10.77	Ener.
9[36]	C(O)-N-morph	OCH₃	2-Cl,3-Cl	C(O)	-10.71	-12.12	Ener.
10[36]	C(O)-N-morph	OCH₃	2-Cl,4-F	C(O)	-9.37	-11.48	Ener.
11[36]	C(O)-N-morph	OCH₃	2-Cl,6-Cl	C(O)	-9.86	-11.71	Ener.
12[36]	C(O)-N-morph	OCH₃	2-Cl,3-Cl	CH₂	>-8.18	-10.80	Pop.
13[36]	CH₂-N-morph	OCH₃	2-Cl,3-Cl	C(O)	-10.80	-11.95	Ener.
14[36]	N-morph	OCH₃	2-Cl,3-Cl	C(O)	-10.44	-11.35	Ener.
15[36]	CH₂-N-morph	H	2-Cl,3-Cl	C(O)	-10.32	-12.04	Ener.

Compd	R	R₁	R₂	G sper	G calc	Cluster
3[36]	CH₂COOMe	1-naphthoyl	OCH₃	-9.34	-10.11	Ener.
16[36]	CH₂CH₂-N-morph	1-naphthoyl	H	-9.09	-10.59	-
17[36]	CH₂-N-morph	1-naphthoyl	H	-11.00	-12.03	Ener.
19[36]	1-naphthoyl	CH₂CH₂-N-morph	H	-10.71	-12.56	Ener.

99

3.1.2.5 Analysis of naphthyridine derivatives. These compounds can form an intramolecular H bond between the carbonylic oxygen atom and the amidic NH (see Table 4). Our studies suggested that this interaction had a high strength, about 30 Kcal/mol (see experimental section for details), and consequently this H bond was considered to be maintained also during interaction in the binding site. For this reason, during the AUTODOCK protocol, we blocked the torsions involved in this intramolecular bond (torsions **d1** and **d2** in Table 4), in order to prevent the loss of this interaction.

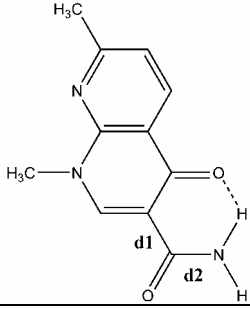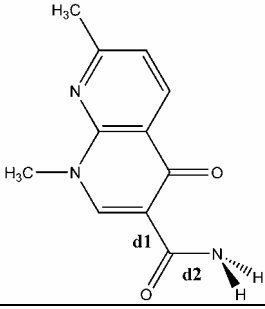**Table 5.** Energy difference between the two optimized models

Model a	Model b

$$\Delta Energy\ (E_{model\ a} - E_{model\ b}) = -31,53\ Kcal/mol$$

Figure 6A shows the docking of **3a** into the CB2 binding site. The ligand position is similar to that of WIN55212-2: **3a** gives H bonds with S3.31(112) through the morpholino substituent, and there is a lipophilic interaction between the cyclohexyl group and W5.43(194) and F5.46(197), in agreement with the mutagenesis data and our binding hypothesis. As regards the naphthyridine ring, it is stabilized by lipophilic interaction with L3.26(107), I3.29(110), M6.55(265) and L6.59(269).

Compound **6a** is 20-fold more potent than **3a** and differs only in the R group, because the ethylmorpholino group is substituted by the *p*-F-benzyl moiety. As shown in Figure 6B, the lowest energy docked conformation of compound **6a** has a binding position very similar to the one shown by **3a** (the rms between the heavy atom positions of the naphthyridine rings of the two ligands in the CB2 binding site is 0.58 Å). As for **3a**, the cyclohexyl group of compound **6a** interacts with W5.43(194) and F5.46(197) while the

R substituent forms an H bond with S3.31(112) (with the fluorine atom). Moreover, the aromatic ring of this substituent is stabilized by the secondary lipophilic pocket formed by L3.27(108), P4.60(168) and L4.61(169), and this could be the reason for the higher affinity shown by this ligand.

The presence of a benzyl substituent in R_1 determines a decrease in the affinity. For example, compound **6e** differs from **6a** only in the presence of the R_1 benzyl instead of the cyclohexyl group, and shows a 10-fold lower affinity for the CB2 receptor.

The superimposition of the binding site of **6e** on that of **6a**, shows that in the docking of **6e,** the presence of an R_1 rigid substituent like benzyl causes a translation of the naphthyridine ring towards the extracellular side of the receptor (the rms between the heavy atoms of **6e** and **6a** naphthyridine rings is 2.12 Å), determining worse interactions with the residues that stabilize the naphthyridine ring (see Figure 6C), with the consequent decrease in affinity.
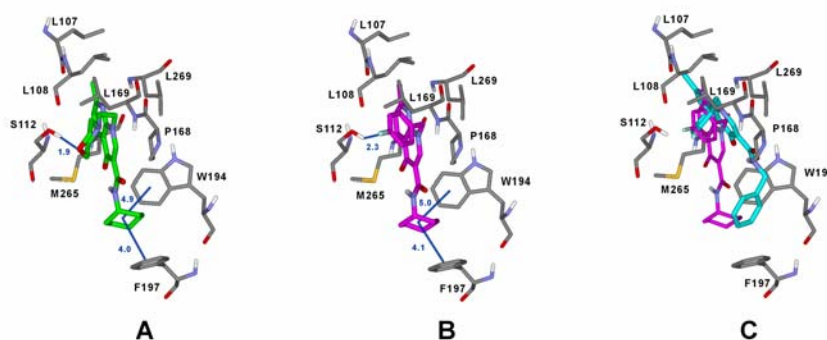

Figure 6. Docking of **3a** (A), **6a** (B), and superimposition (C) of ligands **6a** and **6e** (respectively coloured magenta and sky blue) in the CB2 binding site. The main interatomic distances are reported in blue, all distances are in Angstroms.

3.1.2.6 Analysis of indole derivatives. All the compounds that present a 3-(1-naphthoyl)indole as their central nucleus show a binding position very similar to the one observed for WIN55212-2. The docking of compound **JWH-007**, which is 10-fold less potent than WIN55212-2, shows that the naphthyl ring is stabilized by W5.43(194), F5.46(197) and M6.55(265), while the indole ring interacts with L3.26(107), I3.29(110) and L6.59(269). The R pentyl substituent is inserted into the secondary lipophilic pocket formed by L3.27(108), P4.60(168) and L4.61(169), but of course it is not able to

form the H bond with S3.31(112), and this could be one of the reasons for the lower affinity of this ligand, compared with that of the WIN55212-2 affinity (see Figure 7A).

The compounds (**1-17**) bearing the morpholinic ring linked to the C^3 position of the indole system and the aromatic group linked to the N^1 position (in an opposite manner compared with WIN55212-2, which presents the morpholinic substituent linked to N^1 and the 1-naphthoyl group in the C^3 position), show a different placement of the indole ring in the CB2 binding site, compared with the indole position of the 3-(1-naphthoyl)indole analyzed.

The docking of compound **9**, which presents these structural characteristics, shows that the 3-acetylmorpholine moiety is inserted into the secondary lipophilic pocket and forms the H bond with S3.31(112), while the N^1-2,3-dichlorobenzoyl substituent interacts with W5.43(194), F5.46(197) in the same manner as the naphthyl group of the 3-(1-naphthoyl)indole compounds (see Figure 7B). Comparing the position of **JWH-007** with that of this ligand, we observe that the aromatic and morpholinic groups have the same disposition, but for this reason, unlike from **JWH-007**, the indole ring of compound **9** is upset, with the nitrogen directed towards the intracellular side of the receptor (see Figure 7C).
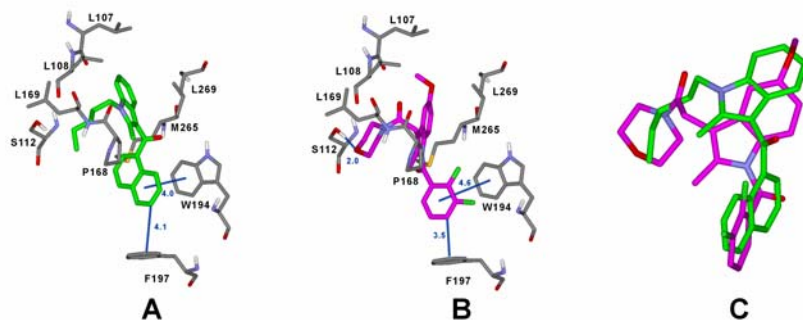

Figure 7. Docking of **JWH-007** (A) and **9** (B) into the CB2 binding site, and superimposition of the two ligands (C) showing the opposite disposition of the indole ring, with the nitrogen pointing towards the intracellular side of the receptor for compound **9** and towards the extracellular side for **JWH-007**.

These observations might suggest that the nitrogen of the indole system should not be important for the interaction, and that the role of the whole indolic system could be only

that of an aromatic core able to place the substituents in the right disposition in the CB2 receptor binding site.

3.1.2.7 Binding free energy estimation. Figure 8 reports the plot of experimental binding energy versus the average estimated binding free energy of the chosen cluster (see experimental section) obtained by using the scoring function of AUTODOCK. The value of the quadratic correlation is very low (R^2=0.31); however we observed that the plot splits the ligands into two different groups. Analyzing the structural characteristics of the compounds belonging to the two groups, we observed that the overestimated one was composed of all the ligands that had the morpholinic substituent.

Considering the plot constituted by two different groups and calculating the two quadratic correlations, we obtained a value of 0.79 for the morpholinic derivatives and a value of 0.70 for all the other ligands. Moreover, the predictive power of the model was tested by leave-one-out (LOO) cross-validation method,[37] where compounds are deleted one after another and prediction of the activity of the deleted compound is based on the QSAR model. This analysis showed a good predictive internal ability for both morpholinic derivatives (q^2=0.74) and the other ligands (q^2=0.69).

As the scoring function of AUTODOCK make use of an empirical approach, and the free energy function is based on the principles of QSAR (quantitative structure-activity relationships) and was parameterized using a large number of protein-inhibitor complexes for which both structure and inhibition constants were known, the split of the ligands into two groups might be due to an overestimation of the AUTODOCK scoring function for the interaction between the morpholinic substituent and our CB2 model.
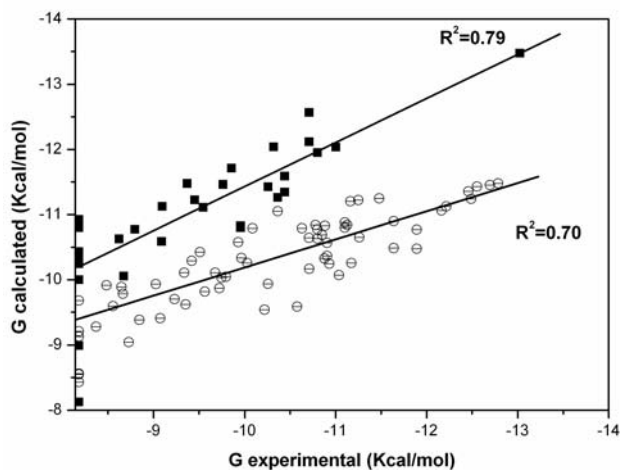
Figure 8. Plot of the experimental binding energy versus the average estimated binding free energy of the chosen cluster. Ligands with a morpholinic group are indicated by ■, while all the others are indicated by ϴ. The quadratic correlations, R^2 were calculated only for the ligands possessing a CB2 affinity greater than 1000 nM.
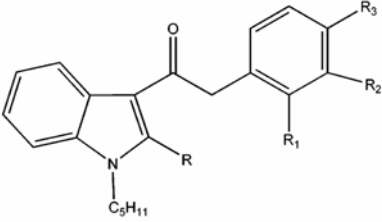
In order to verify if our CB2 model was able to predict the activity of other ligands, we used the 1-penthyl-3-phenylacetylindoles recently published by Huffman et al[38] as test set.

Table 5 shows the 28 compounds tested, their experimental and predicted free energy of binding and the calculated SDEP (Standard Deviation of Errors of Prediction).

The nine ligands with the best CB2 affinity show an *ortho* substituent on the aryl ring (**JWH-252**, **JWH-250**, **JWH-306**, **JWH-311**, **JWH-314**, **JWH-203**, **JWH-204**, **JWH-249** and **JWH-305**). In our model the high affinity of these ligands could be explained by their interaction with T3.37(118), that for most of them consists in the formation of a H bond, since the displacement of the substituent in *meta* or *para* position determined the loss of this interaction.

Table 5. Structure and binding data of the ligands used as test set. In the last column is indicated which cluster was used in the automated docking. (*ener.*= the cluster with the best average of estimated free energy, *pop.* = the best populated cluster, - = the considered cluster don't belong to none of these two classes). "G sper" and "G pred" indicate, respectively, the experimental and predicted binding free energy (Kcal/mol). In the last row is reported the SDEP value.

Compd	R	R_1	R_2	R_3	G exp	G pred	Cluster
JWH-167[38]	H	H	H	H	-9.27	-8.52	Ener.
JWH-205[38]	CH_3	H	H	H	-9.20	-8.45	Ener.
JWH-251[38]	H	CH_3	H	H	-9.32	-8.94	Ener.
JWH-252[38]	CH_3	CH_3	H	H	-9.93	-9.75	Ener.
JWH-208[38]	H	H	H	CH_3	-8.51	-8.15	Ener.
JWH-209[38]	CH_3	H	H	CH_3	-8.00	-8.22	Ener.
JWH-250[38]	H	OCH_3	H	H	-10.20	-9.38	Ener.
JWH-306[38]	CH_3	OCH_3	H	H	-9.66	-10.46	Ener.
JWH-302[38]	H	H	OCH_3	H	-9.61	-8.54	Ener.
JWH-253[38]	CH_3	H	OCH_3	H	-9.65	-9.31	Ener.
JWH-201[38]	H	H	H	OCH_3	-8.66	-7.78	Ener.
JWH-202[38]	CH_3	H	H	OCH_3	-8.44	-8.68	Ener.
JWH-311[38]	H	F	H	H	-10.10	-9.52	Ener.
JWH-314[38]	CH_3	F	H	H	-9.71	-9.10	Ener.
JWH-312[38]	H	H	F	H	-9.60	-8.22	Ener.
JWH-315[38]	CH_3	H	F	H	-9.19	-9.24	Ener.
JWH-313[38]	H	H	H	F	-8.78	-8.15	Ener.
JWH-316[38]	CH_3	H	H	F	-8.33	-8.66	Ener.
JWH-203[38]	H	Cl	H	H	-11.12	-10.03	Ener.
JWH-204[38]	CH_3	Cl	H	H	-10.36	-9.63	Ener.
JWH-237[38]	H	H	Cl	H	-9.51	-9.01	Ener.
JWH-303[38]	CH_3	H	Cl	H	-9.35	-9.93	Ener.
JWH-206[38]	H	H	H	Cl	-8.59	-7.59	Ener.
JWH-207[38]	CH_3	H	H	Cl	-7.40	-7.68	Ener.
JWH-249[38]	H	Br	H	H	-10.50	-10.37	Ener.
JWH-305[38]	CH_3	Br	H	H	-10.28	-10.03	Ener.
JWH-248[38]	H	H	H	Br	-8.43	-7.64	Ener.
JWH-304[38]	CH_3	H	H	Br	-7.60	-7.66	Ener.
SDEP	0.66						

The nine ligands with the best CB2 affinity show an *ortho* substituent on the aryl ring (**JWH-252**, **JWH-250**, **JWH-306**, **JWH-311**, **JWH-314**, **JWH-203**, **JWH-204**, **JWH-249** and **JWH-305**). In our model the high affinity of these ligands could be explained by their interaction with T3.37(118), that for most of them consists in the formation of a H bond, since the displacement of the substituent in *meta* or *para* position determined the loss of this interaction.

3.1.3 Conclusions

We have constructed 3D models of the active conformation of the cannabinoid receptors CB1 and CB2 based on crystallized bovine rhodopsin (1F88).

A model of WIN55212-2 complexed with both receptors is described by means of docking studies, and a comparison of the CB2 and CB1 binding sites showed that the CB2/CB1 selectivity is mainly determined by interaction with the residues S3.31 and F5.46 in the CB2. These residues, corresponding to G3.31 and V5.46 in the CB1, are not conserved, and site-directed mutagenesis suggests that they play an important role.[18,30]

These results suggested that the CB2/CB1 selectivity could be increased by the presence in the ligands of a lipophilic group able to interact in the CB2 with F5.46 and a group able to form a H bond with S3.31.

Using the AUTODOCK programme we docked several ligands into the CB2 model, and their disposition in the receptor confirmed our binding hypothesis. Moreover, the results obtained using this method showed a good correlation between the estimated free energy binding and the experimental binding data.

In order to better verify the predictivity of our CB2 model, an external test set of 28 ligands was used and the SDEP value obtained suggests that this model can be considered quite reliable and predictive.

The cannabinoid receptors are an interesting therapeutic target, in fact many computational studies on these receptors have been recently published.[39-42] Our studies may be very useful in the search for new compounds, and a large database virtual screening analysis using this cannabinoid receptor model is already in progress.

3.1.4 Experimental sections.

3.1.4.1 Amino acid numbering. To refer to specific amino acids sequences, the numbering system suggested by Ballesteros and Weinstein is used.[43]

The most highly conserved residue in each transmembrane helix (TMH) is assigned a value of 0.50 and this number is preceded by the TMH number and followed in parentheses by the sequence number. The other residues in the helix are given a locant value relative to this.

3.1.4.2 Nomenclature of χ1 rotamer. For the χ1 torsion angle we used the nomenclature describe by Shi et al.[44] When the heavy atom in the γ position is in a position opposite the backbone nitrogen, looking from the ß-carbon to the α carbon, the χ1 is said to be trans; from the same viewpoint, when the γ heavy atom is opposite the backbone carbon, the χ1 is said to be gauche+.

3.1.4.3 Methods. The crystal structure of bovine rhodopsin was taken from the Protein Data Bank,[45] while all the primary sequences were obtained from the SWISS-PROT protein sequence database.[46]

As the function of the loops has still not been definite, like other authors[47-49] we modelled only the TM domains of the two receptors. Furthermore, site-directed mutagenesis shows that the perturbation of the first extracellular loop does not affect the binding of WIN55212-2,[50] and in the second extracellular loop, only the mutation of two cysteines determines the loss of binding of the ligand, but the authors hypothesized that the mutation of these conserved residues resulted in an important structural perturbation, perhaps the elimination of a disulfide bridge.[51]

The sequential alignment of rhodopsin and the human cannabinoid receptors CB1 and CB2 was performed by means of CLUSTAL W,[20] using the Blosum series as a matrix, with a gap open penalty of 10 and a gap extension penalty of 0.05. The Psipred programme[21,22] was used in order to verify the presence of α-helices in our TM sequence hypothesis.

Our results were in a perfect agreement with Salo et al.[23] who had taken into consideration several TM receptor sequences for alignment.

All the molecular mechanics and molecular dynamics calculations were performed by means of the Macromodel program[52] by using the AMBER Forcefield. The electrostatic charges were those included in the forcefield and a distance-dependent dielectric constant of 4.0 was used. In molecular mechanics calculations (MM) the minimization algorithms were Steepest Descent followed by Conjugated Gradient until a convergence value of 0.05 kcal/Å•mol; in molecular dynamics simulations the temperature was set at 300 °K and the time step was 1 femtosecond.

All graphic manipulations and visualizations were performed by means of the Maestro[52] and WebLabViewer programmes,[53] while the quantum mechanical calculations were carried out using the Gaussian03 programme.[54]

The backbone conformation of the resulting protein structures was evaluated by inspection of the Psi/Phi Ramachandran plot obtained from PROCHECK analysis,[29] whereas ligand docking was performed using AUTODOCK 3.0.[19]

3.1.4.4 Construction of the activated CB1 and CB2 receptors. The 3D X-ray crystallographic structure of bovine rhodopsin registered in the Protein Data Bank was used as a direct template to construct the 7-TM helical structure of the CB1 and CB2 receptors by means the Maestro programme, on the basis of the alignment obtained from CLUSTAL W and Psipred analysis. Each model helix was capped with an acetyl group at the N-terminus and with an N-methyl group at the C-terminus. Each TM was subjected to preliminary minimization followed by 200 ps of MD, using a constraint of 50 Kcal/mol on the Cα and on the intra-helix H bonds. The final structures were then minimized using the same constraint. The receptor was reassembled on the basis of the rhodopsin structure rotating TM5 by 180° to let W5.43 and V/F5.46 turn towards the intra-helical channel.[16] The whole system was then subjected to preliminary minimization followed by 400 ps of MD, using a constraint of 20 Kcal/mol on the Cα and a constraint with a decreasing force constant (10 to 0.1 Kcal/mol) on the intra-helix H bonds.

The activated states of CB1 and CB2 were created by modification of the rhodopsin-based models thus obtained, by rotating TM3 and TM6 in a counter-clockwise direction (extracellular point of view), straightening TM6 and adjusting the conformation of the χ1 rotamer of W6.48 and F3.36 to trans and gauche+ respectively. The modified TM6 was optimized with the procedure used above for single TMs, and then the whole model was subjected to preliminary minimization, followed by 400 ps of MD, using a constraint of 20 Kcal/mol on the Cα and finally a minimization was applied to the structure obtained as the average of the last 100 ps.

3.1.4.5 Docking of AEA. The ligand was submitted to a conformational search of 1000 steps with an energy window for saving structure of 10 KJ/mol. The algorithm used was the Montecarlo method with MMFFs as the forcefield and a distance-dependent

dielectric constant of 1.0. The ligand was then minimized using the Conjugated Gradient method until a convergence value of 0.05 kcal/Å•mol, using the same forcefield and dielectric constant used for the conformational search. Then the ligand was docked into both receptors using the AUTODOCK 3.0 program.[19] The regions of interest used by AUTODOCK were defined by considering T3.33 as the central residue of a grid of 56, 46, and 50 points in the x, y, and z directions. A grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the energetic map calculations.

Using the Lamarckian Genetic Algorithm, the compound was subjected to 250 runs of the AUTODOCK search, in which the default values of the other parameters were used. Cluster analysis was performed on the docked results using an RMS tolerance of 1.0 Å.

3.1.4.6 Docking of WIN55212-2. The ligand was submitted to a conformational search using the same parameters described above. For WIN55212-2, the best conformation was an s-trans geometry, in agreement with Reggio et al.[55]

The atomic charges of the ligand were calculated by using the RESP method with the 6-31G* wave function on a structure previously minimized at the AM1 level.

WIN55212-2 was docked into both receptors through 400 ps of MD applying a constraint of 20 Kcal/mol on the Cα. Furthermore we applied a constraint on the main ligand-receptor interactions with a decreasing force constant (10, 5, 1 kcal/mol) on the first 300 ps of MD, leaving the ligand free in the last 100 steps. Finally a minimization was applied to the structure obtained as the average of the last 100 ps.

3.1.4.7 Analysis of naphthyridine derivative geometry. In order to measure the strength of the intramolecular H bond, we used a method similar to the one employed by Cuma and co-workers.[56]

We optimized models **a** and **b** of Table 4, using the B3LYP chemical model[57] with a 6-31G+* basis set; as regards the optimization of model **b**, we used a constraint on the torsion involved in the H bond in order to prevent the formation of the H bond during optimization. The energy difference between the two optimized systems was about 30 Kcal/mol, and this value gives us an idea of the high strength of this interaction.

3.1.4.8 Docking of the ligands. Three representative ligands (**3a**, **JWH007**, **9**) were submitted to a conformational search using the same method described for AEA and

the best conformation was used as a general scaffold for the construction of the initial geometry of all the compounds; in all cases, the initial geometries thus obtained were then minimized.

Automated docking was carried out by means of the programme AUTODOCK 3.0;[19] AUTODOCK TOOLS[58] was used to identify the torsion angles in the ligands, add the solvent model and assign partial atomic charges (Gasteiger for the ligands and Kollman for the receptors). The regions of interest used by AUTODOCK were defined by considering WIN55212-2 docked into the CB2 receptor as the central group; in particular, a grid of 54, 50, and 52 points in the x, y, and z directions was constructed centred on the centre of the mass of WIN55212-2. A grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the energetic map calculations.

Using the Lamarckian Genetic Algorithm, all docked compounds were subjected to 100 runs of the AUTODOCK search, in which the default values of the other parameters were used. Cluster analysis was performed on the docked results using an RMS tolerance of 1.0 Å.

As we considered the WIN55212-2 binding geometry to be the one able to stabilize the active form of the CB2 receptor, the selection of the right cluster for each ligand docked was performed mainly on a geometrical basis, i.e. by choosing the best cluster among those in which the ligand had a binding geometry similar to that of WIN55212-2; namely, the one with a substituent inserted between TM3 and TM4 and another substituent directed towards the intracellular side of the receptor.

As regards the most active compounds (those with a Ki lower than 1000 nM), it was found that for 85% of them the chosen cluster was also the one with the best average of estimated free energy, while for 6% of them the chosen cluster was the best populated one.

For ligands **16**, **25b**, **JWH241**, **JWH236**, **JWH079**, **JWH261** and **JWH266** (9%), the chosen cluster did not belong to either of these two types (see Table 3) since neither the cluster with the best average of estimated free energy, or the best populated one possessed a binding geometry similar to that of WIN55212-2.

As regards the ligands that showed a Ki higher than 1000 nM, the chosen cluster for 25% of them was the one possessing the best average of estimated free energy and for 12.5% the best populated one, whereas for 37.5% the chosen cluster did not belong to either of these two types; finally, for four ligands (25%) there were no clusters with a binding geometry similar to that of WIN55212-2, and therefore the cluster with the best average of estimated free energy was chosen.

3.2 DESIGN OF 4-OXO-1,8-NAPHTHYIRIDINE AND QUINOLIN-4(1H)-ON-3-CARBOXAMIDE AS NEW CB2 SELECTIVE AGONISTS

3.2.1 Introduction.

The docking study of naphthyridine derivatives de3scribed in Chapter 3.1 highlighted the ligand-receptor interactions able to determine an increase of the affinity and selectivity. In particular this analysis suggested that for preserving a good CB2/CB1 selectivity and improving the CB2 affinity, it seemed to be necessary: (i) the presence of a non aromatic R2 substituent able to interact into the CB2 with the not conserved residue F5.46(197); (ii) a lipophilic R1 substituent with a H bond acceptor atom able to interact into the CB2 with the not conserved S3.31(112).

For these reasons a new series of 7-methyl-1,8-naphthyridin-4-one-3-carboxamide derivatives were synthesized and tested. For seven compounds of this series the virtual screening into the CB2 model predict an affinity lower than 13nM whereas for one compounds was calculated a poor affinity.

Furthermore our studies suggested that some features of these derivatives seemed to be not important for the interaction with the CB2 receptor; in particular the methyl group (R3) did not seem to strongly interact with any lipophilic residues of the CB2 receptor and the N8 atom of the naphthyridine ring showed a secondary role since it did not interact with any polar residue (see Figure 1). The virtual screening into CB2 receptor of compounds in which the methyl group (R3) was removed or substituted with a chlorine atom and compounds in which the naphthyridin-4-one ring was substituted with a quinolin-4-one as central nucleus, highlighted that these ligands seem to maintain a good CB2 affinity and in some cases the affinity seemed to be greater than their methyl substituted and naphthyridine analogues. In light of these considerations new 1,8-naphthyridin-4(1H)-on-3-carboxamide and quinolin-4(1H)-on-3-carboxamide derivatives in which the R3 methyl group was removed or substituted with a Chlorine atom were synthesized.

All the compounds tested could formed an intramolecular H bond between the carbonylic oxygen and the amidic NH creating a pseudocycle planar to the naphthyridine ring and our studies suggested that this interaction had a high strength. In

order to verify the ability of our CB2 model to discriminate between active and inactive ligands and also to verify if the formation of a planar pseudocycle was important for the interaction inside the CB2 receptor, new compounds characterized by the presence of an hydroxyl group in position 4 of the naphthyridine nucleus instead of the carbonylic oxygen atom and by the absence of the aromaticity of the cycle (which virtual screening predicted a CB2 affinity greater than 750 nM) were synthesized and tested.

3.2.2 Results and Discussion

3.2.2.1 CB1 receptor affinity. The results reported in Table 1 for the 1,8-naphthyridin-4(1H)-on-3-carboxamide derivatives **CB35-CB61** show that, as reported in the previously Chapter, the compounds with an ethylmorpholino group in position 1 (**CB60** and **CB61**) regardless of both the nature of the carboxyamido substituent in position 3 and the substituent in position 7 of the heterocyclic nucleus, exhibited a poor affinity, with Ki values >1000. Analogously the compounds **CB55**, **CB57** and **CB58** bearing in position 1 of the naphthyridine nucleus *N-phenyl-N-ethylpiperazinyl group,* phenethyl group or *p*-methoxybenzyl group respectively possess very low affinity with Ki values >1000. The presence in position 1 of a benzyl group whether substituted or not led to compounds with an interesting affinity; in particular the *p*-fluoro-benzyl derivatives **CB35** and **CB38** showed the highest affinity towards the CB1 receptor, with Ki values of 8.7 nM and 4.3 nM respectively.

As regards the structural modifications at the position 3 of the 1,8-naphthyridine nucleus the substitution of carboxycyclohexylamide group with carboxy-4-methyl-cyclohexylamide or carboxycyclohehepthylamide leads to compounds exhibiting an increase in the affinity towards the CB1 receptor, as is clear from a comparison of compounds **CB35-CB39** with the corresponding 3-carboxycyclohexylamide derivatives previously studied.[35]

Furthermore the substitution of the methyl group in position 7 of the 1,8-naphthyridine nucleus with an atom of chlorine or the lack of any substituent in the same position reduces the CB1 receptor affinity as can be seen from a comparison of compounds **CB41-CB45** and **CB60**, **CB61** with the corresponding 7-methyl-1,8-naphthyridine derivatives previously studied.[35]

Finally the 4-hydroxy-tetrahydro-1,8-naphthyridine derivatives **CB47**, **CB49**, **CB50** and the quinolin-4 (1H)-on-quinoline derivatives **CB64** and **CB65** are lacking of affinity towards CB1 receptor, with Ki value >1000.

The results show that some of the compounds studied possess an interesting affinity at CB1 receptor. In particular the 1,8-naphthyridine derivatives **CB35** and **CB38** exhibit a remerkakable affinity with Ki < 10 nM but are not selective for this receptor.

Table 1. Radioligand binding data of compounds **CB34-CB47**, **CB49**, **CB50**, **CB60**, **CB61**, **CB63**, **CB64** and **CB65**.

CB 34-46,55,57,58,60,61   CB 47,49,50   CB 63-65
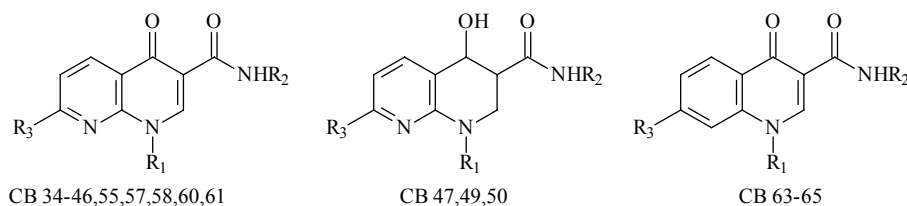
Compd	R	R$_1$	R$_2$	KiCB1[a]	KiCB2[b]	KiCB1/ KiCB2	Pred KiCB2
CB35	p-fluorbenzyl	4-methylcyclohexyl	methyl	8.7 ± 1.6	1.4 ± 0.1	6	3.99
CB36	o-fluorbenzyl	4-methylcyclohexyl	methyl	37.5 ± 5.4	8.4 ± 0.3	4	5.77
CB37	benzyl	cycloheptyl	methyl	143.2 ± 9.1	5.1 ± 1.3	28	7.76
CB38	p-fluorbenzyl	cycloheptyl	methyl	4.3 ± 0.6	1.0 ± 0.1	4.3	2.66
CB39	o-fluorbenzyl	cycloheptyl	methyl	149.4 ± 1.8	13.4 ± 4.7	11	3.57
CB55	N-phenyl-Nethylpiperazine	cyclohexyl	methyl	> 1000	> 1000		2956.88
CB57	phenethyl	cyclohexyl	methyl	> 1000	16.3 ± 1.2	> 62	4.46
CB58	p-methoxyenzyl	cyclohexyl	methyl	> 1000	35.8 ± 2.1	> 28	12.08
CB41	Benzyl	cyclohexyl	cloro	463.6 ± 1.1	24.6 ± 9.7	19	63.61
CB42	p-fluorbenzyl	cyclohexyl	cloro	495 ± 49.4	21.4 ± 1.0	23	9.33
CB43	o-fluorbenzyl	cyclohexyl	cloro	171.2 ± 12.3	18.1 ± 2.7	9.5	7.76
CB44	o-fluorbenzyl	cyclohexyl	H	384.1 ± 45.3	13.0 ± 1.4	29	73.73
CB45	benzyl	cyclohexyl	H	>1000	48.6 ± 18.0	> 21	102.79
CB60	Ethylmorph	4-methylcyclohexyl	Cl	> 1000	40.5 ± 7.7	> 25	125.45
CB61	Ethylmorph	Cyclohexyl	H	> 1000	67.2 ± 11.6	> 15	57.82
CB47	o-fluorbenzyl	cyclohexyl	methyl	>1000	>1000		1013.71
CB49	Ethylmorph	4-methylcyclohexyl	methyl	>1000	>1000		777.38
CB50	benzyl	Cyclohexyl	methyl	>1000	>1000		1763.56
CB64	benzyl	Cyclohexyl	H	>1000	4.8 ± 0.4	> 210	39.37
CB65	Ethylmorph	Cyclohexyl	Cl	>1000	3.3 ± 0.4	> 303	17.87
SDEP							0.69

[a] Affinity of compounds for CB1 receptor was evaluated using mouse brain membranes and [^{3}H]CP55,940.
[b]Affinity of compounds for CB2 receptor was evaluated using mouse spleen and [^{3}H]CP55,940.

3.2.2.2 CB2 receptor affinity. The results obtained indicate that, in according with that previously reported,[35] the N-benzyl-1,8-naphthyridine derivatives possess higher affinity than the N-ethylmorpholino derivatives as is clear from a comparison of compounds **CB35**, **CB36 CB44** and **CB45** with **CB60** and **CB61**. For the N-benzyl-1,8-

naphthyridine derivatives (**CB35-CB39**, **CB41-CB45**, **CB58**), the presence of an atom of fluorine on the benzyl increase the affinity, above all of the substitution is the *para* position. In particular the *p*-fluorobenzyl-1,8-naphthyridine derivatives **CB35** (Ki = 1.4 nM) and **CB38** (Ki = 1 nM) proved to be the compounds with the highest affinity in this series. Furthermore the N-phenethyl-1,8-naphthyridine derivative **CB57** showed a good affinity with a Ki value of 16.3 nM. On the contrary the compound bearing in position 1 of the naphthyridine nucleus a *N-phenyl-N-ethylpiperazinyl group* (**CB55**), possesses very low affinity with Ki value >1000.

As had previously be found for the CB1 receptor, the substitution of carboxycyclohexylamide (cyclohexylamide) group at the position 3 of the 1,8-naphthyridine nucleus with carboxy-4-methyl-cyclohexylamide or carboxycyclohehepthylamide determines an increase in the affinity towards the CB2 receptor, as confirmed by a comparison compounds **CB35**-**CB39** with the corresponding 3-carboxycyclohexylamide derivatives previously studied.[35]

Furthermore the substitution of the methyl group in position 7 of the 1,8-naphthyridine nucleus with an atom of chlorine or the lack of any substituent in the same position determine generally the maintenance or an increase of the affinity (except for compounds **CB42** and **CB45** that showed a fourfold and fivefold decrease of affinity respect to the methyl-analogues).

Analogously to results for the CB1 receptor, the 4-hydroxy-1,8-naphthyridine derivatives **CB47**, **CB49**, and **CB50** exhibit a poor affinity towards CB2 receptor, with Ki value >1000.

Finally compounds **CB64** and **CB65** in which, as suggested by virtual screening, the 4-one- naphthyridine system was substituted by the 4-one-quinoline, possess a remarkable affinity with Ki value of 4.8 and 3.3 nM respectively. The most compounds showed a good affinity for the CB2 receptor. In particular the 1,8-naphthyridin-4-one derivatives **CB35-CB39** and the 4-one-quinoline derivatives **CB64** and **CB65** possess a remarkable affinity with Ki <10 nM. As regards the selectivity towards CB2 receptor the 1,8-naphthyridine **CB44**, **CB45**, **CB57**, **CB58**, and **CB60**, show a good selectivity with KiCB1/KiCB2 >20. Furthermore 4-one-quinoline derivatives **CB64** and **CB65** exhibit a very important CB2 receptor selectivity with KiCB1/KiCB2 >208 and >303

respectively which result higher than that reported for the analogous compounds previously published (KiCB1/KiCB2=143 for the best compound).[59]

3.2.2.3 Molecular modeling. Using the AUTODOCK 3.0 program[19] the compounds showed in Table 1 were docked into the CB2 receptor, their activity were predicted basing on the published computational model described in Chapter 3.1 (see Experimental section for details) and as indicated by the SDEP value (0.69) reported in Table 1 there was a quite good correlation between the experimental and the calculated Ki. Furthermore all the ligands with a CB2 affinity greater than 1000 nM were predicted with an affinity greater than 750 nM. As suggested by our model the compound with the best CB2 affinity was **CB38**. The docking showed that the CB2 binding pocket was delimited by TM3, TM4, TM5 and TM6, and the cycloheptyl substituent of compound **CB38** was directed towards the intracellular side of the receptor interacting with W5.43(194) and F5.46(197) (see Figure 1). As concern the *p*-F-benzyl group it interacted in a lipophilic task constituted by L3.27(108), P4.60(168) and L4.61(169) and the Fluorine atom formed a H bond with S3.31(112).
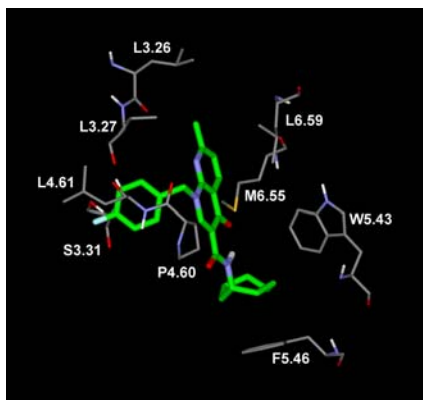

Figure 1. Compound **CB38** docked into CB2 receptor model.

The docking of compound **CB65** (see Figure 2), the most CB2/CB1 selective ligand of this series, revealed also that the presence of the Chlorine atom could contribute to the increase of the CB2 affinity and CB2/CB1 selectivity, since it could interact with the non conserved S6.58(268) (Aspartate in the CB1).
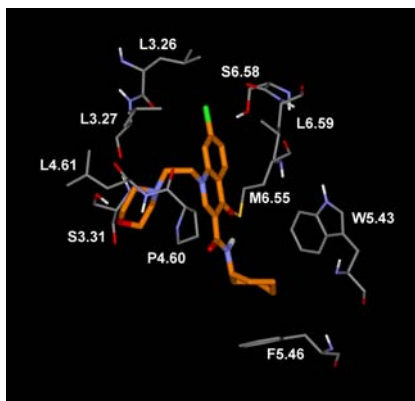
Figure 2. Compound **CB65** docked into CB2 receptor model.

Finally the docking of the inactive compounds **CB47**, **CB49** and **CB50** revealed that the lack of the planarity determined a different position of the central lipophilic core, determining worse interactions with the receptor. As shown in Figure 3 the central core of **CB50**, comparing to the position of the naphthyridine ring of **4a** was upset, far way from TM3 and directed towards TM5 determining worse interaction with M6.55(265) and L3.27(107); furthermore this disposition determined a worse interaction of the cyclohexyl ring with F5.46(197), being at a distance of 4.7 Å (while for **4a** was 4.0 Å).
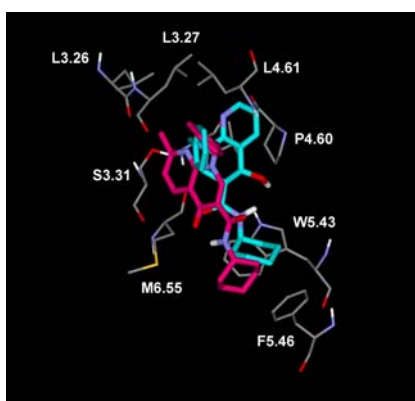

Figure 3. Superimposition of compounds **4a** and **CB50** docked into CB2 receptor.

3.2.3 Conclusions.

In the present study, through a structure based approach, we tried to improve the activity and selectivity of 1,8-naphthyridin-4(1H)-on-3-carboxamide derivatives which had shown to be a new class of CB2 ligands.

For this purpose, following the suggestion obtained from the docking of ligands into a CB2 receptor model, new 4-oxo-1,8-naphthyiridine and quinolin-4(1H)-on-3-carboxamide were designed, synthesized, and tested on the CB1 and CB2 receptors.

Some of these compounds showed a good CB2/CB1 selectivity and a high CB2 affinity in an agreement with the values predicted by the docking study. In particular, **CB38**, which presented the *p*-F-benzyl and cycloheptyl substituents bounded to the 4-oxo-1,8-naphthyridine nucleus showed a CB2 affinity value of 1 nM.

The substitution of the 4-one-naphthyridine central nucleus with the 4-one-quinoline determined a general increase of the CB2 affinity. For compound **CB65** the good CB2 affinity was also accompanied by an high CB2/CB1 selectivity and the docking studies suggested that the interaction of the chlorine substituent with the not conserved residue S268 in the CB2 receptor seemed to be one of the reasons able to explain the high selectivity value.

Finally, the low affinity showed by the new 4-hydroxy-tetrahydro-1,8-naphthyridine derivatives confirmed the hypothesis about the fundamental role of the presence of a planar pseudocycle.

These results provide interesting additions to the currently available SAR for the cannabinoid agonist ligands, opening up a new research field for designing new cannabinoids receptor agonists, characterized by a high CB2/CB1 selectivity.

3.2.4. Experimental section.

Ligands were submitted to a conformational search of 1000 steps with an energy window for saving structure of 10 KJ/mol. The algorithm used was the Montecarlo method with MMFFs as the forcefield and a distance-dependent dielectric constant of 1.0. The ligands were then minimized using the Conjugated Gradient method until a convergence value of 0.05 kcal/Å•mol, using the same forcefield and dielectric constant used for the conformational search. Automated docking was carried out by means of the programme AUTODOCK 3.0;[19] AUTODOCK TOOLS[58] was used to identify the torsion angles in the ligands, add the solvent model and assign partial atomic charges (Gasteiger for the ligands and Kollman for the receptors). The regions of interest used by AUTODOCK were defined by considering the previously published WIN55212-2 docked into the CB2 receptor as the central group of a grid of 54, 50, and 52 points in

the *x*, *y*, and *z* directions. A grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the energetic map calculations. As all the compounds can form an intramolecular H bond and our previously study suggested an high strength for this interaction (see Chapter 3.1), this H bond was considered to be maintained also during interaction in the binding site. For this reason, during the AUTODOCK protocol, we blocked the torsions involved in this intramolecular bond in order to prevent the loss of this interaction. Using the Lamarckian Genetic Algorithm, all docked compounds were subjected to 100 runs of the AUTODOCK search, in which the default values of the other parameters were used. Cluster analysis was performed on the docked results using an RMS tolerance of 1.0 Å, and the cluster with the best average of estimated free energy was chosen. In order to predict the binding affinity of the ligands, a correlation model very similar to that obtained in our previous chapter was used: the only difference was that also the test set previously used for validate the predictivity of the model was now incorporated in the training set. Figure 4 shows the plot of experimental binding energy versus the average estimated binding free energy and the equation used for the prediction calculation.
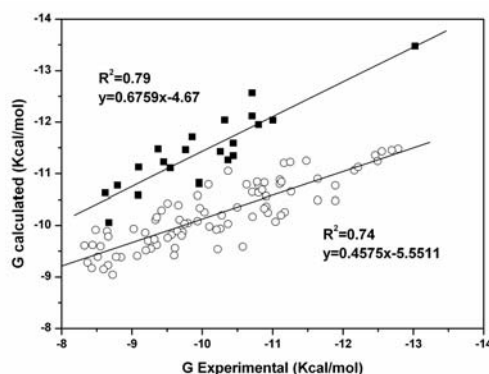

Figure 4. Plot of the experimental binding energy versus the average estimated binding free energy of the chosen cluster. Ligands with a morpholinic group are indicated by ■, while all the others are indicated by ○.

3.3 REFERENCES.

1. Gaoni, Y.; Mechoulam R. J. Am. Chem. Soc. 86 (1964) 1646-16477.

2. Mechoulam, R.; Ben-Shabat, S. Nat. Prod. Rev. 16 (1999) 131-143.

3. Breivogel, C. S.; Griffin, G.; Di Marzo, V.; Martin, B. R. Mol. Pharmacol. 60 (2001) 155-163.

4. Matsuda, L. A.; Lolait, S. J.; Brownstein, M. J.; Young, A. C.; Bonner, T. I. Nature 346 (1990) 561-564.

5. Huffman, J. W.; Lainton, J. A. H. Recent developments in the medicinal chemistry of cannabinoids. Curr. Med. Chem. 3 (1996) 101-116.

6. Herkenham, M.; Lynn, A. B.; Little, M. D.; Johnson, M. R.; Melvin, L. S.; De Costa, D. R.; Rice, K. C. 87 (1990) 1932-1936.

7. Pertwee, R. G. Pharmacology of cannabinoid receptor ligands. Curr. Med. Chem. 6 (1999) 635-664.

8. Galiegue, S.; Mary, S.; Marchand, J.; Dussossoy, D.; Carriere, D.; Carayon, P.; Bouaboula, M.; Shire, D.; Le Fur, G.; Casellas, P. Eur. J. Biochem. 232 (1995) 54-61.

9. Recht, L. D.; Salmonsen, R.; Rosetti, R.; Jang, T.; Pipia, G.; Kubiatowski, T.; Karim, P.; Ross, A. H.; Zurier, R.; Litofsky, N. S.; Burstein S. Biochem. Pharmacol. 62 (2001) 755-763.

10. Felder, C. C.; Joyce, K. E.; Briley, E. M.; Mansouri, J.; Mackie, K.; Blond, O.; Lai, Y.; Ma, A. L.; Mitchell R. L. Mol. Pharmacol. 48 (1995) 443-450.

11. Reggio, P. H. Curr. Pharm. Des. 9 (2003) 1607-1633.

12. Colquhoun D. Br. J. Pharmacol. 125 (1998) 924-947.

13. Gether U. Endocr. Rev. 21 (2000) 90-113.

14. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Science 289 (2000) 739-745.

15. Reggio, P. H. Curr. Med. Chem. 6 (1999) 665-683.

16. Song, Z. H.; Bonner, T. I. Mol. Pharmacol. 49 (1996) 891-896.

17. McAllister, S. D.; Rizvi, G.; Anavi-Goffer, S.; Hurst, D. P.; Barnett-Norris, J.; Lynch D. L.; Reggio, P. H.; Abood, M. E. J. Med. Chem. 46 (2003) 5139-5152.

18. Song, Z. H.; Slowey, C. A.; Hurst, D. P; Reggio, P. H. Mol. Pharmacol. 56 (1999) 834-840.

19. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. J. Comp. Chem. 19 (1998) 1639-1662.

20. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. Nucleic Acids Res. 22 (1994) 4673-4680.

21. McGuffin, L. J.; Bryson, K.; Jones, D. T. Bioinformatics 16 (2000) 404-405.

22. Jones, D. T. J. Mol. Biol. 292 (1999) 195-202.

23. Salo, O. M.; Lahtela-Kakkonen, M.; Gynther, J.; Jarvinen, T.; Poso, J. Med. Chem. 47 (2004) 3048-3057.

24. Ballesteros, J. A.; Jensen, A. D.; Liapakis, G.; Rasmussen, S. G.; Shi, L.; Gether, U.; Javitch, J. A. J. Biol. Chem. 276 (2001) 29171-29177.

25. Sansom, M. S.; Weinstein, H. Trends Pharmacol. Sci. 21 (2000) 445-451.

26. Jensen, A. D.; Guarnieri, F.; Rasmussen, S. G.; Asmar, F.; Ballesteros, J. A.; Gether U. J. Biol. Chem. 276 (2001) 9279-9290.

27. Shi, L.; Liapakis, G.; Xu, R.; Guarnieri, F.; Ballesteros, J. A.; Javitch, J. A. J. Biol. Chem. 277 (2002) 40989-40996.

28. Singh, R.; Hurst, D. P.; Barnett-Norris, J.; Lynch, D. L.; Reggio, P. H.; Guarnieri F. J. Peptide. Res. 60 (2002) 357-370.

29. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thorton, J. M. J. Appl. Crystallogr. 26 (1993) 283–291.

30. Tao, Q.; McAllister, S. D.; Andreassi, J.; Nowell, K. W.; Cabral, G. A.; Hurst, D. P.; Bachtel, K.; Ekman, M. C.; Reggio, P. H.; Abood, M. E. Mol. Pharmacol. 55 (1999) 605-613.

31. GRID v. 22 Molecular Disovery Ltd. (http://www.moldiscovery.com).

32. Huffman, J. W.; Zengin, G.; Wu, M. J.; Lu, J.; Hynd, G.; Bushell, K.; Thompson, A. L.; Bushell, S.; Tartal, C.; Hurst, D. P.; Reggio, P. H.; Selley, D.

E.; Cassidy, M. P.; Wiley, J. L.; Martin, B. R. Bioorg. Med. Chem. 13 (2005) 89-112.

33. Vannacci, A.; Giannini, L.; Passani, M. B.; Di Felice, A.; Pierpaoli, S.; Zagli, G.; Fantappiè, O.; Mazzanti, R.; Masini, E.; Mannaioni, P. F. J. Pharmacol. Exp. Ther. 311 (2004) 256-264.

34. Unpublished results from our laboratory.

35. Ferrarini, P. L.; Calderone, V.; Cavallini, T.; Manera, C.; Saccomanni, G.; Pani, L.; Ruiu, S.; Gessa, G. L. Bioorg. Med. Chem. 12 (2004) 1921-1933.

36. Gallant, M.; Dufresne, C.; Gareau, Y.; Guay, D.; Leblanc, Y.; Prasit, P.; Rochette, C.; Sawyer, N.; Slipetz, D. M.; Tremblay, N.; Metters, K. M.; Labelle, M. Bioorg. Med. Chem. Letters 6 (1996) 2263-2268.

37. Chatterjee, S.; Hadi, A. S.; Price, B. Regression Analysis by Examples 2000, John Wiley & Sons, Inc.

38. Huffman, J. W.; Szklennik, P. V.; Almond, A.; Bushell, K.; Selley, D. E.; He, H.; Cassidy, M. P.; Wiley, J. L.; Martin, B. R. Bioorg. Med. Chem. Letters 15 (2005) 4110-4113.

39. Montero, C.; Campillo, N. E.; Goya, P.; Páez, J. A. Eur. J. Med. Chem. 40 (2005) 75-83.

40. Shim, J. Y.; Welsh, W. J.; Howlett, A. C. Biopolymers 71 (2003) 169-189.

41. Xie, X. Q.; Chen, J. Z.; Billings, E. M. Proteins 53 (2003) 307-319.

42. Salo, O. M. H.; Raitio, K. H.; Savinainen, J. R.; Nevalainen, T.; Lahtela-Kakkonen, M.; Laitinen, J. T.; Järvinen, T.; Poso, A. J. Med. Chem. In press.

43. Ballesteros, J. A.; Weinstein, H. W. Methods in Neuroscience 25 (1995) 366-428.

44. Shi, L.; Liapakis, G.; Xu, R.; Guarnieri, F.; Ballesteros, J. A.; Javitch, J. A. J. Biol. Chem. 277 (2002) 40989–40996.

45. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne P. E. Nucl. Acids Res. 28 (2000) 235-242.

46. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. Nucleic Acids Res. 31 (2003) 3784-3788.

47. Bramblett, R. D.; Panu, A. M.; Ballesteros, J. A.; Reggio, P. H. Life Sci. 56 (1995) 1971-1982.

48. McAllister, S. D.; Tao, Q.; Barnett-Norris, J.; Buehner, K.; Hurst, D. P.; Guarnieri, F.; Reggio, P. H.; Nowell Harmon, K. W.; Cabral, G. A.; Abood, M. E. Biochem. Pharmacol. 63 (2002) 2121-2136.

49. Barnett-Norris, J.; Hurst, D. P.; Lynch, D. L.; Guarnieri, F.; Makriyannis, A.; Reggio, P. H. J. Med. Chem. 45 (2002) 3649-3659.

50. Murphy, J. W.; Kendall, D. A. Biochem. Pharmacol. 65 (2003) 1623-1631.

51. Gouldson, P.; Calandra, B.; Legoux, P.; Kernéis, A.; Rinaldi-Carmona, M.; Barth, F.; Le Fur, G.; Ferrara P.; Shire D. Eur. J. of Pharm. 401 (2000) 17-25.

52. Macromodel ver. 8.5, Schrodinger Inc., 1999.

53. WebLab Viewer Pro 3.7 (Accelrys Inc., San Diego, CA)

54. Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, 2004.

55. Reggio, P. H.; Basu-Dutt, S.; Barnett-Norris, J.; Castro, M. T.; Hurst, D. P.; Seltzman, H. H.; Roche, M. J.; Gilliam, A. F.; Thomas, B. F.; Stevenson, L. A.; Pertwee, R. G.; Abood, M. E. J. Med. Chem. 41 (1998) 5177-87.

56. Cuma, M.; Scheiner, S.; Kar, T. Theochem 467 (1999) 37-49.

57. Bauschlicher, C. W. Chem. Phys. Lett. 246 (1995) 40-44.

58. http://www.scripps.edu/~sanner/python/adt

59. Stern, E.; Muccioli, G.G.; Millet, R.; Goossens, J. F.; Farce, A.; Chavatte, P.; Poupaert, J. H.; Lambert, D. M.; Depreux, P.; Henichart, J. P. J Med Chem. 49 (2006) 70-79.

CHAPTER 4

ESTROGEN RECEPTORS

4.1 PREDICTIONS OF BINDING OF ESTROGEN RECEPTOR LIGANDS BY A COMBINATION OF MOLECULAR DYNAMICS AND CONTINUUM SOLVENT MODELS

4.1.1 Introduction.

The estrogen receptor (ER) is a member of the nuclear receptor superfamily that functions as a ligand-regulated transcription factor.[1] Many physiological processes can be regulated by selectively activating or inhibiting the ER with appropriate agonist or antagonist ligands. Positive effects on the maintenance of bone mineral density,[2] on blood lipid levels, and on vasomotor and central nervous system functions[3] are generally considered to justify the use of estrogen agonists as agents in the treatment of postmenopausal osteoporosis,[4] atherosclerosis,[5] hot flush responses, and Alzheimer's disease.[6] Unfortunately, the activation of the ER also results in an increase in breast and uterine cancer.[7-9] Therefore, molecules, such as tamoxifen and raloxifene, that block the tumor-promoting effects of endogenous estrogens are currently used in the therapy and prevention of breast cancer.[10-12] These drugs, however, are not pure estrogen antagonists, because they promote estrogen-like effects in certain tissues; rather, they are termed selective estrogen receptor modulators (SERMs).[13,14] SERMs are particularly attractive as therapeutic agents because they are able to block estrogen action at those sites where stimulation would be undesirable, such as the breast and uterus, but at the same time stimulate estrogen actions in other tissues where they are desired, such as the bone and liver.[11,15] A great deal of effort has been devoted to the task of understanding the processes by which SERMs are able to exert tissue-specific estrogen agonist and antagonist effects and to improve upon their already rather favorable profile of selectivity. More recently, a second subtype of the estrogen receptor, named ER, was discovered.[16,17] The two ERs share about 95% homology in their DNA binding domains; as regards the ligand binding domains (LBD), they share only a modest homology (58% identity), but their ligand binding cavities are nearly identical, differing by only two amino acid residues (ERα L384 and M421 are respectively replaced by ERβ M336 and I373).[18]

The tissue distribution patterns of ERα and ERβ however, are rather different,[15,19] as are their biological functions, some of which have been revealed by in vivo studies on receptor subtype-specific knockout mice.[20-22] The tissue-selective pharmacology of SERMs is thought to result from their different action on each ER subtype and/or by the different interactions that the ER-ligand complex might have with the cellular coregulatory proteins or effector components that vary from tissue to tissue.[23,24]

A new series of anthranylaldoximes derivatives (see Table 1) bearing various substituents were synthesized and tested and the results indicate that this new class of compounds showed interesting ER binding properties on both receptors subtypes (ERα and ERβ).

A docking study of these anthranylaldoximes compounds with both the ERs was carried out and the results were further investigated through the recently developed MM/PBSA (molecular mechanics-Poisson Boltzmann surface area) approach.

Table 1. Ligand used for the docking study (Kd is espressed as nM)

Compd	R_1	R_2	R_3	R_4	R_5	Kd ERα	Kd ERβ	α/β ratio
L4	H	H	H	H	CH_3	4	9	2.2
L5	OH	H	H	H	CH_3	3	42	14.0
L6	CH_3	H	CH_3	H	CH_3	17	25	1.5
L7	H	CH_3	H	CH_3	CH_3	4	22	5.5
L11	H	H	OH	H	CH_3	13	64	4.9
L12	OH	H	OH	H	CH_3	8	323	40.4
L13	OCH_3	H	OCH_3	H	CH_3	7	97	13.9
L14	H	OH	H	OH	CH_3	112	1004	9.0
L15	H	OCH_3	H	OCH_3	CH_3	24	325	13.5
L16	H	H	H	H	CH_2CH3	33	963	29.2
L17	H	H	H	H	H	7	17	2.43

4.1.2 Results and Discussion.

X-ray structures of ERα (1G50[25]) and ERβ (1X7J[26]) were complexed with estradiol and the complexes were subjected to 900 ps of molecular dynamics (MD) simulation (see Experimental section for details). In ERβ the missed loop M410-N421 were constructed by means of Modeller program[27] using 1QKN[18] as template. As shown in Figure 1A, after about 100 ps of MD, both systems reached an equilibrium, since the total energy for the last 800 ps remained constant. Analysing the RMS deviation (RMSD) from the X-ray structures of all the heavy atoms of the proteins, we observed that after an initial increase, in the last 800 ps the RMSD remained approximately constant around the range of 0.9-1.0 Å for ERα and 0.7-0.8 Å for ERβ (see Figure 1B).
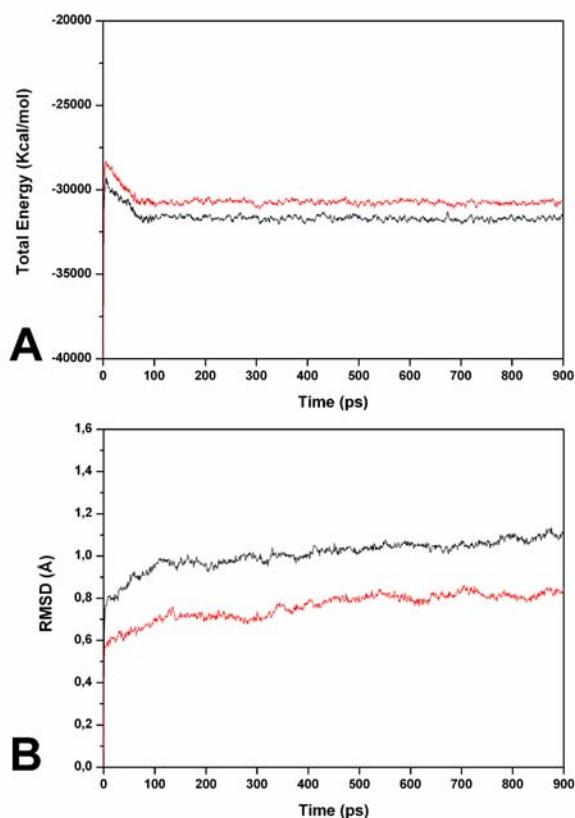

Figure 1. Analysis of the MD simulation of estradiol complexed with ERα (black) and ERβ (red). In the first plot the Total energy of the system vs the time was reported; in the second plot the RMSD between all the heavy atoms of the receptors and the two X-ray starting structures was reported.

129

Figure 2 shows the minimized average structures of the binding site of the two receptors complexed with estradiol. As also confirmed by the H bond analysis of the MD simulations (see Table 2), estradiol presented a very similar interaction scheme into both receptors: by means of 17'-hydroxy group it interacted with H524(475), while through the 3'-hydroxy substituent it formed a H bond with E353(305). Furthermore, as already suggested by Brzozowski and Pike[18,28] a water molecule seems to be the central core of a H bond network system between E353(305), R394(346) and estradiol (that interacted with water molecule through the 3'-hydroxy substituent).
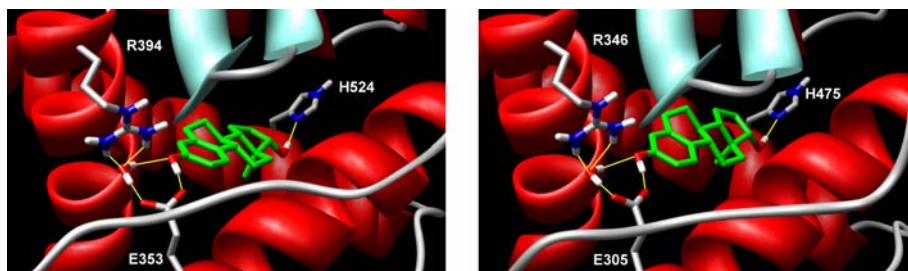
Figure 2. Hydrogen-bonding network of the estradiol in the ERα (left) and ERβ (right).

Table 2. Hydrogen bonds analysis of the E353(305)-R394(346)-Water-Estradiol-H524(475) system during the MD simulation into both receptors.

Donor	AcceptorH	Acceptor	Distance (Å)	% Occupied
ERα				
E353@OE2	EST@HO3	EST@O3	2.55	100
H524@ND1	EST@OH1	EST@O17	3.03	97
WAT@O	R394@HH12	R394@NH1	2.98	94.95
E353@OE1	WAT@H1	WAT@O	2.70	91.20
EST@O3	WAT@H2	WAT@O	2.95	65.75
WAT@O	R394@HH22	R394@NH2	3.24	63.10
ERβ				
E305@OE2	EST@HO3	EST@O3	2.60	100
H475@ND1	EST@OH1	EST@O17	3.08	94.20
WAT@O	R346@HH12	R346@NH1	2.99	95.20
E305@OE1	WAT@H1	WAT@O	2.79	84.25
EST@O3	WAT@H2	WAT@O	3.1	59.21
WAT@O	R346@HH22	R346@NH2	3.18	70.80

Using the two minimized average structures reported above as receptors, compounds showed in Table 2 were docked using an automated docking procedure by means of AUTODOCK[29] (see Experimental section).

Conformational search of these ligands revealed the presence of an intramolecular H bond with the formation of a pseudocycle. The quantomechanical optimization of **L4** suggested that this interaction had a good strength, about 4 Kcal/mol (see Experimental section for details), and consequently this H bond was considered to be maintained also during interaction in the binding site. For this reason, during the AUTODOCK protocol, we blocked the torsions involved in this intramolecular bond, in order to prevent the loss of this interaction.

The docking studies revealed that, as already suggested for other ER ligands[30] these compounds could interact into ERs through two different binding orientations.

Figure 3 shows these two possible binding orientations, in both cases the originated pseudocycle the phenolic ring of estradiol, interacting with the E353(305)-R394(346)-WAT system. Concerning the other parts of the ligands in the A binding orientation the aromatic ring 1 was directed towards H524(475) and interacted with L/M384(336), M388(340), M/I421(373), I424(376) and L525(476), while the aromatic ring 2 interacted with L346(298), F404(356), M343(295). Conversely in the B binding orientation the aromatic ring 1 interacted with L346(298), F404(356), and M343(295) while the aromatic ring 2 with M388(340), L/M384(336), I424(376) and L525(476).
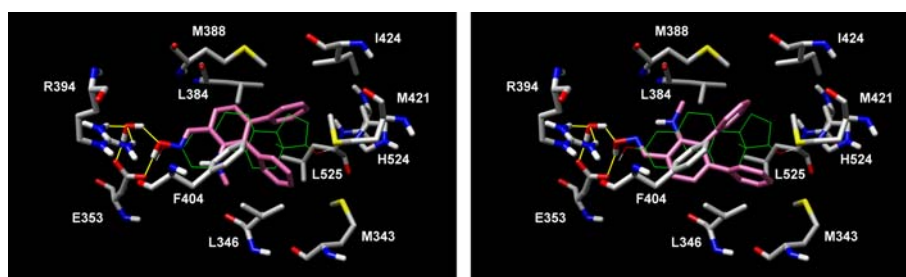

Figure 3. Docking of compound **L4** into ERα, the binding mode A is shown on the left side, while on the right is shown the binding mode B.

As shown in Table 3 all the ligands showed in the ERα, orientation A as the preferred one while in the ERβ for six ligands the preferred binding orientation was B.

However for **L4** in ERα and **L4**, **L5**, **L7**, **L11**, **L12**, **L13**, **L14**, **L16** and **L17** in ERβ, the difference between the estimated free energy of binding for the two orientation was lower than 0.5 Kcal/mol. Thus, while the AUTODOCK results could be considered reliable for indicating the preferred orientation for ERα, this was not the case for ERβ.

Table 2. Preferred binding conformation (Pref) of the ligands into ERα and ERβ. The free energy difference between orientation A and B, calculated by means of AUTODOCK scoring function, was also reported (ΔE_{A-B}). When AUTODOCK found only one of the two orientations the free energy difference was indicated as "--".

Compd	ERα		ERβ	
	Pref. Or.	**ΔE_{A-B}**	**Pref. Or.**	**ΔE_{A-B}**
L4	A	-0.39	B	+0.37
L5	A	-2.00	A	-0.44
L6	A	--	A	-0.85
L7	A	-0.66	B	+0.31
L11	A	-1.44	A	-0.37
L12	A	--	A	-0.47
L13	A	-0.99	B	+0.49
L14	A	--	B	+0.15
L15	A	--	B	+0.91
L16	A	--	B	+0.48
L17	A	--	A	-0.26

Therefore, in order to make a more precise analysis of the ligand-receptor interaction and to find more reliable results about the preferred orientation of the ligand into ERα and ERβ, all the ligand-receptor complexes were used as starting structure for 500 ps of MD simulation.

Successively the MD trajectories were further analyzed through the MM-PBSA method that has shown to accurately estimate the ligand-receptor energy interaction. This approach averages contributions of gas-phase energies, solvation free energies and solute entropies calculated for snapshots of the complex molecule as well as the unbound components, extracted from MD trajectories according to the procedure fully described in the Experimental section.

In particular, for each ligand was taken into account as starting point both the A and B orientations found by AUTODOCK in both ERs for a total of 44 calculations; in the case of the ligands for which AUTODOCK did not find any B orientation into the ERα; they were built on the basis of the B orientation found for the unsubstituted compound **L4**.

In Table 4 were reported the MM-PBSA results calculated for each ligand (A and B conformation) into both receptors. The analysis of the energies difference between the two orientations for each ligand highlighted that all the compounds showed into both receptor A orientation as the preferred one, only **L12**-ERβ complex showed an higher interaction energy when **L12** was positioned in the B orientation.

After the MD simulations the ligands interaction observed in the starting structures (obtained by means of AUTODOCK) were generally maintained. In both orientations the pseudocycle interacted with the E353(305)-R394(346)-WAT system, while the two aromatic group interacted with M343(295), L346(298), L/M384(336), M388(340), F404(356), M/I421(373), I424(376) and L525(476), as described above.

The biological data showed for all the compounds tested an ERα selectivity and, as already mentioned above, the ERα and ERβ binding site are identical except for the presence of L384 and M421 in ERα that are substituted by Met336 and I373 in ERβ. In our receptor models the compounds showed a good interaction with M421 in ERα, as for all the tested ligands the distance between the aromatic ring and the methyl group of methionine was about 3.4 Å while in ERβ the distance between I373 and aromatic ring was greater than 4 Å. Regarding L/M384(336), into both receptors this residue was far away from ligands (the distance was greater than 4 Å) and thus it affected only slightly the ligand-receptor interaction.

Therefore the interaction with the not conserved residue L384 in the ERα could be one of the reasons of the ERα selectivity of these ligands.

Table 4. MM-PBSA binding contributions for each ligand-receptor complex. In the last column was indicated the energy difference between the two orientations, calculated as difference between the calculated $\Delta G-(T\Delta S)$ of orientation A and the $\Delta G-(T\Delta S)$ of orientation B.

αER

Compd	Orientation A					Orientation B					
	ΔE_{vdW}	ΔE_{ele}	ΔG_{PB}	ΔG_{SA}	$\Delta G-(T\Delta S)$	ΔE_{vdW}	ΔE_{ele}	ΔG_{PB}	ΔG_{SA}	$\Delta G-(T\Delta S)$	$\Delta\Delta E_{A-B}$
L4	-41.69	-16.66	31.91	-3.69	-30.13	-41.54	-6.99	29.32	-3.70	-22.92	-7.21
L5	-43.32	-22.53	39.14	-3.79	-30.51	-36.76	-37.23	48.01	-3.69	-29.67	-0.84
L6	-43.70	-20.15	32.28	-4.06	-35.62	-40.37	-17.82	29.87	-4.01	-32.33	-3.29
Lg7	-45.40	-20.37	34.70	-3.94	-35.02	-44.36	-9.29	30.42	-3.89	-27.12	-7.90
L11	-40.59	-21.03	37.18	-3.78	-28.23	-40.40	-18.11	41.32	-3.78	-20.97	-7.26
L12	-42.47	-28.06	41.18	-3.87	-33.22	-42.84	-17.20	44.72	-3.76	-19.08	-14.14
L13	-45.85	-23.99	40.97	-4.17	-33.04	-46.84	-17.84	40.86	-3.99	-27.81	-5.23
L14	-44.39	-23.10	42.89	-3.81	-28.41	-40.22	-34.86	53.45	-3.81	-25.45	-2.96
L15	-46.50	-21.11	36.14	-4.01	-35.49	-47.44	-12.90	35.94	-3.98	-28.38	-7.11
L16	-44.56	-17.62	33.01	-3.79	-32.97	-40.12	-22.38	38.74	-3.80	-27.56	-5.41
L17	-40.16	-20.94	35.87	-3.54	-28.76	-38.10	-9.62	33.13	-3.58	-18.16	-10.60

βER

Compd	Orientation A					Orientation B					
	ΔE_{vdW}	ΔE_{ele}	ΔG_{PB}	ΔG_{SA}	$\Delta G-(T\Delta S)$	ΔE_{vdW}	ΔE_{ele}	ΔG_{PB}	ΔG_{SA}	$\Delta G-(T\Delta S)$	$\Delta\Delta E_{A-B}$
L4	-44.55	-13.00	30.51	-3.65	-30.69	-43.06	-5.48	28.56	-3.73	-23.71	-6.98
L5	-46.17	-18.23	36.73	-3.73	-31.40	-42.23	-34.85	51.55	-3.75	-29.29	-2.11
L6	-47.00	-13.06	30.60	-4.00	-33.46	-44.60	-29.40	46.00	-4.11	-32.12	-1.34
L7	-48.39	-10.28	31.54	-4.03	-31.16	-46.34	-8.87	30.77	-4.07	-28.52	-2.64
L11	-45.06	-13.83	34.21	-3.75	-28.43	-44.06	-8.33	36.15	-3.86	-20.10	-8.33
L12	-45.41	-25.85	47.35	-3.78	-27.70	-44.73	-38.48	57.50	-3.84	-29.55	+1.85
L13	-50.20	-17.59	37.28	-4.13	-34.64	-51.23	-12.00	39.09	-4.19	-28.33	-6.31
L14	-47.24	-20.39	44.77	-3.80	-26.67	-43.62	-14.58	41.50	-3.82	-20.51	-6.16
L15	-51.85	-14.22	34.13	-4.06	-36.01	-50.03	-10.32	34.14	-4.07	-30.29	-5.72
L16	-46.10	-11.74	30.13	-3.76	-31.47	-44.58	-10.17	32.09	-3.87	-26.53	-4.94
L17	-42.74	-12.98	34.00	-3.50	-25.22	-39.93	-8.78	31.58	-3.55	-20.68	-4.54

Figure 4a shows the superposition of the binding sites of ERα and ERβ complexed with **L5**, with the two not conserved residues L/M384(336) and M/I421(373) represented as stick. The compound formed H bonds with the E353(305)-R394(346)-WAT system through the pseudocycle group into both receptors; furthermore the R_1 hydorxy substituent formed a H bond with H524(475), simulating the interaction of 17'-hydroxy group of estradiol.

The substitution in the meta position of the two aromatic rings with an hydroxyl group determined a decrease of affinity into both receptors; Figure 4b shows the superposition of the binding sites of ERα and ERβ complexed with **L14**, which was the less active into both receptors among all the compounds tested. The figure highlights that both the meta hydroxyl substituents were unable to interact with any residue, confirming the biological results.
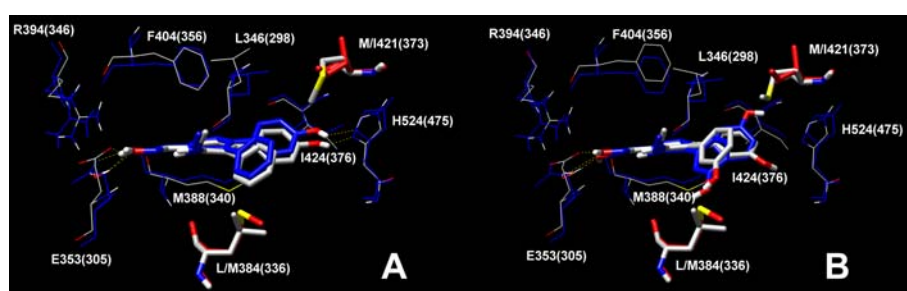
Figure 4. Superposition of the binding sites of ERα and ERβ complexed with **L5** (on the left side) and **L14** (on the right side). The ligand-ERα complex was coloured byatom while the ligand-ERβ in blue. In stick were reported the two not conserved residue L/M384(336) and M/I421(373) (coloured red in the ERβ)

L12 is the ligand with the highest ERα/ERβ selectivity and our models suggested only for this compound that the B orientation in which the R_3 hydroxy group formed a H bond with H524(475) could be the preferred orientation (see Figure 5) in the ERβ.

In order to verify if using the MM-PBSA method it was possible to extrapolate a quantitative correlation between the calculated free energy and the experimental affinity of the ligands for ERs, using the AMBER module *nmode* (see Experimental section for details) the entropic contribute was estimated for the best 11 ligand-receptor complexes. As shown in Figure 6, reporting the experimental versus the calculated free energy of

binding (see Table 5) and considering **L15** as an outlier there is a good quadratic correlation for both receptors ($R^2 = 0.83$ for ERα and $R^2 = 0.81$ for ERβ).
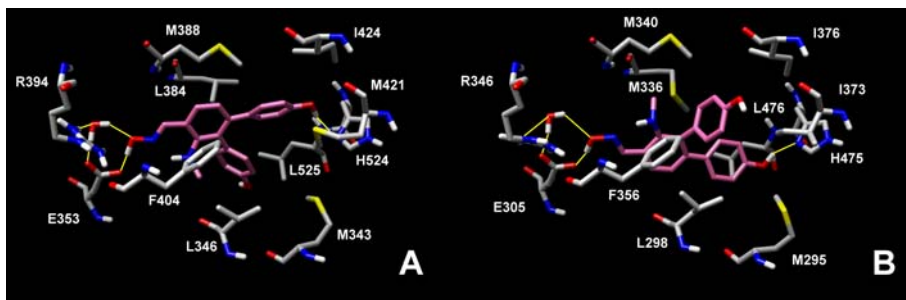

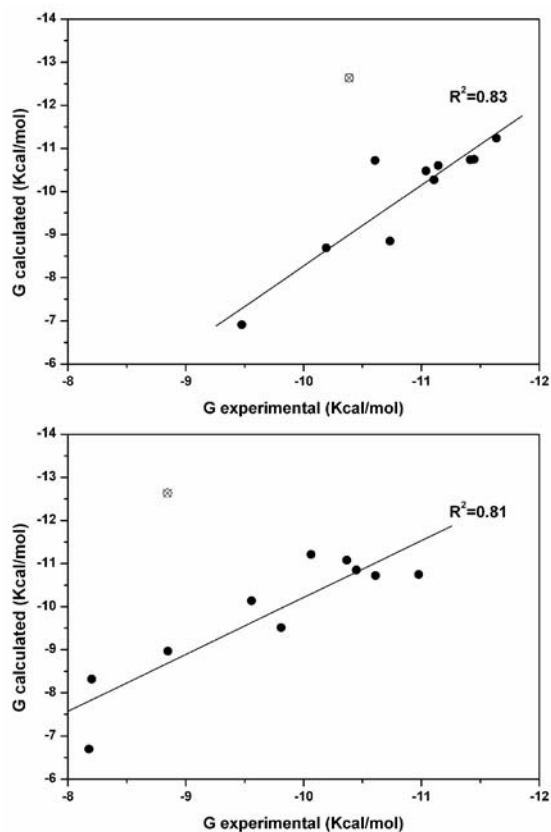**Figure 5.** Best binding orientation for **L12** into ERα (on the left side) and ERβ(on the right side).

Figure 6. Experimental free energies of binding (ΔGexp) vs predicted (ΔGcalc) for the eight ligands with ERα (left) and ERβ (right). The values are expressed as Kcal/mol.

Table 5. Energy contributions to the free energy of binding (ΔG_{cal}, expressed as Kcal/mol) for the best ligand-ER orientation of each compound. The energy difference ($\Delta \Delta G$) between the calculated (ΔG_{cal}) and the experimental free energy of binding (ΔG_{exp}) was also reported.

Compd	ERα						ERβ					
	ΔE_{MM}	ΔG_{PBSA}	$-T\Delta S$	ΔG_{cal}	ΔG_{exp}	$\Delta\Delta G$	ΔE_{MM}	ΔG_{PBSA}	$-T\Delta S$	ΔG_{cal}	ΔG_{exp}	$\Delta\Delta G$
L4	-58.35	28.22	19.39	-10.74	-11.42	0.68	-57.55	26.86	19.94	-10.75	-10.98	0.23
L5	-65.85	35.35	19.26	-11.24	-11.64	0.40	-64.40	33.00	20.19	-11.21	-10.06	-1.15
L6	-63.85	28.22	24.91	-10.72	-10.61	-0.11	-60.06	26.60	22.38	-11.08	-10.37	-0.71
L7	-65.77	30.76	24.26	-10.75	-11.45	0.70	-58.67	27.51	20.31	-10.85	-10.45	-0.40
L11	-61.62	33.40	19.37	-8.85	-10.73	1.88	-58.89	30.46	18.92	-9.51	-9.81	0.30
L12	-70.53	37.31	22.74	-10.48	-11.04	0.56	-83.21	53.66	20.58	-8.97	-8.85	-0.12
L13	-69.84	36.80	22.44	-10.60	-11.14	0.54	-67.79	33.15	24.50	-10.14	-9.56	-0.58
L14	-67.49	39.08	21.5	-6.91	-9.48	2.57	-67.63	40.97	19.96	-6.70	-8.18	1.48
L15	-67.61	32.13	22.85	-12.63	-10.39	-2.24	-66.07	30.07	22.53	-13.47	-8.85	-4.62
L16	-62.18	29.22	24.27	-8.69	-10.19	1.5	-57.84	26.37	23.15	-8.32	-8.20	-0.12
L17	-61.10	32.33	18.50	-10.27	-11.11	0.84	-58.72	30.50	17.5	-10.72	-10.61	-0.11

4.1.3 Conclusions.

In this study we present a docking analysis of estrogen ligands in ERα and ERβ. The results obtained show that the pseudocycle of the ligands used for this study is able to interact with the Arg/Water/Glu system like the estradiol molecule. Furthermore we report calculations of binding free energies between these ligands and both ERs through the recently developed MM/PBSA approach; the calculated binding free energy is in fine agreement with the experimentally determined value, showing for both receptors a quadratic correlation greater than 0.80.

Since the estrogen receptors are an interesting therapeutic target, these studies may be very useful in the search for new compounds, more active and selective for both receptors.
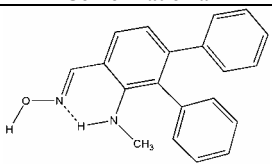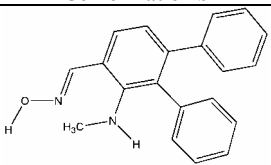
4.1.4 Experimental Section.

4.1.4.1 Amino acid numbering. For ERα and ERβ was used respectively the 1G50 and 1X7J reference numbering, when the amino acid was referred to both receptors in parenthesis was indicated the ERβ numbering.

4.1.4.2 Analysis of the compounds geometry. Conformational search of the tested compounds revealed the presence of two low energetic conformations (see Table 6) and conformation **a** showed an intramolecular H bond and the formation of a pseudocycle.

In order to measure the different stability of the two conformations and the strength of the intramolecular H bond, both conformation were optimized using the B3LYP/6-31G+* method. The energy difference between the two optimized system was about 4 Kcal/mol and this value gave us an idea of the strength of the intramolecular H bond and the better energetic stability of the conformation **a**.

Table 6. Energy difference between the two optimized conformation of **lig4**.

Conformation a	Conformation b
Δ**Energy** ($E_{conf. a}$-$E_{conf. b}$) = -4.00 Kcal/mol	

4.1.4.3 Automated docking procedure. The ligands were submitted to a conformational search of 1000 steps with an energy window for saving structure of 10 KJ/mol. The algorithm used was the Montecarlo method with MMFFs as the forcefield and a distance-dependent dielectric constant of 1.0. The first conformation that showed the intramolecular H bond was then minimized using the Conjugated Gradient method until a convergence value of 0.05 kcal/Å•mol, using the same force field and dielectric constant used for the conformational search.

Automated docking was carried out by means of the programme AUTODOCK 3.0;[29] AUTODOCK TOOLS was used to identify the torsion angles in the ligands, add the solvent model and assign the Kollman and the Gasteiger partial atomic charges to the protein and to the ligands respectively. The regions of interest used by AUTODOCK were defined by considering estradiol into both receptors as the central group; in particular, a grid of 46, 44, and 44 points in the x, y, and z directions was constructed centred on the centre of the mass of this compound. A grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the energetic map calculations. Using the Lamarckian Genetic Algorithm, the docked compounds were subjected to 100 runs of the AUTODOCK search, in which the default values of the other parameters were used. Cluster analysis was performed on the results using an RMS tolerance of 1.0 Å.

4.1.4.4 MD simulations. All simulations were performed using AMBER 8.0.[31] MD simulations were carried out using the parm94 force field at 300 K. An explicit solvent model TIP3P water was used and the complexes were solvated with a 10 Å water cap. Sodium ions were added as counterions to neutralize the system. Prior to MD simulations, two steps of minimization were carried out; in the first stage, we kept the protein fixed with a constraint of 500 Kcal/mol and we just minimized the positions of the water molecules; then in the second stage, we minimized the entire system applying a constraint of 20 Kcal/mol on the α Carbon. The two minimization stages consisted of 5000 steps in which the first 1000 were Steepest Descent (SD) and the last 4000 Conjugate Gradient (CG). Molecular dynamics trajectories were run using the minimized structure as a starting input. The time step of the simulations was 2.0 fs with a cutoff of 12 Å for the non-bonded interaction and SHAKE was employed to keep all

bonds involving hydrogen atoms rigid. Constant-volume was carried out for 100 ps, during which the temperature was raised from 0 to 300 K (using the Langevin dynamics method); then 500 ps (800 ps for the ERα and Erβ complexed with estradiol) of constant-pressure MD were carried out at 300 K. All the α Carbons were blocked with a harmonic force constant of 10 Kcal/mol•Å. The final structure of the complexes was obtained as the average of the last 400 ps (700 ps for the Erα and Erβ complexed with estradiol) of MD minimized with the CG method until a convergence of 0.05 Kcal/Å•mol.

For the ligands the force fields parameters were taken from GAFF while the atomic partial charges were derived by semiempirical AM1 geometry optimization and subsequent single point Hartree-Fock (HF)/6-31G* calculation of the electrostatic potential, to which the charges were fitted using the RESP procedure.[32]

In order to verify if using these parameters the ligands were able to maintain the intramolecular H bond and consequently the pseudocycle, the simulation protocol applied to the receptor-ligand complex as described above, were applied on compound **L4** immersed in an explicit solvent model TIP3P water with a 20 Å water cap. The minimized average of the last 400 ps of MD showed the presence of the intramolecular H bond and the pseudocycle, furthermore the superimposition of this structure with the one optimized with the QM method (conformation **a** of Table 6) showed a value of 0.40 of RMSD (evaluated on all the atoms).

4.1.4.5 Energy evaluation. We extracted from the last 400 ps of MD of the 44 ligand-ER complexes, 200 snapshots (at time intervals of 2 ps) for each species (complex, receptor and ligand). The various MM-PBSA energy terms in equation **1** were computed as follows.

$$\textbf{eq1} \qquad G = G_{polar} + G_{nonpolar} + E_{mm} - TS$$

Electrostatic, van der Waals and internal energies (E_{mm}) were obtained using the SANDER module in Amber8.0. Polar energies (G_{polar}) were obtained from the PBSA module of Amber8.0 program (using the Poisson-Boltzman method) applying a

dielectric constants of 1 and 80 to represent gas and water phases, respectively. Nonpolar energies ($G_{nonpolar}$) were determined using the MOLSURF program.

In order to compare the energetic interactions of the two orientations (A and B in Figure 3) of the same ligand into the same receptor we took into account only the first three terms of equation 1, considering approximately constant the entropic value.

Solute entropy was evaluate only for the best orientation of each ligand into the two receptors, in order to correlate the predicted free energy of binding (calculated as in equation **2**) with the experimental one. It was estimated using the NMODE module of Amber8.0 on a total of 10 snapshots. Prior the normal mode calculations, each species (complex, receptor, or ligand) was subjected to a CG energy minimization using a distance dependent dielectric, until a convergence of 0.00001 Kcal/mol Å.

$$\textbf{eq2} \qquad \Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand})$$

4.2 REFERENCES.

1. Katzenellenbogen, J. A.; Katzenellenbogen, B. S. Chem. Biol. 3 (1996) 529-536.

2. Cauley, J. A.; Stelley, D. G.; Ensrud, K.; Ettinger, B.; Black, D.; Cummings, S. R. Ann. Intern. Med. 122 (1995) 9-16.

3. Sherwin, B. S. Neurology 48 (1997) S21-26.

4. Cosman, F.; Lindsay, R. Endocr. Rev. 20 (1999) 418-434.

5. Barrett-Connor, E.; Cox, D. A.; Anderson, P. W. Trends Endocrinol. Metab. 10 (1999) 320-325.

6. Yaffe, K.; Sawaya, G.; Lieberburg, I.; Grady, D. J. Am. Med. Assoc. 279 (1998) 688-695.

7. Davidson, N. E.; Lippman, M. E. Crit. Rev. Oncog. 1 (1989) 89-111.

8. Zumoff, B. Proc. Soc. Exp. Biol. Med. 217 (1998) 30-37.

9. Beresford, S. A.; Weiss, N. S.; Voigt, L. F.; McKnight, B. Lancet 349 (1997) 458-461.

10. Meegan, M. J.; Hughes, R. B.; Lloyd, D. G.; Williams, D. C.; Zisterer, D. M. J. Med. Chem. 44 (2001) 1072-1084.

11. Grese, T. A.; Dodge, J. Curr. Pharm. Des. 4 (1998) 71-92.

12. Lin, X.; Huebner, V. Curr. Opin. Drug Discovery Dev. 3 (2000) 383-398.

13. Shang, Y.; Brown, M. Science 295 (2002) 2465-2468.

14. Katzenellenbogen, B. S.; Katzenellebogen, J. K. Science 295 (2002) 2380-2381.

15. Kuiper, G. G. J. M.; Carlsson, B.; Grandien, K.; Enmark, E.; Haggblad, J.; Nilsson, S.; Gustafsson, J.-Å. Endocrinology 138 (1997) 863-870.

16. Mosselman, S.; Polman, J.; Dijekema, R. FEBS Lett. 392 (1996) 49-53.

17. Kuiper, G. G. J. M.; Enmark, E.; Pelto-Huikko, M.; Nilsson, S.; Gustafsson, J.-Å. Proc. Natl Acad. Sci. USA 93 (1996) 5925-5930.

18. Pike, A. C. W.; Brzozowski, A. M.; Hubbard, R. E.; Bonn, T.; Thorsell, A.-G.; Engstrom, O.; Ljunggren, J.; Gustafsson, J.-A¡ .; Carlquist, M. EMBO J. 18 (1999) 4608-4618.

19. Kuiper, G. G. J. M.; Gustafsson, J.-Å. FEBS Lett. 410 (1997) 87-90.

20. Hewitt, S. C.; Couse, J. F.; Korach, K. S. Breast Cancer Res. 2 (2000) 345-352.

21. Couse, J. F.; Korach, K. S. Endocr. Rev. 20 (1999) 358-417.

22. Dupont, S.; Krust, A.; Gansmuller, A.; Dierich, A.; Chambon, P.; Mark, M. Development 127 (2000) 4277-4291.

23. Katzenellenbogen, J. A.; O'Malley, B. W.; Katzenellenbogen, B. S. Mol. Endocrinol. 10 (1996) 119-131.

24. McKenna, N. J.; O'Malley, B. W. Cell 108 (2002) 465-474.

25. Eiler, S.; Gangloff, M.; Duclaud, S.; Moras, D.; Ruff, M. Protein Expr.Purif. 22 (2001) 165-173.

26. Manas, E.S.; Xu, Z.B.; Unwalla, R.J.; Somers, W.S. Structure 12 (2004) 2197-2207.

27. Fiser A., Do R.K., Sali A. Science 9 (2000) 1753-1773.

28. Brzozowski, A.M., Pike, A.C., Dauter, Z., Hubbard, R.E., Bonn, T., Engstrom, O., Ohman, L., Greene, G.L., Gustafsson, J.A., Carlquist, M. Nature 389 (1997) 753-758.

29. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. J. Comp. Chem. 19 (1998) 1639-1662.

30. Mewshaw, R.E.; Edsall, R.J. Jr; Yang, C.; Manas, E.S.; Xu, Z.B.; Henderson, R.A.; Keith, J.C. Jr; Harris, H.A. J Med Chem. 16 (2005) 3953-3979.

31. Amber version 8.0; University of California at San Francisco: San Francisco, CA.

32. Bayly, C. I.; Cieplak, P.; Cornell, W.D.; Kollman, P. A. J. Phys. Chem. 97 (1993) 10269-10280.

CHAPTER 5

ANGIOTENSIN RECEPTORS

146

5.1 PROPOSAL OF A NEW BINDING ORIENTATION FOR NON-PEPTIDE ANTAGONISTS: AT1 HOMOLOGY MODELING, DOCKING AND 3D-QSAR ANALYSIS.

5.1.1 Introduction.

Renin-Angiotensin-Aldosterone System (RAAS) is a proteolytic cascade that plays an important role in the electrolyte homeostasis and in the regulation of blood pressure but is also involved in the pathogenesis of hypertension and renal diseases.

The RAAS begins with the release of the aspartic protease renin from the juxtaglomerular cells of the kidney. This enzyme is responsible for the conversion of angiotensinogen to the inactive decapeptide angiotensin I . In turn, angiotensin I is cleaved by Angiotensin Converting Enzyme (ACE) to produce the octapeptide Angiotensin II (AngII) which constitutes the main effector hormone of RAAS.

AngII is the major regulator of blood pressure, electrolyte balance, and endocrine functions related to cardiovascular diseases such as hypertension. Moreover, it has been shown that AngII plays a role in various pathological situations involving tissue remodeling, such as cardiac hypertrophy. Recent findings;[1,2] indicate the involvement of this peptide also in cancer.

AngII affects most of the biological functions by activating selective membrane-bounded receptors. Two distinct subtypes of AngII receptors [type 1 (AT1) and type 2 (AT2)] have been identified, and both belong to the G protein-coupled receptor superfamily (GPCRs).

AT1 and AT2 are seven transmembrane-spanning receptors comprising an extracellular glycosilated region connected to the seven transmembrane α-helices linked by three intracellular and three extracellular loops. The carboxy-terminal domain of the protein is cytoplasmic and is a regulatory site. AT1 is a 359 aminoacids protein, while AT2 is formed by 363 amino acids and is 30% homologous to AT1. Both receptors are N-linked glycosylated post-translationally.

AT1 receptors are expressed in various parts of the body and mediate all of the known effects associated to AngII such as vasocostriction, aldosterone release, and other functions that tend to elevate blood pressure and cause hypertrophy and hyperplasia of

target cells. In contrast, AT2 receptors are highest expressed in fetal tissue in which are responsible to mediate organ remodelling, and are poorly represented in the adult, appearing only at the site of inflammation, tissue damage or other forms of cellular stress such as ischaemia. AT2 receptor seems to be responsible of both the inhibition of cell growth and the promotion of apoptosis.[3] Although the role of AT2 receptor is less understood, it is usually retained that it exerts opposing effects to the AT1 receptor.

Given the important role played by RAAS in the hypertension, this system constitutes the principal target of an effective therapy. The first choice class of drugs influencing the RAAS targeting is represented by ACE-inhibitors.

These drugs block the formation of AngII and also prevent the conversion of bradykinin to inactive peptides. Although bradykinin may contribute to the beneficial effects of ACE inhibitors through its vasorelaxing effect, its accumulation determines some disadvantages such as the development of cough and angioedema which are side-effects often associated with the ACE-inhibitors therapy. Moreover ACE-inhibitors do not completely suppress AngII, because its formation is ensured also by ACE-indipendent pathways. For these reasons, it became important the development of AngII subtype 1 receptor antagonists (AT1 antagonists or sartans) as a new class of antihypertensive used in the treatment of hypertension,[4] heart failure[5] and renal diseases.[6] Although precise mechanisms have not yet been elucidated to explain all of the beneficial effects, sartans are unique in their ability to provide such benefits with a limited side-effect profile.[7]

The first non-peptide AT1 antagonist which represents the prototype of the sartans was losartan. The major active metabolite of losartan, EXP3174, generated by the oxidation of the 5-hydroxymethyl group on the imidazole ring, is 10-40 times more potent than losartan itself and is therefore responsible for the majority of its pharmacological activity.

At this time, many selective, potent and orally available sartans have been developed and are used to treat both hypertension and damage associated with diseases such atherosclerosis and diabetes.

Although all sartans bind with high affinity to the AT1 receptor and share a common mechanism of action, there are differences in the modes of interactions with the

receptor.[8,9] Sartans such as losartan, eprosartan and tasosartan bind to receptor displaying degrees of surmountability. Valsartan, irbesartan, candesartan and the active metabolite of losartan (EXP3174) are termed insurmountable antagonist, in contrast to surmountable antagonists which do not impair the maximum response to AngII.

A possible explanation for this different response is that the surmountable antagonists interfere with receptor activation by occupying an intramembrane site that overlaps with the space occupied by the agonist; while insurmountable antagonists induce conformational changes that prevent the agonist binding. Another proposal hypothesizes that surmountable antagonists dissociate rapidly from receptor, whereas insurmountable antagonists bind tightly and dissociate so slowly as to cause functional loss of the occluded receptors.[10]

A knowledge of the 3D structure of AT1 receptor could be of great help in the task of understanding the antagonists interaction and in the rational design of specific ligands; however since GPCRs are membrane-bound proteins, high-resolution structural characterization is still an extremely difficult task.

Several studies have been performed to investigate the binding site of the AT1 receptor and the disposition of its residues; using all these information in the present study a homology model of the human AT1 receptor was constructed and surmountable and insurmountable antagonists were docked.

Furthermore in order to obtain a quantitative model able to measure the reliability of the constructed receptor and also to provide a predictive system exploitable for designing of new AT1 antagonists, a 3D-QSAR model was calculated, basing on the alignment obtained docking several ligands into the AT1 receptor.

5.1.2 Results and Discussion.

5.1.2.1 Homology modeling. AT1 receptor model was generated using the recent bovine rhodopsin crystal structure determined at 2.2 Å (1U19[11]) as the template. The sequence alignment was studied on several Angiotensin receptors. As shown in Figure 1 the alignment was guided by the highly conserved amino acid residues, including the asparagine residues N1.50, the LA-AD (L2.46, A2.47, A2.49, and D2.50) and D/ERY--V (D/E3.49, R3.50, Y3.51, and V3.54) motif, the highly conserved triptophane W4.50,

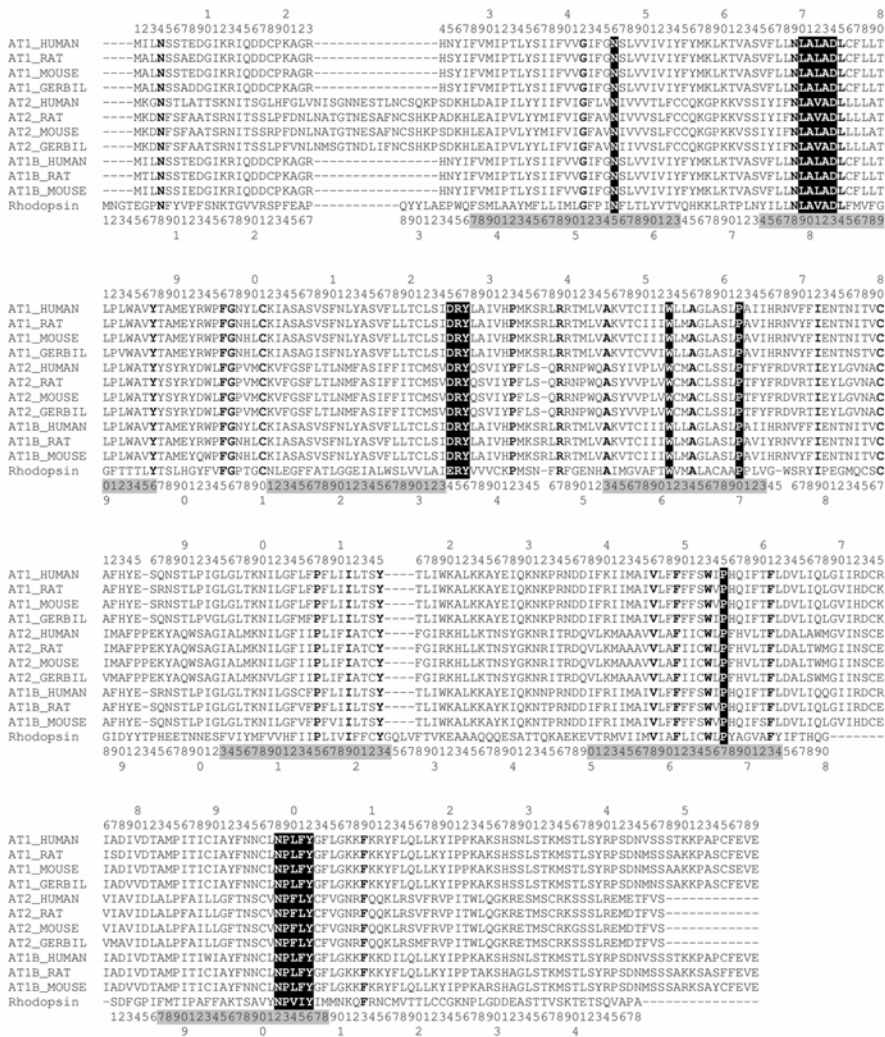the two proline P4.59 and P6.50, and the NPXXY motif in TM7 (N7.49, P7.50, and Y7.53).[12]

```
                1         2                    3         4         5         6         7         8
         123456789012345678901234567890123    456789012345678901234567890123456789012345678901234567890
AT1_HUMAN   ----MILNSSTEDGIKRIQDDCPKAGR--------------HNYIFVMIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLT
AT1_RAT     ----MALNSSAEDGIKRIQDDCPKAGR--------------HSYIFVMIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLT
AT1_MOUSE   ----MALNSSTEDGIKRIQDDCPRAGR--------------HSYIFVMIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLT
AT1_GERBIL  ----MALNSSADDGIKRIQDDCPKAGR--------------HSYIFVMIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLT
AT2_HUMAN   ----MKGNSTLATTSKNITSGLHFGLVNISGNNESTLNCSQKPSDKHLDAIPILYYIIFVIGFLVNIVVVTLFCCQKGPKKVSSIYIFNLAVADLLLLAT
AT2_RAT     ----MKDNFSFAATSRNITSSLPFDNLNATGTNESAFNCSHKPADKHLEAIPVLYYMIFVIGFAVNIVVVSLFCCQKGPKKVSSIYIFNLAVADLLLLAT
AT2_MOUSE   ----MKDNFSFAATSRNITSSRPFDNLNATGTNESAFNCSHKPSDKHLEAIPVLYYMIFVIGFAVNIVVVSLFCCQKGPKKVSSIYIFNLAVADLLLLAT
AT2_GERBIL  ----MKDNFSFAATSRNITSSLPFVNLNMSGTNDLIFNCSHKPSDKHLEAIPVLYLIFVIGFAVNIIVVSLFCCQKGPKKVSSIYIFNLAVADLLLLAT
AT1B_HUMAN  ----MILNSSTEDGIKRIQDDCPKAGR--------------HNYIFVMIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLT
AT1B_RAT    ----MTLNSSTEDGIKRIQDDCPKAGR--------------HNYIFVMIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLT
AT1B_MOUSE  ----MILNSSIEDGIKRIQDDCPKAGR--------------HNYIFVMIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLT
Rhodopsin   MNGTEGPNFYVPFSNKTGVVRSPFEAP----------QYYLAEPWQFSMLAAYMFLLIMLGFPINFLTLYVTVQHKKLRTPLNYILLNLAVADLFMVFG
         1234567890123456789012345678      890123456789012345678901234567890123456789012345678901234567890
                1         2                       3         4         5         6         7         8
```

```
                 9         0         1         2         3         4         5         6         7         8
         123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890
AT1_HUMAN   LPLWAVYTAMEYRWPFGNYLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTCIIIWLLAGLASLPAIIHRNVFFIENTNITVC
AT1_RAT     LPLWAVYTAMEYRWPFGNHLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTCIIIWLMAGLASLPAVIHRNVYFIENTNITVC
AT1_MOUSE   LPLWAVYTAMEYRWPFGNHLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTCIIIWLMAGLASLPAVIHRNVYFIENTNITVC
AT1_GERBIL  LPVWAVYTAMEYRWPFGNHLCKIASAGISFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTCVVIWLLAGLASLPAVIHRNVYFIENTNSTVC
AT2_HUMAN   LPLWATYYSYRYDWLFGPVMCKVFGSFLTLNMFASIFFITCMSVDRYQSVIYPFLS-QRRNPWQASYIVPLVWCMACLSSLPTFYFRDVRTIEYLGVNAC
AT2_RAT     LPLWATYYSYRYDWLFGPVMCKVFGSFLTLNMFASIFFITCMSVDRYQSVIYPFLS-QRRNPWQASYVVPLVWCMACLSSLPTFYFRDVRTIEYLGVNAC
AT2_MOUSE   LPLWATYYSYRYDWLFGPVMCKVFGSFLTLNMFASIFFITCMSVDRYQSVIYPFLS-QRRNPWQASYVVPLVWCMACLSSLPTFYFRDVRTIEYLGVNAC
AT2_GERBIL  LPLWATYYSYRYDWLFGPVMCKVFGSFLTLNMFASIFFITCMSVDRYQSVIYPFLS-QRRNPWQASYVVPLVWCMACLSSLPTFYFRDVRTIEYLGVNAC
AT1B_HUMAN  LPLWAVYTAMEYRWPFGNYLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTCIIIWLLAGLASLPAVIHRNVFFIENTNITVC
AT1B_RAT    LPLWAVYTAMEYRWPFGNHLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTCIIIWLMAGLASLPAVIYRNVYFIENTNITVC
AT1B_MOUSE  LPLWAVYTAMEYQWPFGNHLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTCIIIWLMAGLASLPAVIHRNVYFIENTNITVC
Rhodopsin   GFTTTLYTSLHGYVFGPTGCNLEGFFATLGGEIALWSLVVLAIERYVVVCKPMSN-FRFGENHAIMGVAFTWVMALACAAPPLVG-WSRYIPEGMQCSC
         012345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345
                 9         0         1         2         3         4         5         6         7         8
```

```
                 9         0         1              2         3         4         5         6         7
         12345 67890123456789012345       6789012345678901234567890123456789012345678901234567890123
AT1_HUMAN   AFHYE-SQNSTLPIGLGLTKNILGFLFPFLIILTSY----TLIWKALKKAYEIQKNKPRNDDIFKIIMAIVLFFFFSWIPHQIFTFLDVLIQLGIIRDCR
AT1_RAT     AFHYE-SRNSTLPIGLGLTKNILGFLFPFLIILTSY----TLIWKALKKAYEIQKNKPRNDDIFRIIMAIVLFFFFSWVPHQIFTFLDVLIQLGVIHDCK
AT1_MOUSE   AFHYE-SRNSTLPIGLGLTKNILGFLFPFLIILTSY----TLIWKALKKAYEIQKNKPRNDDIFRIIMAIVLFFFFSWVPHQIFTFLDVLIQLGVIHDCK
AT1_GERBIL  AFHYE-SQNSTLPVGLGLTKNILGFMFPFLIILTSY----TLIWKALKKAYEIQKNKPRNDDIFRIIMAIVLFFFFSWVPHQIFTFLDVLIQLGVIHDCK
AT2_HUMAN   IMAFPPPEKYAQWSAGIALMKNILGFIIPLIFIATCY----FGIRKHLLKTNSYGKNRITRDQVLKMAAAVVLAFIICWLPFHVLTFLDALAWMGVINSCE
AT2_RAT     IMAFPPPEKYAQWSAGIALMKNILGFIIPLIFIATCY----FGIRKHLLKTNSYGKNRITRDQVLKMAAAVVLAFIICWLPFHVLTFLDALTWMGIINSCE
AT2_MOUSE   IMAFPPPEKYAQWSAGIALMKNILGFIIPLIFIATCY----FGIRKHLLKTNSYGKNRITRDQVLKMAAAVVLAFIICWLPFHVLTFLDALTWMGIINSCE
AT2_GERBIL  VMAFPPPEKYAQWSAGIALMKNVLGFIIPLIFIATCY----FGIRKHLLKTNSYGKNRITRDQVLKMAAAVVLAFIICWLPFHVLTFLDALSWMGIINSCE
AT1B_HUMAN  AFHYE-SRNSTLPIGLGLTKNILGSCFPFLIILTSY----TLIWKALKKAYEIQKNNPRNDDIFRIIMAIVLFFFFSWIPHQIFTFLDVLIQQGIIRDCR
AT1B_RAT    AFHYE-SQNSTLPIGLGLTKNILGFVFPFLIILTSY----TLIWKALKKAYKIQKNTPRNDDIFRIIMAIVLFFFFSWVPHQIFTFLDVLIQLGIIRDCE
AT1B_MOUSE  AFHYE-SQNSTLPIGLGLTKNILGFVFPFVIILTSY----TLIWKALKAYKIQKNTPRNDDIFRIIMAIVLFFFFSWVPHQIFSFLDVLIQLGVIHDCE
Rhodopsin   GIDYYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTVKEAAAQQQESATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQG-------
         890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890
                 9         0         1              2         3         4         5         6         7
```

```
                8         9         0         1         2         3         4         5
         67890123456789012345678901234567890123456789012345678901234567890123456789
AT1_HUMAN   IADIVDTAMPITICIAYFNNCLNPLFYGFLGKKFKRYFLQLLKYIPPKAKSHSNLSTKMSTLSYRPSDNVSSSTKKPAPCFEVE
AT1_RAT     ISDIVDTAMPITICIAYFNNCLNPLFYGFLGKKFKKYFLQLLKYIPPKAKSHSSLSTKMSTLSYRPSDNMSSSAAKKPASCFEVE
AT1_MOUSE   IADIVDTAMPITICIAYFNNCLNPLFYGFLGKKFKKYFLQLLKYIPPKAKSHSSLSTKMSTLSYRPSDNMNSSAAKKPASCSEVE
AT1_GERBIL  IADVVDTAMPITICIAYFNNCLNPLFYGFLGKKFKKYFLQLLKYIPPKAKSHSSLSTKMSTLSYRPSDNMNSSAAKKPASCFEVE
AT2_HUMAN   VIAVIDLALPFAILLGFTNSCVNPLFYCFVGNRFQQKLRSVFRVPITWLQGKRESMSCRKSSSLREMETFVS-----------
AT2_RAT     VIAVIDLALPFAILLGFTNSCVNPLFYCFVGNRFQQKLRSVFRVPITWLQGKRETMSCRKGSSLREMDTFVS-----------
AT2_MOUSE   VIAVIDLALPFAILLGFTNSCVNPLFYCFVGNRFQQKLRSVFRVPITWLQGKRETMSCRKGSSLREMDTFVS-----------
AT2_GERBIL  VMAVIDLALPFAILLGFTNSCVNPLFYCFVGNRFQQKLRSMFRVPITWLQGKRETMSCRKSSSLREMDTFVS-----------
AT1B_HUMAN  IADIVDTAMPITIWIAYFNNCLNPLFYGFLGKKFKDILQLLKYIPPKAKSHSNLSTKMSTLSYRPSDNVSSSTKKPAPCFEVE
AT1B_RAT    IADIVDTAMPITICIAYFNNCLNPLFYGFLGKKFKKYFLQLLKYIPPTAKSHAGLSTKMSTLSYRPSDNMSSSAAKKSASFFEVE
AT1B_MOUSE  IADVVDTAMPITICIAYFNNCLNPLFYGFLGKKFKRYFLQLLKYIPPKAKSHAGLSTKMSTLSYRPSDNMSSSARKSAYCFEVE
Rhodopsin   -SDFGPIFMTIPAFFAKTSAVYNPVIYIMMNKQFRNCMVTTLCCGKNPLGDDEASTTVSKTETSQVAPA----------------
         123456789012345678901234567890123456789012345678012345678
                8         9         0         1         2         3
```

Figure 1. Alignment of the angiotensine receptors and bovine rhodopsin (OPSD_BOVIN) amino acid sequences. The highly conserved patterns of the LA-AD (L2.46, A2.47, A2.49, and D2.50), the D/ERY motif (D/E3.49, R3.50, and Y3.51), the highly conserved triptophane W4.50, the two proline P4.59 and P6.50, and the NPXXY motif in TM7 (N7.49, P7.50, and Y7.53) are marked with black. The other identical residues are in bold. In the first and last line of the alignment scheme were reported the numeration of the human AT1 and bovine rhodopsin respectively, while the TM of bovine rhodopsin were reported in grey.

Using this alignment the AT1 receptor model was built and was then subjected to a simulated annealing protocol by means of Modeller program.[13] The backbone conformation of the best scored structure was evaluated with the PROCHECK software[14] (see Experimental section for details) and the analysis of the Psi/Phi Ramachandran plot indicated that only two amino acids of the loop fragments (H24 and S186) had a disallowed geometry.

The obtained results suggested that the created molecular model of the AT1 receptor could be used for further studies.

5.1.2.2 Docking of losartan. Site directed mutagenesis suggested an important role for many residues, in particular the affinity of losartan seemed to be principally influence by the presence of K3.24(102), K5.42(199),[15] V3.32(108), A4.60(163),[16] and N7.46(295)[17] (even if this last residue could be responsible of the conformational changes that occur in the AT1 receptor activation[18,19]).

The analysis of the disposition of these residues in the AT1 receptor model highlighted that, with the exception of K3.24(102) which was far away and directed toward the extracellular side of the receptor, all the others residues were comprised in a limited region compatible for the interaction with losartan. Furthermore aligning the AT1 receptor model with the bovine rhodopsin crystal structure came out that the residues listed above (principally V3.32(108), K5.42(199) and N7.46(295)) delimited a region which corresponded to the binding site of retinal.

Keeping all these things in mind, the non peptide antagonist losartan was docked into the AT1 receptor model using the AUTODOCK program[20] and building a "docking box" that included the main mutagenesis data (see Experimental section for details).

The best docked structure of losartan highlighted an interaction of the hydroxymethyl substituent with K5.42(199), V3.32(108) interacted with the biphenyl system, A4.60(163) with the *n*-butyl substituent, while N7.46(295) seemed to not interact with the ligand.

Finally the tetrazole ring did not appear to interact with any residue suggested important by mutagenesis studies but instead it was stabilized by a H bond with Y184.

All the recent AT1 receptor models suggested for the tetrazole ring an interaction with K5.42(199),[17,21] however as there were none experimental evidence able to exclude others binding dispositions we proceeded with further refinements and evaluations.

The losartan-AT1 receptor complex obtained through the automated docking procedure was refined by means of 1 ns of molecular dynamics (MD) simulation.

A vacuum MD simulation can lead to severe distortions, especially of the loop structures, and the primary source of these distortions appear to be formation of artifactual H-bonds.[22] Furthermore a vacuo MD simulation requires the use of a set of restraints, to replace the natural stabilizing effects of the membrane bilayer on the TM domains, reducing the free movement of the helices. In order to avoid these problems we carried out the simulation in a fully hydrated phospholipid bilayer environment constituted by dipalmitoylphosphatidylcholine (DPPC) molecules solvated by TIP3 water molecules; as described in the experimental section.

The system contains 192 DPPC molecules, 7410 water molecules, 16 Chlorine atoms, the AT1 receptor and losartan, for a total of 52940 atoms.

The stability of the model was evaluated by calculating the total energy of the system; as shown in Figure 2, after 400 ps of MD, the system reached an equilibrium, since the total energy for the last 600 ps remained approximately constant. Analysing the root mean square deviation (RMSD) of all the α Carbons of the TM helices from the starting AT1 model structure, we observed that after an initial increase, in the last 400 ps the RMSD remained included between the value of 1.5 and 1.8 Å (see Figure 2B), suggesting that our MD procedure was correct.
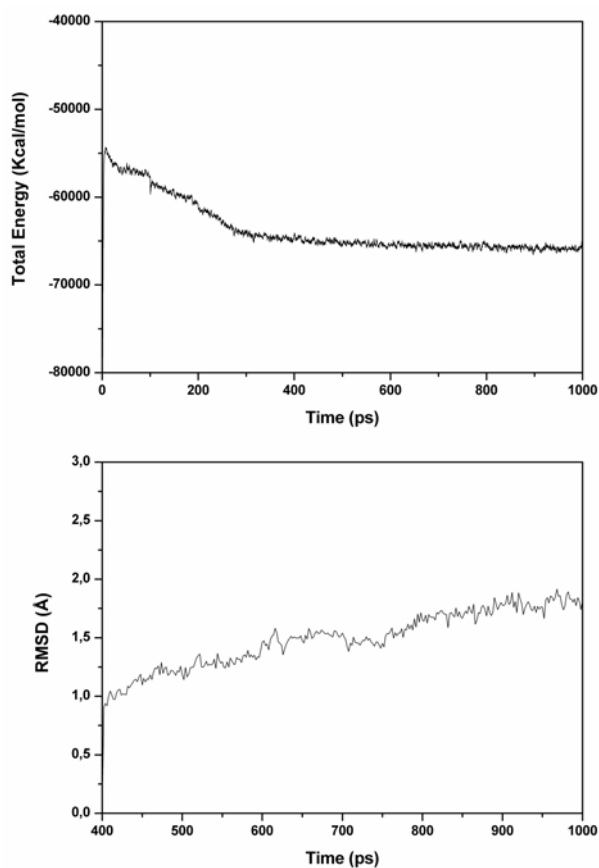
Figure 2. Analysis of the MD simulation of losartan complexed with AT1. In the first plot the Total energy of the system vs the time was reported; in the second plot the RMSD between all the heavy atoms of the receptor was reported

Figure 3 shows the AT1-losartan complex embedded into DPPC bilayer, the binding site was limited by TM3, TM4, TM5, TM6, TM7 and the second extracellular loop (EL2). Regarding losartan it showed the tetrazole ring and the hydroxymethyl substituent turned in the direction of the extracellular side of the receptor.

Biophysical studies showed that losartan interacts with the interface of phospholipid membranes;[23,24] basing on these studies Zoumpoulakis et al. hypothesized that losartan bound onto the receptor after a first step which involves incorporation and interaction with membrane bilayers. Regarding the location of losartan into phospholipid core, it was found to locate itself near the interface and showed a H bond between the

hydroxymethyl group and the phospholipid phosphate group while the tetrazole substituent interacted with four $N^+(CH_3)_3$ head groups of the phospholipid. The binding conformation of losartan obtained in our MD simulation and its disposition inside the receptor were compatible with this possible mechanism of interaction. The position of the antagonist was in the upper part of the receptor, in the region comprised between the TM3-7 helices and furthermore the losartan conformation inside the receptor was similar to the one suggested able to interact into the membrane.[21]
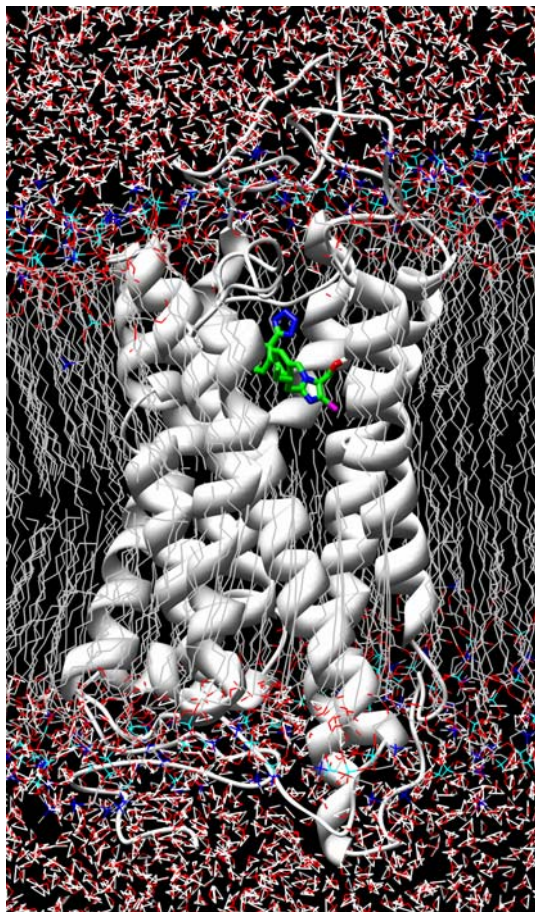
Figure 3. Losartan-AT1 receptor complex inserted into the DPPC bilayer model

Figure 4 shows the binding site of losartan into AT1 receptor. The biphenyl ring of the antagonist was positioned among TM3, TM6 and TM7, in a lipophilic cavity principally

delimited by V3.32(108), V179, W6.48(253), H6.51(256), I7.39(288) and Y7.43(292). The anionic tetrazole ring was directed towards the extracellular side of the receptor and interacted with T175 and Y184, that are two residue of EL2, and formed a third H bond with H6.51(256). Regarding the 2'-butyl substituent it was directed towards TM4 and interacted in a secondary lipophilic pocket constituted by S3.33(109), L3.36(112), Y3.37(113), A4.60(163), F171 and F182 of EL2 while the hydroxymethyl group formed a H bond with K5.42(199).
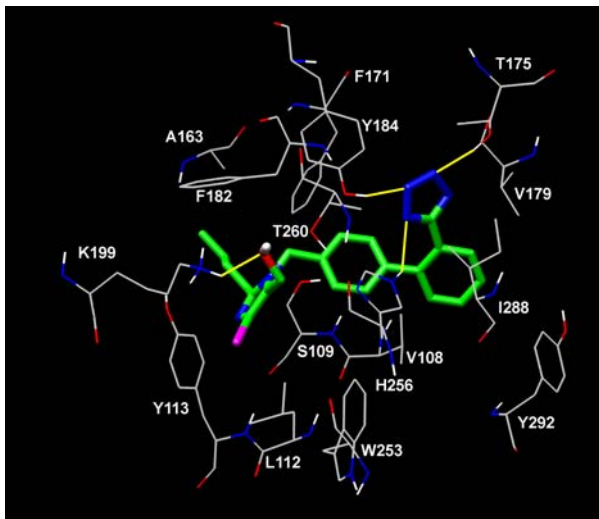
Figure 4. Losartan docked into the AT1 binding sites. Interatomic distances between H-bonded atoms are indicated in yellow.

In the last three years studies on the AT1 receptor using the substituted-cysteine accessibility method (SCAM)[25,26] and the methionine proximity assay[27] investigated the disposition of the residues belonging to TM3, TM6, TM7 and part of EL3. More in details these studies suggested that A3.28(104), N3.35(111), L3.36(112), Y3.37(113), F6.44(249), W6.48(253), H6.51(256), T6.55(260), A277, V280, T7.x(282), A7.34(283), I7.37(286), A7.42(291), F7.44(293), N7.45(294), N7.46(295), C7.47(296), L7.48(297), and F7.52(301) should oriented themselves within the water-accessible crevice of the AT1 receptor and in our receptor model among all these residues only A7.34(283), I7.37(286) and F7.44(293) were not oriented within the water-accessible crevice. However regarding F7.44(293), the authors suggested that the terminal part of the TM7

structure is somewhat destabilized, permitting outward pointing residues to turn intermittently inwards.[27]

As regards the interaction of losartan into AT1 receptor, site directed mutagenesis suggested the importance of V3.32(108).[16] In 1995 Noda et al[15] revealed that mutation of K5.42(199) with alanine, glutamate and arginine determined more than tenfold decreased of binding affinity, while the mutation of lysine with glutamine determined only a twofold decrease of losartan affinity. Regarding H6.51(256) Takezako et al[17] highlighted that the mutation with alanine determined an important affinity decrease only when accompanied by the K5.42(199)A mutation, suggesting a complementary role for H6.51(256), otherwise Noda et al[15] for the mutation H6.51(256)A reported a twofold decreases of affinity of losartan. Thus all these studies were in agreement with our AT1-losartan model.

Other mutagenesis studies suggested for the residues D2.50(74), N3.35(111), S3.39(115), N7.45(294) and N7.46(295) a fundamental role in the activation process of the receptor.[19,28-32] In our model these residues were not directly involved in the interaction with losartan, however as shown in Figure 5 they interacted each other constituting a H bond network system able to connect TM2, TM3 and TM7, probably controlling the inactive-active state of the receptor.

Thus these observations were in agreement with the hypothesis of a "structural role" for N7.46(295) instead of a direct interaction with losartan.[18,19]
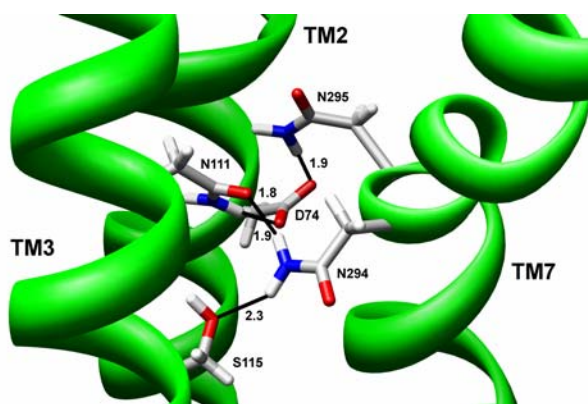
Figure 5. H bond network system of TM2, TM3 and TM7

5.1.2.3 Docking of insurmountable antagonists. In order to investigate the binding characteristics derived by the interaction of an insurmountable antagonist, starting from the AT1 receptor model obtained through the simulated annealing protocol we docked by means of Autodock the active metabolite of losartan (EXP3174) (see Experimental section for details). Then, the complex of best docked structure with the AT1 receptor were subjected to 1 ns of MD using the same protocol applied for losartan.

The results obtained from the MD simulation showed that the binding disposition of the ligand was similar to the one of losartan, with the biphenyl ring located into the lipophilic task delimited by V3.32(108), V179, W6.48(253), H6.51(256), I7.39(288) and Y7.43(292) and 2'-butyl substituent that interacted in the secondary lipophilic pocket constituted by S3.33(109), L3.36(112), Y3.37(113), A4.60(163), F171 and F182. However as shown in figure 6, respect to losartan, EXP3174 was shifted about 2.7 Å towards TM5. This disposition determined the loss of the H bond of tetrazole with T175, which resulted far away (5.3 Å from the tetrazole ring), while the interaction with Y184 and H6.51(256) were maintained.

Regarding the imidazole ring, it showed a new H bond with Y3.37(113) while the carboxylate group interacted with K5.42(199) and moreover it formed a second H bond with Q6.52(257)
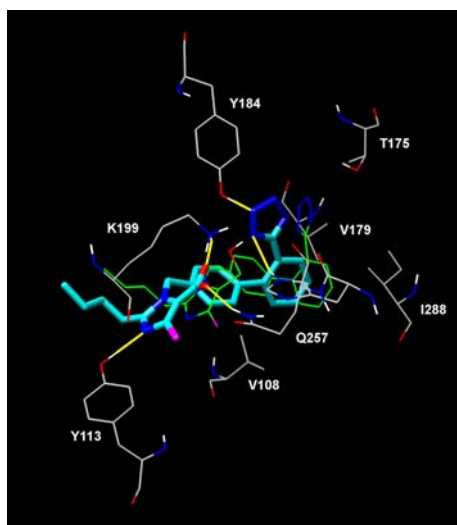
Figure 6. Superimposition between EXP3174 and losartan (coloured blue and green respectively) docked into the AT1 binding sites.

From this analysis came out that even though EXP3174 did not interact with T175, like losartan; however it showed two additional interaction with Y3.37(113) and Q6.52(257), moreover the carboxylate group formed a ionic interaction with K5.42(199) which should be stronger with respect to the H bond interaction of the hydroxymethyl group of losartan.

Overall, it seemed to appear that the EXP3174-AT1 binding interaction was stronger than the losartan-AT1 interaction; therefore these observation were in agreement with the hypothesis that, differently from surmountable, insurmountable antagonists could bind tightly and dissociate so slowly as to cause functional loss of the occluded receptors.

Furthermore the interaction of the carboxylate group with K5.42(199) and Q6.52(257) were in agreement with site directed mutagenesis data, which suggested for these two residues a fundamental role in the insurmountable antagonism binding.[14,33]

In order to verify if the interaction with K5.42(199) and Q6.52(257) was only a peculiarity of EXP3174 or, as suggested by mutagenesis data it was showed also by other insormountable antagonists, using the receptor obtained through the simulation with EXP3174, the other three insormountable antagonists candesartan, irbesartan and valsartan (see Figure 7) were docked through an automated docking procedure (see Experimental section for details).
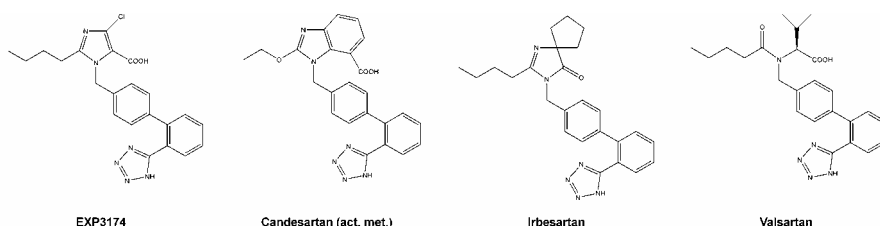
Figure 7. Insurmountable antagonists used in the docking study.

As shown by Figure 8 all the ligands exhibited a binding disposition very similar to EXP3174, furthermore K5.42(199) and Q6.52(257) were involved in the binding of all the three ligands, interacting with the carboxylate group of candesartan and valsartan and the carbonylic oxygen of irbesartan.
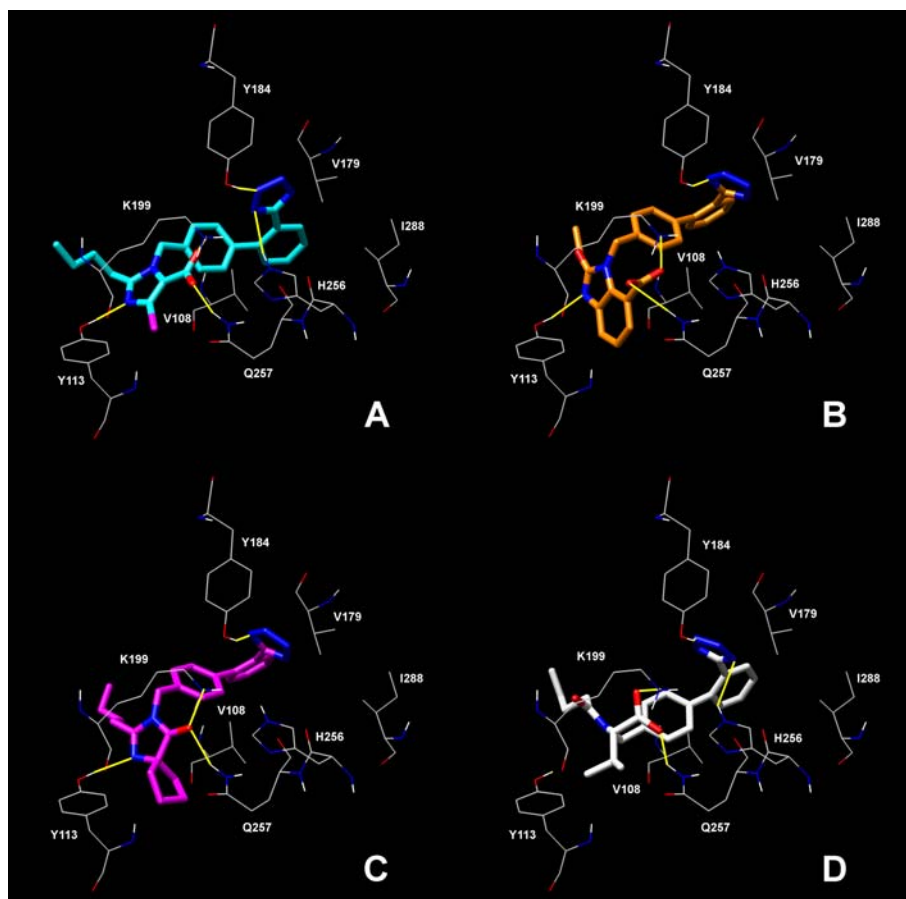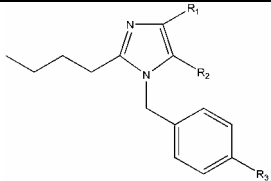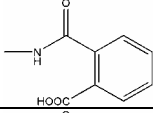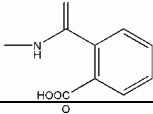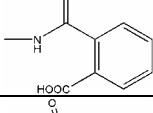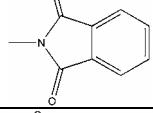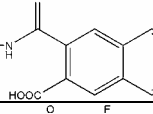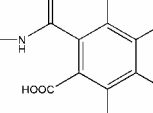
Figure 8. Docking of EXP3174 (A), candesartan (B), ibesartan (C), valsartan (D) docked into the AT1 binding site.
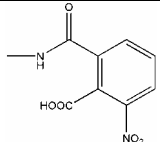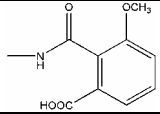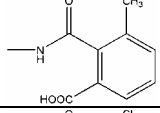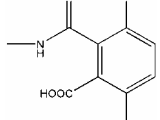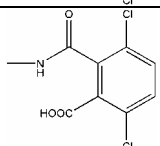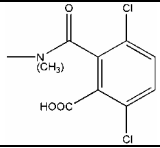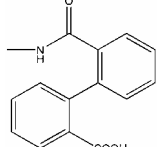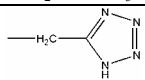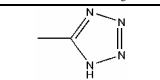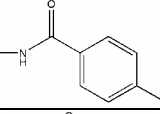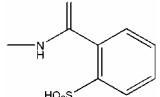
5.1.2.4 3D-QSAR. In order to verify the reliability of our modelled AT1 receptor and also to obtain a quantitative model able to predict the activity of new ligands, 63 non-peptide antagonists[34-36] (see Table 1) were docked into the AT1 receptor, and for each ligand the best docked conformation was used for the development of a 3D-QSAR model (see Experimental section).
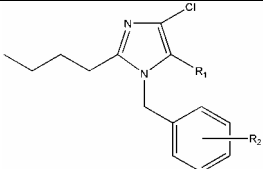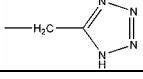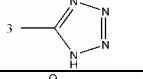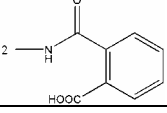
Table 1. Ligands used for the 3D-QSAR study.
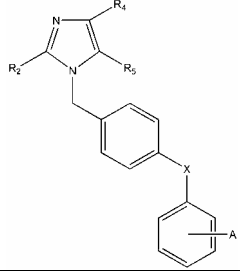
Compd	R_1	R_2	R_3	IC_{50} (μM)
4	Cl	CH_2COOH	COOH	1.6
5a	Cl	CH_2COOCH_3	NH_2	100
6	CH_2COOH	Cl	COOH	19
7[*]	Cl	CH_2OH	COOH	1.7
8	CH_2OH	Cl	COOH	100
9	Cl	CH_2OAc	COOH	5.3
10	Cl	CH_2COOH	CH_2COOH	13
11	Cl	CH_2COOCH_3	NO_2	100
12	Cl	CH_2COOCH_3	$NHCH_2Ph$	40
13	Cl	CH_2OH	CHO	28
15[*]	Cl	CH_2OH	OCH_3	100
17	Cl	CH_2COOCH_3	$NHCO(CH_2)_2COOH$	46
18[*]	Cl	CH_2COOCH_3	$NHCO(CH_2)_3COOH$	32
19	Cl	CH_2COOCH_3	NHCOCH=CHCOOH (cis)	11
22	Cl	CH_2COOCH_3		0.14
23[*]	CH_2COOCH_3	Cl		0.42
24	Cl	CH_2OCH_3		0.28
25	Cl	CH_2COOCH_3		3.1
26	Cl	CH_2COOCH_3		5.8
27	Cl	CH_2COOCH_3		0.79

160
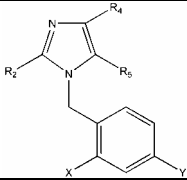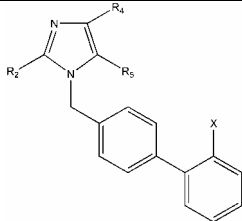
30	Cl	CH$_2$COOCH$_3$		0.40
31	Cl	CH$_2$COOCH$_3$		0.08
32	Cl	CH$_2$COOCH$_3$		0.032
33	Cl	CH$_2$COOCH$_3$		0.018
34[*]	Cl	CH$_2$OCH$_3$		0.042
35[*]	Cl	CH$_2$OCH$_3$		5.70
36	Cl	CH$_2$COOCH$_3$		0.19
38	Cl	CH$_2$COOCH$_3$	NHCOCF$_3$	27
39	Cl			1.2
48	Cl	CH$_2$COOCH$_3$		11
53	Cl	CH$_2$OCH$_3$		0.012

Compd	R₁	R₂	IC₅₀ (μM)
54	CH_2COOH	3-COOH	100
55	CH_2COOH	2-COOH	38
56			3.8
59	CH_2COOCH_3		100

Compd	R₂	R₄	R₅	X	A	IC₅₀ (μM)
2a	n-C₄H₉	Cl	CH_2OH	Single bond	3-CO₂H	0.49
2b	n-C₄H₉	H	CH_2OH	Single bond	3-CO₂H	1.1
2c	n-C₄H₉	Cl	CH_2OCOCH_3	Single bond	3-CO₂H	2.5
2d	n-C₄H₉	Cl	CH_2OCH_3	Single bond	3-CO₂H	2.9
3a	n-C₄H₉	Cl	CH_2OH	CO	2-CO₂H	0.16
3b	n-C₄H₉	CH_2OH	Cl	CO	2-CO₂H	0.34
3c	n-C₄H₉	CH_2OCOCH_3	Cl	CO	2-CO₂H	1.4
3d	n-C₄H₉	Cl	$CH_2NHCO_2CH_3$	CO	2-CO₂H	0.27
3e[*]	n-C₄H₉	Cl	CH_2OCH_3	CO	2-CO₂H	0.15
4x	n-C₄H₉	Cl	CH_2OH	O	2-CO₂H	0.40
5	n-C₄H₉	Cl	CH_2OH	S	2-CO₂H	0.40
6a	n-C₄H₉	Cl	CH_2OH	OCH₂	2-CO₂H	0.92
6b	n-C₄H₉	H	CH_2OH	OCH₂	2-CO₂H	0.31
6c	n-C₄H₉	Cl	CH_2OCOCH_3	OCH₂	2-CO₂H	1.8
6d	n-C₄H₉	Cl	CH_2OCH_3	OCH₂	2-CO₂H	1.2
6e	n-C₃H₇S	H	CH_2OH	OCH₂	2-CO₂H	5.9
6f	n-C₂H₅S	H	CH_2OH	OCH₂	2-CO₂H	12
7x	n-C₄H₉	Cl	CH_2OH	Trans-CH=CH	2-CO₂H	5.4

162

Compd	R$_2$	R$_4$	R$_5$	X	Y	IC$_{50}$ (μM)
S-8307[*]	n-C$_4$H$_9$	Cl	CH$_2$CO$_2$Na	Cl	H	40
S-8308[*]	n-C$_4$H$_9$	Cl	CH$_2$CO$_2$Na	NO$_2$	H	13
EXP6155	n-C$_4$H$_9$	Cl	CH$_2$CO$_2$Na	H	CO$_2$Na	1.6

Compd	R$_2$	R$_4$	R$_5$	X	IC$_{50}$ (μM)
EXP7711	n-C$_4$H$_9$	Cl	CH$_2$OH	CO$_2$H	0.3
EXP8823	n-C$_4$H$_9$	Cl	CO$_2$H	CO$_2$H	0.092
EXP9019[*]	n-C$_4$H$_9$	Cl	CHO	CO$_2$H	0.94
EXP9020	n-C$_4$H$_9$	H	CHO	CO$_2$H	0.55
EXP9270	C$_2$H$_5$CH=CH	Cl	CH$_2$OH	CO$_2$H	0.08
EXP9654[*]	C$_2$H$_5$CH=CH	Cl	CHO	CO$_2$H	0.33
losartan	n-C$_4$H$_9$	Cl	CH$_2$OH		0.019

The 3D-QSAR analysis was carried out using C3, OH probe and a combination of them. To measure the goodness of the different probes, the three 3D-QSAR models were characterized by correlation coefficient (r^2), predictive correlation coefficient (q^2) and cross-validated standard deviation of errors of prediction (SDEP$_{CV}$). As summarized in Table 2 the model obtained through OH probe seemed to be the best one.

Furthermore an external test set of ten antagonists (marked with a "[*]" in Table 1) were used to test the predictive ability of the models and this analysis confirmed that the OH probe was the best selection (SDEP$_{Test-Set}$=0.54, see Table 2 and Figure 9).

Table 2. Statistical results of the 3D-QSAR models.

Model	Grid probes	Vars	PC	r^2	q^2	$SDEP_{cv}$	$SDEP_{Test-Set}$
1	C3	1168	5	0.98	0.72	0.57	0.64
2	OH	1086	4	0.97	0.76	0.52	0.54
3	C3 and OH	2011	5	0.98	0.74	0.54	0.60
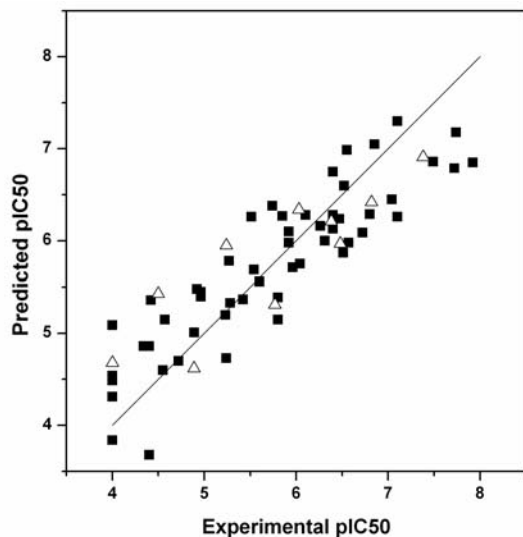
Figure 9. Plot of model 2: experimental/predicted pIC_{50} is reported, the test set is represented as Δ.

One important feature of 3D QSAR analysis is the graphical representation of the model, usually aimed at making its interpretation easier. In the GOLPE program, there are several options for displaying the final model. Among these PLS pseudocoefficient and the activity contribution plots are very useful. PLS coefficient plot allows the visualization of favourable and unfavourable interactions between the probes and the molecules under study, while the activity contribution plot is different for every molecule within the training set and gives the possibility to display spatial regions that are individually important for the selected molecule. In Figure 10A and 10B are reported the PLS coefficients plot of model 2. Figure 10A shows negative PLS coefficients; in particular there is eight principal regions (A-H) with negative values in which a favourable interaction between a substituent and the probe determines an increase of activity whereas unfavourable interaction between a substituent and the probe determines a decrease of activity. Otherwise positive PLS coefficients reported in

Figure 10B (yellow surface) show areas where a favourable interaction between a substituent and the probe determines a decrease of activity whereas unfavourable interaction between a substituent and the probe determines an increase of activity; in this picture three main regions (A', G', H') and other six secondary regions (B'-F', I') where recognized.
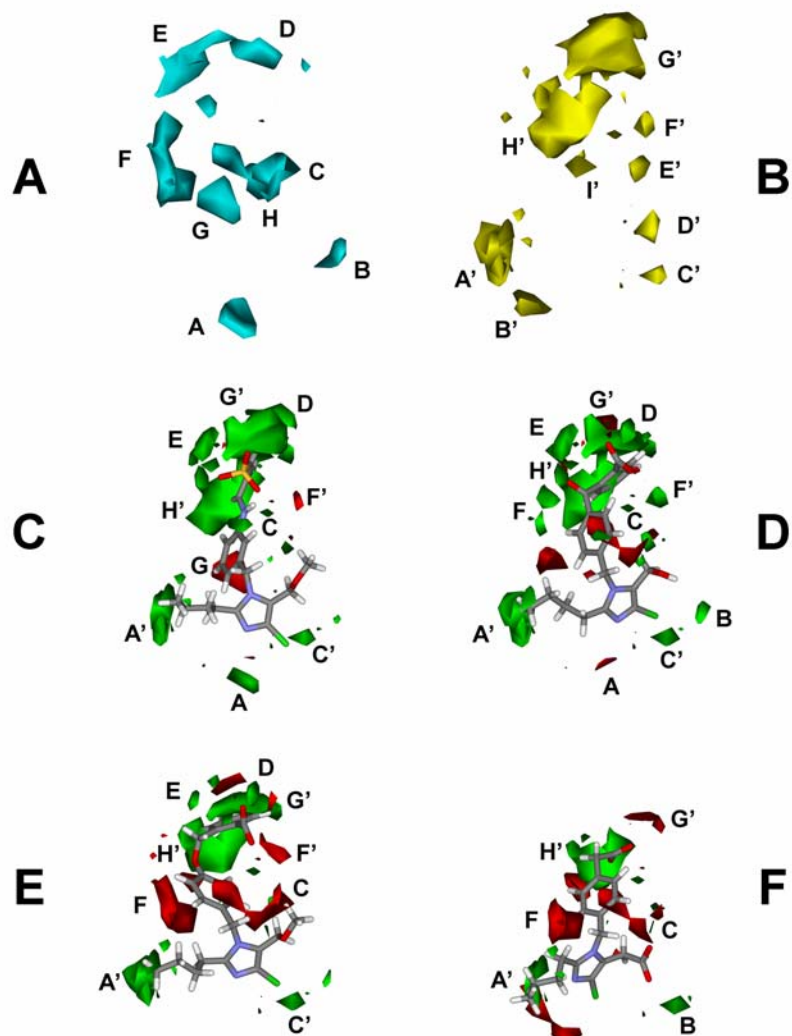

Figure 10. Negative (A) and positive (B) regions of the PLS coefficient plot obtained with the OH probe. The activity contributions plots (green polyhedrons represent positive contributions whereas red polyhedrons negative contributions to the activity) for compounds **53** (C), **3a** (D), **6a** (E) and **10** (F) were also reported.

165

Figure 10C-10F reported compounds **53**, **3a**, **6a** and **10** embedded in their activity plot contributions; positive contribution to the activity is coloured as a green polyhedron while negative contribution to the activity is coloured as a red polyhedron.

Figure 10C highlights that the sulfonate group of **53** (characterized by the best IC_{50} value, 0.012 µM) make favourable interaction with the regions C, D and E, while the unsubstituted nitrogen of the imidazole group interacts with region A. Concerning the yellow regions that give a positive contribution to the activity plot, they could be considered as lipophilic contributions that increase the activity of the ligand; for **53** the n-butyl chain make favourable interactions with region A', the sulfonate phenyl ring with H' and G' while the chlorine atom with region C'.

Compound **3a** is about 13 fold less potent then **53** and it is characterized by the presence of 2-benzoylbenzoate substituent instead of the biphenyl-2-sulfonate and the hydroxymethyl substituent instead of the methoxymethyl. Figure 10D shows that the carboxylate function makes favorable interactions with E and D regions, while the hydroxymethyl group interacts with B region and the carbonylic function makes favourable interactions with the F zone. However, respect to **53** the presence of 2-benzoylbenzoate group determines the shift of the ligand with the loss of the interaction with C and A regions. As regards the steric contributions the interaction with A', C', G' and H' were maintained.

The presence of 2-(phenoxymethyl)benzoate instead of 2-benzoylbenzoate determines a decrease of activity, compound **6a** differs from **3a** only for the presence of this group and it is about six fold less active. As shown in figure 10E the absence of the carbonylic function determines the lost of the favourable interaction with the F region and a reduction of the interaction of the carboxylic group with the E region.

Compound **10** shows a low activity (IC50=13 µM), and in fact as shown in figure 10F the substitution of the benzoate group with an acetate determines the loss of the interaction with the E, D and G' regions.

As the alignment of the ligands were performed using the structures docked into the AT1 receptor, it would be useful to check for matching between the AT1 receptor and the 3D-QSAR maps.

In Figure 11 the binding site of the AT1 receptor is overlapped with the 3D PLS coefficient maps. There is a close match between the receptor and the PLS surfaces: the A region is in proximity of Y113, region B correspond to Q6.52(257), while C is comprised between K5.42(199) and Y184. Region D corresponds to T175, the E region is comprised between the backbone of C180 and S3.29(105) while F is comprised between S3.29(105) and S3.33(109). As regards the yellow surfaces, regions A' and B' correspond to the secondary lipophilic pocket principally constituted by Y3.3.37(113), A4.60(163) and F171 and region E' corresponds to H6.51(256). Finally the main regions G' and H' are in close proximity of V179 and I7.39(288) the first one while region H' corresponds to V3.32(108).
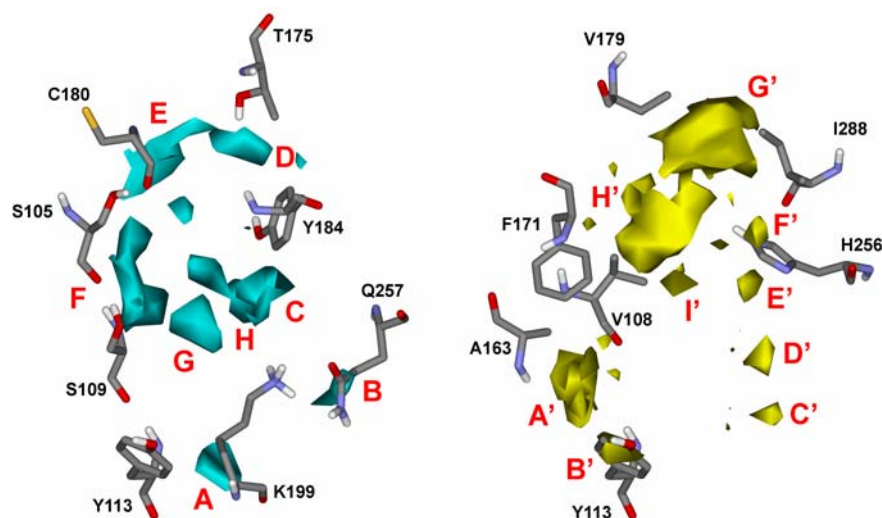

Figure 11. PLS coefficient plots obtained with the OH probe superimposed to the AT1 receptor binding site

5.1.3 Conclusions.

We have constructed the 3D model of the inactive conformation of the AngII receptor AT1, based on crystallized bovine rhodopsin (1U19[11]).

The docking of losartan into AT1 receptor confirmed that K5.42(199), V3.32(108), and A4.60(163)[15,16] interacted with the ligand, in agreement with mutagenesis data. Regarding the anionic tetrazole ring of losartan, it did not appear to interact with any residue suggested important by mutagenesis data and in particular, in contrast with the

recent published AT1 homology model, it did not interact with K5.42(199).[17,21] Otherwise it was principally stabilized by a H bond with Y184 and further interactions with T175 and H6.51(256).

Residues D2.50(74), N3.35(111), S3.39(115), N7.45(294) and N7.46(295) interacted each other constituting a H bond network system able to connect TM2, TM3 and TM7: as site direct mutagenesis suggested for them a fundamental role in the activation process of the receptor[18,19,29-31] they probably controlled the inactive-active state of the receptor interacting each other.

Regarding the insurmountable interaction of EXP3174, it was explained through a strong binding interaction, moreover the model was in agreement with the important role of the interaction of K5.42(199) and Q6.52(257) with the insurmountable antagonists EXP3174, candesartan, valsartan and irbesartan.

Finally in order to measure the reliability of the constructed AT1 receptor and also to provide a quantitative model useful for the prediction of the affinity of new designed antagonists, using the alignment obtained through an automated docking procedure, 63 non-peptide antagonists were included into a 3D-QSAR model and the results showed a good correlation.

In conclusion our AT1 receptor model proposed a new hypothesis about the binding interaction of the non-peptide antagonists inside the AT1 receptor, encouraging further analyses for new residues which could be fundamental for the ligand-receptor interaction. Furthermore as the AT1 antagonists are an interesting therapeutic target the information obtained, combined with the calculated 3D-QSAR model could allow a quantitative evaluation of new designed non-peptide antagonists.

5.1.4 Experimental Section.

5.1.4.1 Amino acid numbering. To refer to specific amino acids sequences, the numbering system suggested by Ballesteros and Weinstein is used.[37]

The most highly conserved residue in each transmembrane helix (TMH) is assigned a value of 0.50 and this number is preceded by the TMH number and followed in parentheses by the sequence number. The other residues in the helix are given a locant value relative to this.

5.1.4.2 Homology modeling. The crystal structure of bovine rhodopsin was taken from the Protein Data Bank,[38][nota 45 cann] while all the primary sequences were obtained from the SWISS-PROT protein sequence database.[39]

The sequential alignment of bovine rhodopsin and the AT receptors was performed by means of CLUSTAL W,[40] using the Blosum series as a matrix, with a gap open penalty of 10 and a gap extension penalty of 0.05.

The TM helices, the first and second intracellular and extracellular loop of the AT1 receptor were constructed directly from the coordinates of the corresponding amino acids in rhodopsin by means of Modeller program.[13] The other loop regions, as the amino acid length differs from the template were constructed by means of the "Loop optimization method" of Modeller, applying the "very_slow" loop refinement method. During the construction of the receptor and the loop refinement, the presence of a disulfide bridge between C101 and C180 was taken into account as it was present in the bovine rhodopsin; furthermore, as suggested by mutagenesis studies[41] we considered also the formation of a disulfide bridge between C18 and C274.

Starting from this receptor ten structures were generated, through the "very slow MD annealing" refinement method, as implemented in Modeller and on the basis of the DOPE assess method the best receptor model was chosen. The backbone conformation of the resulting receptor structure was evaluated by inspection of the Psi/Phi Ramachandran plot obtained from PROCHECK analysis.[14]

5.1.4.3 Docking of losartan and EXP3174. The ligands were submitted to a conformational search of 1000 steps with an energy window for saving structure of 10 KJ/mol by means of MACROMODEL program.[42] The algorithm used was the Montecarlo method with MMFFs as the forcefield and a distance-dependent dielectric constant of 1.0. The ligands were then minimized using the Conjugated Gradient method until a convergence value of 0.05 kcal/Å•mol, using the same forcefield and dielectric constant used for the conformational search. Both ligands were docked into AT1 receptor using the AUTODOCK 3.0 programme.[20] The regions of interest used by AUTODOCK were defined by considering atom CZ3 of W6.48(253) as the central residue of a grid of 50, 44, and 48 points in the x, y, and z directions so that the main residues suggested important by site directed mutagenesis were considered. A grid

spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the energetic map calculations.

Using the Lamarckian Genetic Algorithm, the compound was subjected to 100 runs of the AUTODOCK search, in which the default values of the other parameters were used. Cluster analysis was performed on the docked results using an RMS tolerance of 1.0 Å. The best two docking conformations were complexed with AT1 receptor and then subjected to MD simulations.

5.1.4.4 MD simulations. All simulations were performed using AMBER 8.0.[43] The two complexes were embedded into a phospholipid bilayer previously stabilized constituted by DPPC molecules. The receptor-ligand complexes were manually inserted into the enter of the DPPC bilayer in such a way that the α helices of the receptor were oriented approximately parallel to the hydrocarbon chains of the phospholipids. After that all phospholipids within a radius of 1 Å around the receptor were deleted.

MD simulations were carried out using the modified parm94 force field at 300 K. An explicit solvent model TIP3P water was used and the system were solvated on the "extracellular" and "intracellular" side with a 15 Å water cap. Chlorine ions were added as counterions to neutralize the system. Prior to MD simulations, three steps of minimization were carried out; in the first stage, we kept the protein and phospholipids fixed with a constraint of 500 Kcal/mol and we just minimized the positions of the water molecules; then in the second stage, we minimized the phospholipids-water system applying a constraint of 500 Kcal/mol on the protein, and finally in the last step we apply a constraint of 50 Kcal/mol only on the Cα of the receptor. The three minimization stages consisted of 5000 steps in which the first 1000 were Steepest Descent (SD) and the last 4000 Conjugate Gradient (CG). Molecular dynamics trajectories were run using the minimized structure as a starting input, and the particle mesh Ewald (PME) algorithm was used for dealing with long-range interactions.[44] The time step of the simulations was 2.0 fs with a cutoff of 12 Å for the non-bonded interaction and SHAKE was employed to keep all bonds involving hydrogen atoms rigid. Constant-volume was carried out for 100 ps, during which the temperature was raised from 0 to 300 K (using the Langevin dynamics method); then 900 ps of constant-pressure MD were carried out at 300 K. In the first 400 ps of MD all the α Carbons of

the receptor were blocked with a harmonic force constant that during these 400 ps decrease from 50 to 1 Kcal/mol•Å, while in the last 600 ps there were no constraints. The final structure of the complexes was obtained as the average of the last 500 ps of MD minimized with the CG method until a convergence of 0.05 Kcal/ Å•mol.

General Amber Force Field (GAFF) parameters were assigned to ligands and DPPC molecules, while the partial charges were calculated using the AM1-BCC method as implemented in the Antechamber suite of Amber 8.

The phospholipid bilayer system was previously stabilized by 600 ps of MD using the same parameters described above. Prior to MD simulations, two steps of minimization were carried out; in the first stage, we kept the phospholipids fixed with a constraint of 500 Kcal/mol and we just minimized the positions of the water molecules; then in the second stage, we minimized the phospholipids-water system applying a constraint of 100 Kcal/mol on the heavy atom of the phospholipids. In the first 200 ps of MD all the heavy atoms of the DPPC molecules were blocked with a harmonic force constant that during these 200 ps decreases from 100 to 10 Kcal/mol•Å, while in the last 400 ps there were no constraints. The structure of the bilayer system in which were embedded the two AT1 receptor complexes was obtained as the average of the last 300 ps of MD minimized with the CG method until a convergence of 0.05 Kcal/ Å•mol.

5.1.4.6 Docking of insurmountable antagonists. Candesartan, irbersartan and valsartan were docked into AT1 receptor using the minimized average of the last 500 ps of MD simulations of the AT1-EXP3174 complex as receptor. The regions of interest used by AUTODOCK were defined by considering EXP3174 complexed into AT1 as the central group; in particular, a grid of 40, 40, and 48 points in the *x*, *y*, and *z* directions was constructed centred on the centre of the mass of this antagonist. A grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the energetic map calculations.

Using the Lamarckian Genetic Algorithm, the docked compounds were subjected to 100 runs of the AUTODOCK search, in which the default values of the other parameters were used. Cluster analysis was performed on the results using an RMS tolerance of 1.0 Å, and the best docked structures were considered.

5.1.4.7 3D QSAR. *5.1.4.7.1 Alignment of the molecules.* The ligands showed in Table 1 were docked into AT1 receptor using the same procedure seen above for the insurmountable antagonists, but the minimized average of the last 500 ps of MD simulations of the AT1-Losartan complex were used as receptor. For each ligand the best docked structure was chosen and this receptor-based alignment was used for the further studies.

All the ligands reported in literature that showed more than 6 AUTODOCK atom types were not taken into consideration for the limits of the software.

5.1.4.7.2 Data set. The GOLPE program[46] was used to define three 3D QSAR models, using GRID interaction fields [nota] as descriptors (see below). The training set were composed by 53 compounds characterized by affinity values spanning about 4 orders of magnitude, the minimum value of 4.00 (expressed as $-\log IC_{50}$) being associated with compounds **11**, **54**, **59**, **5a** and **8** and the maximum value of 7.92 being associated with compound **53**. Similarly, compounds belonging to the test set showed affinity value ranging from 4.0 (compound **15**) to 7.38 (compound **34**) and were uniformly distributed along the activity range (see Figure 9).

5.1.4.7.3 Probe selection. The GRID program[46] was used to describe the previously superposed molecular structure. Interaction energies between selected probes and each molecule were calculated using a grid spacing of 1 Å. The C3 (corresponding to a methyl group), OH (corresponding to a phenolic OH group) and a combination of them were used to calculate the molecular interaction fields (MIFs).

5.1.4.7.4 Variable selection. The MIFs of the training set were imported in GOLPE; it is well known that many of the variables derived from GRID analysis could be considered as noise, which decreases the quality of the model. For this reason variable selection was operated by zeroing values with absolute values smaller than 0.06 Kcal/mol and removing variables with standard deviation below 0.1. Moreover, variables which exhibited only two values and had a skewed distribution were also removed.

The smart region definition (SRD) algorithm[47] was applied with 10% of the active variables as number of seed (selected in the PLS weights space), a critical distance cutoff of 2.5 Å, and collapsing distance cutoff of 4.0 Å. The groups were then used in the Fractional Factorial Design (FFD) procedure. FFD selection was applied two time,

until the r^2 and q^2 value did not increase significantly using the cross-validation routine with five random sets of compounds.

5.2 REFERENCES

1. Juillart-Jeanneret, L. Br. J. Cancer 90 (2004) 1059-1068.

2. Fogarty, D. Glia 39 (2002) 304-313.

3. Herion, D.; Kubis, N.; Levy, B. Hypertension 38 (2001) 1150-1157.

4. Weber, A. J. Renin Angiotensin Aldosterone Syst. 4 (2003) 62.73.

5. Manohar, P.; Pina, I.L. Mayo Clin. Proc. 78 (2003) 334-338.

6. Norris, K.; Vaughn, C. Expert Rev. Cardiovasc. Ther. 1 (2003) 51-63.

7. Mimran, A.; Ribstein, J.; J. Am. Soc. Nephrol. 10(Suppl.12) (1999) S273-S277).

8. Panek, R.I.; Lu, G.H.; Overhisser, R.W. J. Pharm. Exp. Ther. 273 (1995) 753-761.

9. Timmermans, P.B. Can. J. Cardiol. 15(supp.F) (1999) 26F-28F.

10. De Gasparo, M.; Catt, K.J.; Inagami, T.; Wright, J.W.; Unger, T. Pharm. Rev. 52 (2000) 415-472.

11. Okada, T.; Sugihara, M.; Bondar, A.N.; Elstner, M.; Entel, P., Buss, V. J.Mol.Biol. 342 (2004) 571-583.

12. Baldwin, J.M.; Schertler, G.F.X.; Unger, V.M. J. Mol. Biol. 272 (1997) 144-164.

13. Fiser, A.; Do, R.K.; Sali, A. Science 9 (2000) 1753-1773.

14. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thorton, J.M. J. Appl. Crystallogr. 26 (1993) 283–291.

15. Noda, K.; Saad, Y.; Kinoshita, A.; Boyle, T.P.; Graham, R.M.; Husain, A.; Karnik, S.S. J. Biol. Chem. 270 (1995) 2284-2289.

16. Ji, H.; Leung, M.; Zhang, Y.; Catt, K.J.; Sandberg K. J. Biol. Chem. 269 (1994) 16533-16536.

17. Takezako, T.; Gogonea, C.; Saad, Y.; Noda, K.; Karnik, S.S. J. Biol. Chem. 279 (2004) 15248-15257.

18. Balmforth, A.J.; Lee, A.J.; Warburton, P.; Donnelly, D.; Ball, S.G. . J. Biol. Chem. 272 (1997) 4245-4251.

19. Hines, J.; Fluharty, S.J;. Yee, D.K. Biochem. Pharm. 66 (2003) 251-262.

20. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. J. Comp. Chem. 19 (1998) 1639-1662.

21. Zoumpoulakis, P.; Daliani, I.; Zervou, M.;. Kyrikou, I.; Siapi, E.; Lamprinidis, G.; Mikros, E.; Mavromoustakos, T. Chem. Phys.Lip. 125 (2003) 13-25.

22. Mehler, E.L.; Periole, X.; Hassan, S.A.; Weinstein, H. J comp-aid mol des 16 (2002) 841-853.

23. Mavromoustakos, T.; Daliani, I.; Matsoukas, J.; a. Bioactive Peptides in Drug Discovery and Design: Medical Aspects 22. (1999) IOS Press-Ohmsha, pp. 13–24.

24. Theodoropoulou, E.; Marsh, D. Biochim. Biophys. Acta 1461 (1999) 135–146.

25. Boucard, A.A.; Roy, M.; Beaulieu, M.E.; Lavigne, P.; Escher, E.; Guillemette, G.; Leduc, R. J Biol Chem. 278 (2003) 36628-36636.

26. Martin, S.S.; Boucard, A.A.; Clement, M.; Escher, E.; Leduc, R.; Guillemette, G. J Biol Chem. 279 (2004) 51415-51423.

27. Clement, M.; Martin, S.S.; Beaulieu, M.E.; Chamberland, C.; Lavigne, P.; Leduc, R.; Guillemette, G.; Escher, E. J Biol Chem. 280 (2005) 27121-27129.

28. Balmforth, A.J.; Lee, A.J.; Warburton, P.; Donnelly, D.; Ball S.G. J Biol Chem. 272 (1997) 4245-4251.

29. Doan, T.N.; Ali, M.S.; Bernstein, K.E. J Biol Chem. 276 (2001) 20954-20958.

30. Le, M.T.; Vanderheyden, P.M.; Szaszak, M.; Hunyady, L.; Kersemans, V.; Vauquelin, G. Biochem Pharmacol. 65 (2003) 1329-1338.

31. Hunyady, L.; Ji, H.; Jagadeesh, G.; Zhang, M.; Gaborik, Z.; Mihalik, B.; Catt, K.J. Mol Pharmacol. 54 (1998) 427-434.

32. Monnot, C.; Bihoreau, C.; Conchon, S.; Curnow, K.M.; Corvol, P.; Clauser, E. J Biol Chem. 271 (1996) 1507-1513.

33. Fierens, F.L.; Vanderheyden, P.M.; Gaborik, Z.; Minh, T.L.; Backer, J.P.; Hunyady, L.; Ijzerman, A.; Vauquelin, G. J Renin Angiotensin Aldosterone Syst. 3 (2000) 283-288.

34. Duncia, J.V.; Chiu, A.T.; Carini, D.J.; Gregory, G.B.; Johnson, A.L.; Price, W.A.; Wells, G.J.; Wong, P.C.; Calabrese, J.C.; Timmermans, P.B. J Med Chem. 33 (1990) 1312-1329.

35. Carini, D.J.; Duncia, J.V.; Johnson, A.L.; Chiu, A.T.; Price, W.A.; Wong, P.C.; Timmermans, P.B. J Med Chem. 33 (1990) 1330-1336.

36. Wong, P.C.; Price, W.A. Jr; Chiu, A.T.; Carini, D.J.; Duncia, J.V.; Johnson, A.L.; Wexler, R.R.; Timmermans, P.B. Hypertension. 15 (1990) 823-834.

37. Ballesteros, J. A.; Weinstein, H. W. Methods in Neuroscience 25 (1995) 366-428.

38. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne P. E. Nucl. Acids Res. 28 (2000) 235-242.

39. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. Nucleic Acids Res. 31 (2003) 3784-3788.

40. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. Nucleic Acids Res. 22 (1994) 4673-4680.

41. Heerding, J.N.; Hines, J.; Fluharty, S.J.; Yee, D.K. Biochemistry. 40 (2001) 8369-8377.

42. Macromodel ver. 8.5, Schrodinger Inc., 1999.

43. Amber version 8.0; University of California at San Francisco: San Francisco, CA.

44. Lovejoy, B.; Welch, A. R.; Carr, S.; Luong, C.; Broka, C.; Hendricks, R. T.; Campbell, J. A.; Walker, K. A.; Martin, R.; Van Wart, H.; Browner, M. F. Nat.Struct.Biol. 6 (1999) 217-221.

45. GOLPE 4.5. Multivariate Infometric Analysis Srl., Viale dei Castagni 16, Perugia, Italy (1999).

46. Cruciani G, Goodford P. Copyright Molecular Discovery Ltd. 2001-2003 GREATER graphical interface for GRID version, 1.1.7 GRIB, UPF/IMIM, Barcelona (Spain) http://www.moldiscovery.com.

47. Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S. *Quant. Struct.-Act. Relat*. 12 (1993) 9-20.