Università degli Studi di Palermo

Ministero dell'Università e della Ricerca Scientifica e Tecnologica

***Facoltà di Farmacia***

*PhD in Pharmaceutical Sciences*

*Bando A.A. 2009-2011 (CYCLE XXIII)* SSD:CHIM/08

---

# Molecular Modelling Studies on Molecular Pathways Related to Tumorigenesis in the Discovery of New Lead Compounds

*Dr. Licia Pantano*

**Supervisor**

*Chiar.mo Prof. Anna Maria Almerico*

**PhD Coordinator**

*Chiar.mo Prof. Girolamo Cirrincione*

---

Dipartimento di Scienze e Tecnologie Molecolari e Biomolecolari (STEMBIO)

# INDEX

# Introduction

Drug design is an iterative process which begins when a chemist identifies a compound that displays an interesting biological profile and ends when both the activity profile and the chemical synthesis of the new chemical entity are optimized. One of the basic tenets of medicinal chemistry is that biological activity is dependent on the three-dimensional placement of specific functional groups. Computational chemistry/molecular modeling is the science of representing molecular structures numerically and simulating their behavior with the equations of quantum and classical physics. Computational chemistry programs allow scientists to generate and present molecular data including geometries (bond lengths, bond angles, torsion angles), energies (heat of formation, activation energy, etc.), electronic properties (moments, charges, ionization potential, electron affinity), spectroscopic properties (vibrational modes, chemical shifts) and bulk properties (volumes, surface areas, diffusion, viscosity, etc.). Comparison to experimental data, where available, is also important to guide both laboratory and computational work.

Virtual screening is the application of computational models to select or prioritize compounds for experimental screening. There are two broad categories of virtual screening techniques: ligand-based design and structure-based design.

Ligand-based design methods capitalize on the fact that ligands similar to an active ligand are more likely to be active than random ligands. Ligand-based approaches commonly consider two- or three-dimensional chemistry, shape, electrostatic, and interaction points (e.g., pharmacophore points) to assess similarity. Structure-based design attempts to use the 3D protein structure to predict which ligands will bind to the target. The amount and quality of information required to apply these techniques varies. Ligand similarity approaches require only a single active molecule. Ligand-based Quantitative Structure-Activity Relationship (QSAR) approaches require a number of active molecules spanning a wide range of activity against the target receptor. The quality of the QSAR model depends to a large extent on the quality of the activity data, so that reliable QSAR models are usually built based on carefully acquired binding or inhibition data. Structure-based approaches, of which the best known is docking, require a protein structure or homology model as starting point. Pharmacophore models that include receptor information require an experimental structure of the complex between an active molecule and its target protein.

## Homology Modeling

Homology modeling is an increasingly efficient way to obtain useful information about the proteins of interest. For example, in designing mutants it can be helpful to test hypotheses about a protein function, identifying active and binding sites, identifying, designing and improving ligands for a given binding site, modeling substrate specificity, predicting antigenic epitopes, simulating protein–protein docking, refining models based on NMR constraints and

rationalizing known experimental observations. This computational approach is based on the notion that primary structure of proteins is conseved, through evolution, to a lesser extent than the higher level structures (secondary, tertiary and quaternary). The aim of homology protein structure modeling is to build a three-dimensional (3D) model for a protein of unknown structure (the target) on the basis of sequence similarity to proteins of known structure (the templates).[1–5] Two conditions must be met to build an useful model. First, the similarity between the target sequence and the template structure must be detectable. Second, a substantially correct alignment between the target sequence and the template structures must be calculated. The obtained structure can be structurally refined with different protocols such as energy minimization. Comparative modeling is possible because small changes in the protein sequence usually result in small changes in its 3D structure.[6] Although considerable progress has been made in *ab initio* protein structure prediction,[7] comparative protein structure modeling remains the most accurate prediction method.

**Molecular Dynamics**

Molecular Dynamic simulations describe the time evolution of a molecular system, e.g., a protein, by numerically solving Newton's equations of motion for all atoms in the system.
Molecular dynamics combines energy calculations from force field methodology with the laws of Newtonian mechanics. Such simulations can accurately describe the dynamics of biological relevant systems by using three approximations; (a) the Born-Oppenheimer approximation, where nuclear and electronic motions are decoupled, (b) the approximation that nuclei can be treated as classical particles, and (c) the use of an empirical force field to describe the interaction between particles. The simulation of a protein or nucleic acid requires the explicit spatial coordinates and initial velocities of every atom in such a molecule. While the initial velocities can be obtained from a Maxwell-Boltzmann distribution at given temperature, the spatial coordinates were obtained from structures stored in the Protein Data Bank[8] or from homology models. In molecular dynamics, successive configurations of the system are generated by integrating Newton's laws of motion (Equation 1). The result is a trajectory that specifies how the positions and velocities vary with time.

$$F_i = m_i a_i$$

**Equation 1: Newton's law of motion**

where $F_i$ is the force exerted on particle $i$, $m_i$ is the mass of particle $i$ and $a_i$ is the acceleration of particle $i$.
The force can also be expressed as the gradient of the potential energy (Equation 2).

$$F_i = -\nabla_i V$$

**Equation 2**

Combining these two equations yields Equation 3.

$$-\frac{dV}{dx_i} = m_i \frac{d^2 x_i}{dt^2}$$

**Equation 3**

where *V* is the potential energy of the system.
Newton's equation of motion can then relate the derivative of the potential energy to the changes in position as a function of time.
Numerous numerical algorithms have been developed for integrating the equations of motion:

- Verlet algorithm
- Leap-frog algorithm
- Velocity Verlet
- Beeman's algorithm

All the integration algorithms assume the positions, velocities, and accelerations can be approximated by a Taylor series expansion (Equation 4).

$$x(t+\delta t) = x(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 + \ldots$$
$$v(t+\delta t) = v(t) + a(t)\delta t + \frac{1}{2}b(t)\delta t^2 + \ldots$$
$$a(t+\delta t) = a(t) + b(t)\delta t + \ldots$$

**Equation 4**

Where *x* is the position, *v* is the velocity (the first derivative with respect to time), *a* is the acceleration (the second derivative with respect to time).
The *Verlet algorithm* uses positions and accelerations at time *t* and the positions from time *t-δt* to calculate new positions at time *t+δt*. The Verlet algorithm uses no explicit velocities. The advantages of the Verlet algorithm are simplicity, but the algorithm is of moderate precision.
In the *Leap-frog algorithm*, the velocities are first calculated at time *t+1/2δt*; these are used to calculate the positions, *x*, at time *t+δt*. In this way, the velocities leap over the positions, then the positions *leap* over the velocities. The advantage of this algorithm is that the velocities are explicitly calculated; however, the disadvantage is that they are not calculated at the same time as the positions.
The *velocity-Verlet integrator algorithm* yields positions, velocities, and accelerations at time *t*. There is no compromise on precision.
The *Beeman's algorithm* is closely related to the Verlet one. The advantage of this algorithm is that it provides a more accurate expression for the velocities and a better energy

conservation. The disadvantage is that the more complex expressions make the calculation more expensive.

**Molecular Docking**

Molecular Docking is a method that predicts structure of the intermolecular complex formed between two or more molecules. Frequently it is used to predict binding orientation of drug candidates to their protein targets in order to predict affinity and activity. Prediction of the binding affinity will be useful when compounds are being synthesed whereby it is possible to predict the affinity of the desired compound towards a certain target (say a protein or DNA; with particular interest to stop the function of the enzyme/protein or to block certain reaction).

The orientation of the ligand (small molecule or substrate protein) will be "fitted" to the receptor of interest using either two approaches: matching technique, and simulation processes.

Molecular docking can be divided into two separate sections:

1) **Search algorithm** – The algorithm should create an optimum number of configurations that include the experimentally determined binding modes. So the search space consists of all possible orientations and conformations of the protein paired with the ligand. In practice it is impossible to exhaustively explore the search space because this would involve enumerating all possible distortions of each molecule and all possible rotational and translational orientations of the ligand.
2) **Scoring Function** – Mathematical methods used to predict the strength of the non-covalent interaction, called as binding affinity, between two molecules after they have been docked. Scoring functions have also been developed to predict the strength of other types of intermolecular interactions, for example between two proteins or between protein and DNA or protein and drug. These configurations are evaluated using scoring functions to distinguish the experimental binding modes from all other modes explored through the searching algorithm.

In standard virtual docking studies, ligands are docked into the binding site of a receptor where the receptor is held rigid and the ligand is free to move. However, the assumption of a rigid receptor can give misleading results, since in reality many proteins alter their binding site to conform to the shape and binding mode of the ligand. These changes allow the receptor to alter its binding site so that it more closely conforms to the shape and binding mode of the ligand. This is often referred to as "induced fit".

# 1. Hypoxia and cancer

Tumorigenesis in human is a multistep process that involves the sequential acquisition of a number of genetic, epigenetic, or somatic alterations as a result of increasing genomic instability caused by defects in cell cycle checkpoint controls.[9] These alternations enable cancer cells to acquire characteristics different from normal cells: resistance to growth inhibitory factors, proliferation in the absence of exogenous growth factors, evasion of apoptosis, limitless replication potential *via* the reactivation of telomerase, abnormal angiogenesis, evasion of destruction by the immune system, invasion and metastasis.[10] In addition to the genetic, epigenetic, or somatic changes that occur in cancer, the tumor microenvironment is considered to be a critical factor in malignancy progression and metastasis, and it influences the response to conventional anti-tumor therapies.[11]

As one of the most pervasive microenvironmental stresses and common features of solid tumors, hypoxia has been recognized as playing a key role in several cellular physiological processes, from cell proliferation, to cell survival, angiogenesis, metabolism and tumor progression and metastasis (Fig. 1).

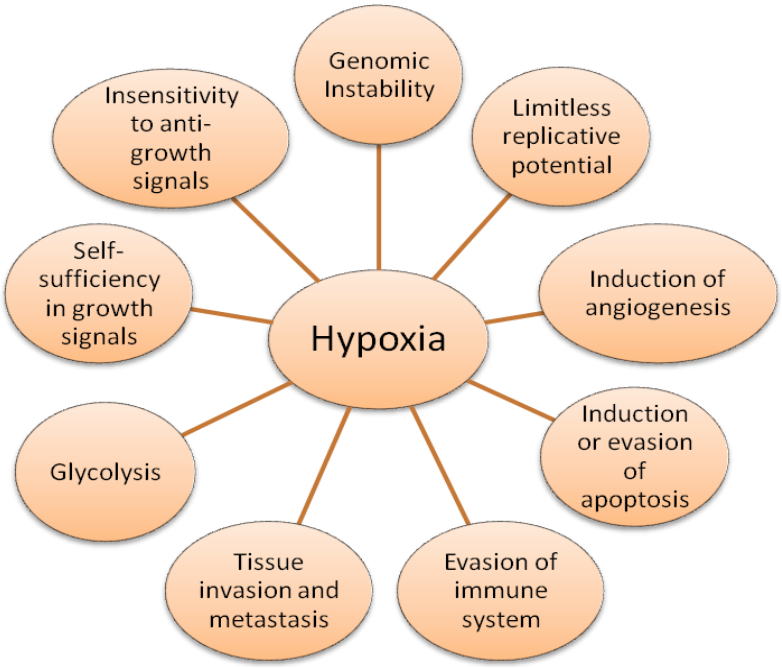



**Fig. 1: The role of hypoxia in the hallmarks of human cancer.**

To survive and grow in this hypoxic microenvironment, tumor cells co-opt adaptive mechanisms to switch to a glycolytic metabolism, promote proliferation, become resistant to apoptosis, obtain unlimited replication potential and genomic instability, evade immune attack, induce angiogenesis, and migrate to less hypoxic areas of the body.

Hypoxia can be divided into acute and chronic hypoxia. The acute variant is usually caused by a temporary disruption to the blood flow and does not last long. On the other hand, chronic hypoxia is durable and can have lasting effects.[12]

Hypoxic cells have been shown to be more resistant to radiotherapy and chemotherapy, and they usually have been associated with increased risk of invasion and metastasis, and a poor clinical prognosis of solid tumors.[13] Therefore, since HIF-1 is a key regulator of the response of cells to oxygen deprivation and plays critical roles in the adaptation of tumor cells to a hypoxic microenvironment, HIF-1 inhibition is an attractive anticancer target. Knowledge of the mechanisms of action of all the actors in the hypoxic pathway is thus becoming a priority in identifying new agents capable of specifically targeting HIF-1.

**Mdm2 and HIF-1α interaction in tumor cells during hypoxia**

The importance of the HIF-1 response pathway in human tumorigenesis is underscored by the finding that HIF-1α is overexpressed in multiple human cancers, because tumor cells, unlike normal cells from the same tissue, are often chronically hypoxic.[14]

In normal unstressed cells, p53 is a very unstable protein with a short half-life, which is present at very low cellular levels owing to continuous degradation largely mediated by Mdm2. In contrast, the p53 protein is stabilized, and its level increases in response to various stresses such as DNA damage and hypoxia.[15,16] Mdm2 an p53 are linked to each other through an autoregulatory negative feedback loop aimed at maintaining low cellular p53 levels in the absence of stress. Mdm2 inhibits p53 activity because it stimulates its degradation in the nucleus and the cytoplasm, blocks its transcriptional activity interfering with the ability of p53 to contact transcriptional coactivators such as p300/CBP, and promotes its nuclear export to the cytoplasm, where p53 is then degraded by cytoplasmic proteasomes.[17]

Evidence for interaction between HIF-1 and p53 network is substantial,[18–21] the precise mechanism by which HIF-1α regulates p53-mediated function remains unknown (Fig. 2).

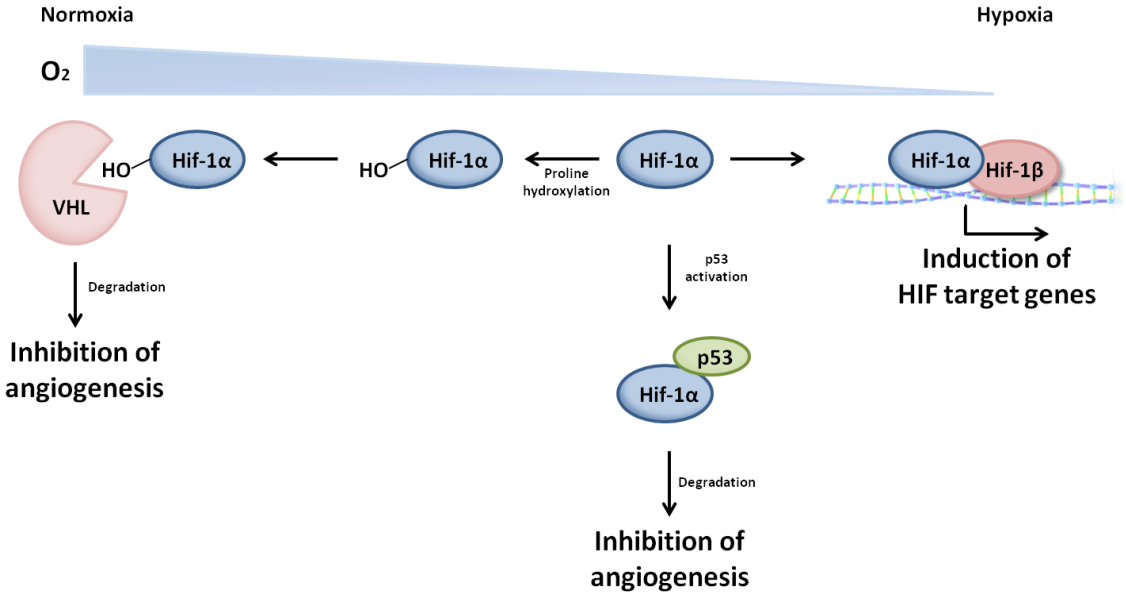



**Fig. 2**

Mdm2 may act as a bridge and mediate the indirect interaction between HIF-1α and p53 in cells. HIF-1α protects p53 from degradation mediated by Mdm2 and can abrogate p53 transcriptional repression by Mdm2. Also, HIF-1α interacts with the wild-type p53 protein but not the tumor-derived p53 mutant form in cells.[18] Since wild-type p53 protein is capable of inducing Mdm2 expression in cells, in contrast, because the tumor-derived p53 mutant is completely inactive in transcriptional activation of endogenous Mdm2, HIF-1α fails to interact with p53 since there is no (or very low levels) Mdm2 in cells expressing mutated p53.

**A3 adenosine receptor induces HIF-1*α* protein accumulation in hypoxia**

Purine nucleosides, such as adenosine, are critical mediators of physiological responses to acute and chronic hypoxia. Adenosine is the final metabolite in the stepwise dephosphorylation of ATP and it is produced and released in response to ischemia and hypoxia in the central nervous system.[22]

There is a strong link between adenosine and hypoxia-related signaling. The expression levels of adenosine and adenosine receptors are regulated in conditions of cellular stress, and signal transduction increases *via* one or more of the adenosine receptors. Hypoxia apparently induces a program that shifts the tissue phenotype toward an increase in extracellular adenosine. In turn, adenosine receptor activation tends to limit the potential damage incurred by hypoxia.

Adenosine modulates a variety of cellular functions through occupancy of four cell surface G-protein–coupled receptors, named A1, A2A, A2B, and A3.[23,24] In particular, adenosine was found to exert its effects on cell proliferation, clone formation ability, UV resistance, and cell death mainly through the A3 subtype,[25–28] which is highly expressed in tumor cells.[29–33] These findings confirm recent data indicating that A3 receptor overexpression may be a good candidate as a tumor cell marker.

Several reports demonstrate that adenosine is able to increase HIF-1α protein expression in response to hypoxia in a dose-dependent and time-dependent manner in human melanoma cells, whereas HIF-1β protein levels are not affected. A3 receptor subtype mediates the observed adenosine effects on HIF-1α regulation in this cell line. The effects of adenosine on HIF-1α protein accumulation are not mediated by A1, A2A, or A2B receptors but through A3 receptors.[34]

So given the ability of A3 adenosine receptor antagonists to block HIF-1α protein expression accumulation in hypoxia, this pathway could be a new approach for the treatment of cancer, based on the cooperation between hypoxic and adenosine signals, that ultimately may lead to the increase in HIF-1–mediated effects in cancer cells.

## 2. HIF-1α

HIF-1 is a heterodimer composed of HIF-1α and HIF-1β subunits.[35] Whereas HIF-1β is constitutively expressed, HIF-1α expression is induced in hypoxic cells with an exponential increase in expression as cells are exposed to $O_2$ concentrations of less than 6%. Two other homologues of the α subunit have been cloned (HIF-2α or EPAS-1 and HIF-3α), but there appears to be little redundancy in the hypoxic response. The three related forms of human HIF-α (HIF-1α, HIF-2α, and HIF-3α) are encoded by a distinct genetic locus. HIF-1α and HIF-2α possess similar domain structures that are regulated in a related manner by oxygen, although each isoform has distinct and separate roles. The role of HIF-3α is not fully understood, although a truncated form of murine HIF-3α, known as inhibitory Per/Arnt/Sim (PAS) domain protein (IPAS), has been found to act as an inhibitor of HIF *via* dimerization with HIF-β. HIF-1α is the best characterized and forms a heterodimer with the HIF-1β subunit, initially identified as the Aryl hydrocarbon Receptor Nuclear Translocator (ARNT). HIF-2α and 3α compete for binding to ARNT. These proteins belong to the basic helix-loop-helix (bHLH)–PER-*A*RNT-*S*IM (PAS) protein family (Fig. 3).

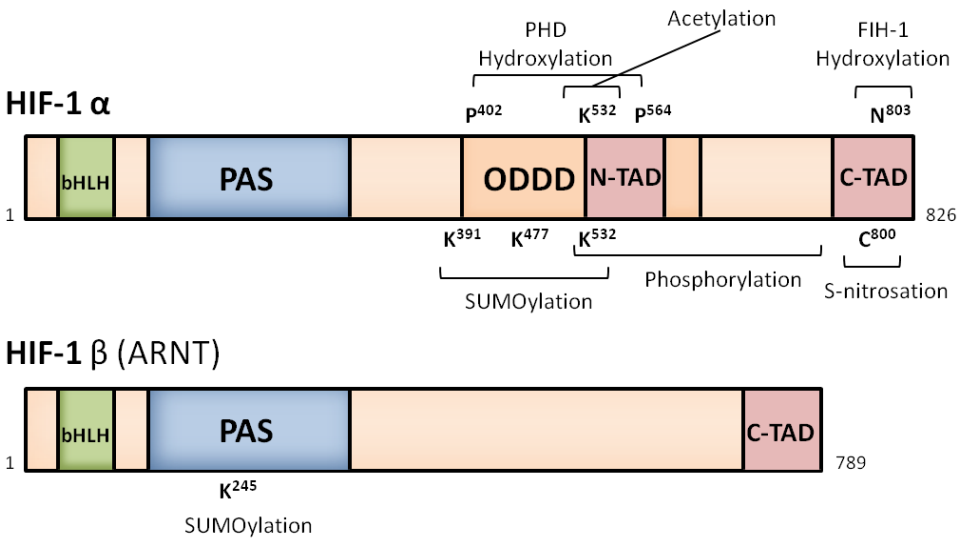



**Fig. 3: Domain structure and structural location of the post-translational modifications of HIF-1α and β.**

The bHLH and PAS motifs are required for dimerization while the downstream basic region affords specific binding to the HRE DNA sequence 5′-ACGTG-3′. The stability and subsequent transactivational function of the α subunit of HIF-1 is regulated by its post-translational modification, in particular hydroxylation and phosphorylation. The Oxygen-Dependent Degradation Domain (ODDD) of HIF-1α regulates its stability through the hydroxylation of proline 402 and 564 by Prolyl Hydroxylase Domain (PHD) proteins and the acetylation of lysine 532 by ARrest Defective-1 protein (ARD1), which favour binding to and subsequent

ubiquitination by the von Hippel-Lindau (VHL) E3 ligase complex. The α subunit contains two transactivation domains, the N- and C-terminal activation domains, respectively N-TAD and C-TAD, while the β subunit contains only one TAD. A number of co-activators including CBP/p300 have been identified to interact with the C-TAD and to enhance transactivation. HIF-1α is phosphorylated, and phosphorylation has been shown to enhance transcriptional activity though possibly through modification of co-activators. The transcriptional activity is negatively regulated by hydroxylation of asparagine 803 by the Factor Inhibiting HIF-1 (FIH-1) in the C-TAD which abrogates binding of the co-activator CBP/p300. S-nitrosation of cysteine 800 increases interaction with CBP/p300 and enhances transactivation, while the transcriptional activity of both subunits appears to be negatively regulated by SUMOylation.[36]

In normoxia, HIF-1α constitutively transcribed and translated, but immediately directed for degradation. This is achieved by the hydroxylation of proline residues (P402 and P564) by the three PHDs, which depend on oxygen, 2-oxoglutarate, $Fe^{+2}$, and ascorbate as substrates and cofactors for their activity. Prolyl hydroxylation permits the binding of the von Hippel-Lindau protein (pVHL), a recognition component of the E3 ligase complex[37–42] together with elongin B, elongin C, cullin-2, and ring-box 1. This complex ubiquinates HIF-1α subunits and targets them for proteosomal degradation.[43] In addition, a conserved asparagine residue undergoes hydroxylation, by Factor Inhibiting HIF (FIH), which blocks activation of HIF target genes, as well as having other functions.

In addition to ubiquitin, there exist a number of related polypeptides which are covalently attached to target proteins and regulate their function.[44] The Small Ubiquitin-like MOdifier (SUMO) is one of these polypeptides, but in contrast to ubiquitin, SUMO does not signal protein destruction but instead may even protect proteins from ubiquitination and influence intracellular localization and protein–protein interactions.[45] A growing number of transcription factors including p53, heat shock transcription factor, c-Myb, GRIP1, Sp3 and AP-2 are being reported to undergo SUMO post-translational modification. Often this modification negatively regulates the transcriptional activity and this has been demonstrated for the β subunit of HIF-1.[46] Under hypoxic conditions, prolyl hydroxylation of HIF-1α is blocked and acetylation is down-regulated, permitting thus HIF-1α protein stabilization. HIF-1α protein accumulates and translocates to the nucleus where is then free to bind with HIF-1β to form the HIF-1 transcription complex. The heterodimer can then bind to hypoxic response elements (HREs) in the promoters of a host of genes and activate expression of these genes (Fig. 4).

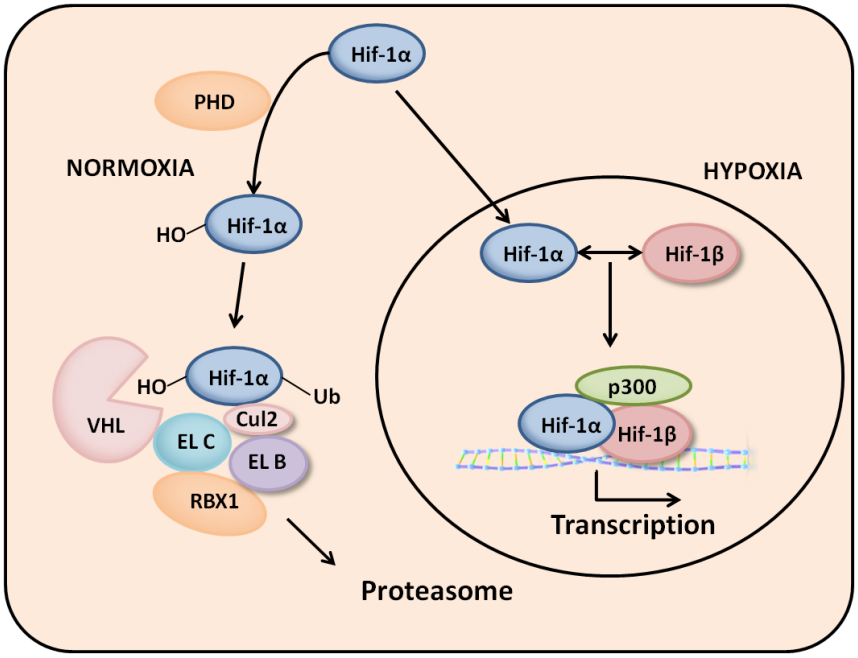



**Fig. 4**

The HIF-1α protein is an excellent example of multiple post-translational modifications as switches in function. Hydroxylation by PHD proteins and acetylation by ARD1 protein increase interaction of HIF-1α with the E3 ligase pVHL complex which earmarks HIF-1α with ubiquitin. Hydroxylation by FIH-1 decreases the transacriptional activity as does modification with SUMO. HIF-1α is phosphorylated by Mitogen Activated Protein Kinase (MAPK) and phosphorylation increases transactivation as does S-nitrosation.

**Regulation of HIF-1α protein stability under hypoxia**

Hypoxia-inducible factor can be activated by physiological or pathological activation of growth factor and cell adhesion pathways. Growth-factor-induced activation of receptor tyrosine kinases (RTKs) leads to HIF-1α stabilization and activation. Activated RTKs interact with p85, the regulatory subunit of Phosphatidyl Inositol 3-Kinase (PI3K), which leads to its activation. Activated PI3K triggers a phosphorylation cascade that results in the phosphorylation/activation of AKT, a serine/ threonine kinase that promotes antiapoptotic and pro-survival responses of a cell.[47] Activation of AKT has been shown to lead to an increase in HIF-1 protein translation by the AKT/FRAP/mTOR pathway.[48,49] Activated RTKs also signal through the MAPK pathway, and phosphorylated p38 and extracellular-signal-regulated kinase 1/2 (ERK1/2) can further phosphorylate and activate HIF-1.[50] Inhibition of ERK activity leads to inhibition of HIF activity without affecting HIF stabilization.[51] In addition to growth factor–mediated RTK activation, the PI3K/AKT pathway is also activated by extracellular matrix (ECM) adhesion mediated by integrins.[52] Integrin ligation causes an activation of the integrin-linked kinase (ILK) leading to increased HIF-1, as

well as increased VEGF production by the PI3K/AKT/FRAP/mTOR pathway.[53] Additionally, activation of PI3K/AKT also leads to an increase in steady-state concentrations of heat shock proteins (HSP) 90 and 70, both of which interact with and stabilize HIF-1.[54] p53 negatively modulates this process by inducing Mdm2, which can ubiquitinate and lead to HIF-1 degradation by the proteasome pathway (Fig. 5).

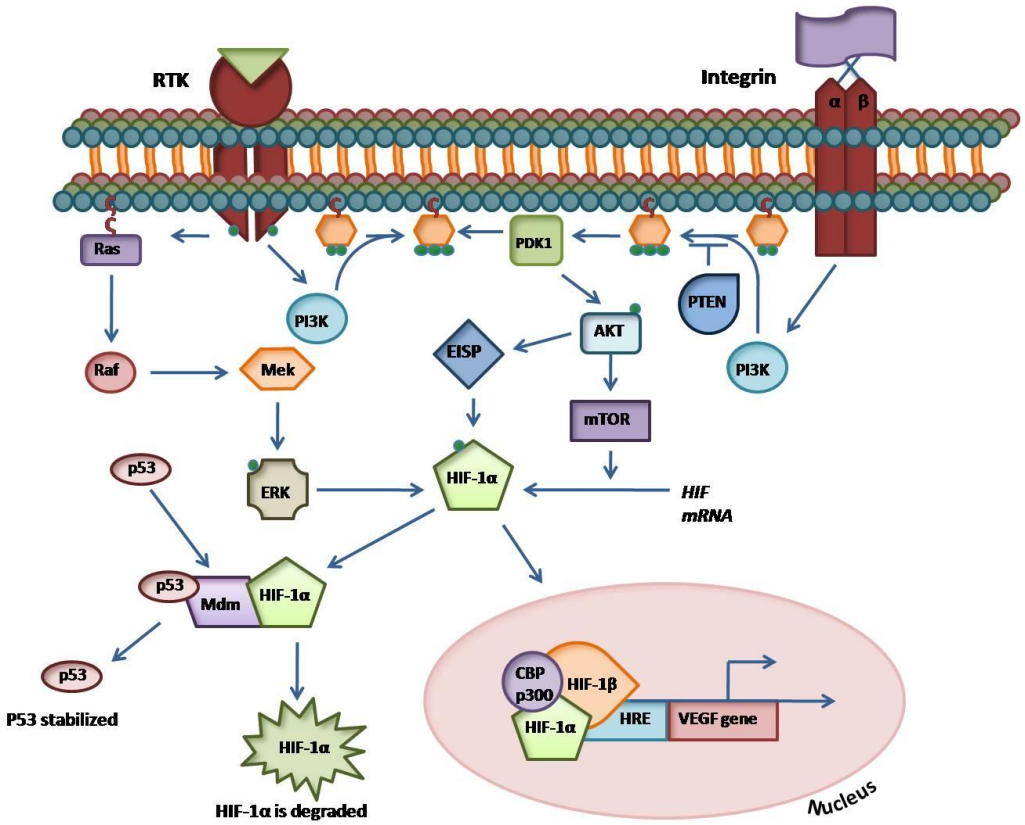



**Fig. 5**

## HIF-1 target genes

Once activated by hypoxia, HIF-1 binds to the consensus HIF-1 DNA binding site (HBS) ACGTG present in the hypoxia-response elements (HREs) of many oxygen-regulated genes. Table 1 shows a compilation of the HIF-1 target genes identified. These genes are involved in oxygen homeostasis at the cellular, local and systemic levels. Erythropoietin activates erythropoiesis to enhance the systemic oxygen transport capacity. Because iron is a limiting factor in haem formation, erythropoiesis is sustained by increased expression of transferrin and transferrin receptor to enhance iron supply to erythroid cells. At the local level, HIF-1 activates VEGF, as well as one of its receptors (Flt-1), which induce angiogenesis leading to an increase in the vascular density and hence a decrease in the diffusion distance for oxygen. Local blood circulation is also controlled by modulation of the vascular tone through the production of NO (nitric oxide synthase), CO (haem oxygenase 1), endothelin-1, adrenomedullin or activation of the $\alpha_{1B}$-adrenergic receptor. At the cellular level, loss of ATP production in mitochondria is compensated by anaerobic glycolysis. Therefore, glucose uptake (glucose

transporters) and glycolysis (glycolytic enzymes) are upregulated by HIF-1. Besides hypoxia, both insulin and insulin-like growth factors, (IGF)-1 and 2, induce HIF-1α expression.[55] In addition to glucose transporters and glycolytic enzymes, HIF-1 activates IGF-2 and the IGF-binding proteins (IGFBPs) 1, 2 and 3 but not 4, 5 and 6.[56,57] Other pleiotropic growth factors and cytokines capable of inducing HIF-1 include epidermal growth factor, fibroblast growth factor-2, interleukin-1β and tumour necrosis factor-α..[58,59]

| **Hypoxia-inducible HIF-1 target gene** |
|---|
| Oxygen transport: erythropoiesis |
| Erythropoietin |
| Transferrin (iron transport) |
| Transferrin receptor (iron uptake) |
| Oxygen transport: angiogenesis and vascular tone |
| Vascular endothelial growth factor (VEGF) |
| Flt-1 (VEGF receptor 1) |
| Plasminogen activator inhibitor-1 |
| Endothelin-1 |
| Inducible nitric oxide synthase (NO production) |
| Haem oxygenase 1 (CO production) |
| Adrenomedullin |
| $\alpha_{1B}$-adrenergic receptor |
| Anaerobic energy: glycolysis and glucose uptake |
| Phosphofructokinase L |
| Aldolase A |
| Glyceraldehyde-3-phosphate dehydrogenase |
| Phosphoglycerate kinase 1 |
| Enolase 1 |
| Lactate dehydrogenase A |
| Glucose transporter-1 |
| Negative feedback regulation of HIF-1 function |
| p35srj (CBP/p300 antagonist) |
| Others |
| Insulin-like growth factor binding protein-1 |
| Retrotransposon VL30 |

**Table 1: Identified HIF-1 target genes.**

## 3. Mdm2 and p53

The protein p53 plays a key role in maintaining the genomic integrity of cells. In response to DNA damage and other types of *stress stimuli*, p53 causes cell-cycle arrest[60] or activates apoptosis.[61,62] In normal cells, p53 is held in check until needed by Mdm2 (the murine double-minute clone 2, more appropriately termed human double-minute clone 2, or Hdm2) [63]. Harmful mutations of p53 are common mechanisms for the loss of p53 wild-type activity in tumor cells.[64] But another important mechanism is overexpression of Mdm2, which leads to constitutive inhibition of p53; this is commonly seen in cancerous cells containing wild-type (WT) p53.[65,66] Because of its importance in cancer development, the p53–Mdm2 complex is a really interesting target for anticancer drug design. It has been shown that a p53 homologue is sufficient to induce p53-dependent cell death in cells overexpressing Mdm2,[67] and that a peptide as short as six residues could bind to Mdm2 in the same manner.[68] Medicinal chemistry modifications to the same 6-residue peptide dramatically increased its inhibitory activity.[69] Small p53 mimics would be expected to disrupt the p53-Mdm2 complex, consequently liberating p53 to initiate cell-cycle arrest or apoptosis.

The MDM2 gene was originally identified on double-minute chromosomes of spontaneously transformed mouse 3T3 fibroblasts (mouse double minute),[70] and the Mdm2 protein was later found to be physically associated with p53.[71]

Mdm2 is a member of the really interesting new gene 1 (RING) domain family of E3 ubiquitin ligases. The full-length transcript of the MDM2 gene encodes a protein of 491 amino acids with a predicted molecular weight of 56kDa. This protein contains several conserved structural domains including an N-terminal p53 interaction domain (Fig. 6). The Mdm2 protein also contains a central acidic domain (residues 230-300). The phosphorylation of residues within this domain appears to be important for regulation of Mdm2 function. In addition, this region contains nuclear export and import signals that are essential for proper nuclear-cytoplasmic trafficking of Mdm2. Another conserved domain within the Mdm2 protein is a Zinc finger domain. Mdm2 also contains a C-terminal RING domain (amino acid residues 430-480), which contains a consensus sequence that coordinates two molecules of zinc.

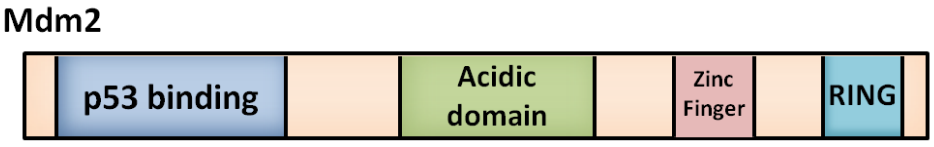



**Fig. 6**

Like other RING domain proteins, Mdm2 functions as an adaptor protein, as an E3 ubiquitin ligase responsible for the ubiquitination and degradation of p53.[72–74] Ubiquitination of proteins occurs through a complex series of steps that involve E1, E2, and E3 proteins.[75,76] The E1 enzyme binds ubiquitin, a 76-amino acid protein, activating ubiquitin for further

processing. The E2 conjugating enzyme accepts the activated ubiquitin from E1 and transfers it to the E3 enzyme, a ligase that covalently binds the ubiquitin to the substrate. Mdm2 functions as the E3 ligase to ubiquitinate p53 at several lysine residues.[77,78] It also has the ability to ubiquitinate itself.[79,80] Mdm2 together with the p300, transcriptional co-activator protein, mediates the ubiquitination and proteasome-dependent degradation of the p53 tumor suppressor protein and other growth regulatory proteins.[81–83] In addition to mediating degradation of p53, Mdm2 blocks the interaction of p53 with the transcriptional apparatus [71], mediates translocation of p53 to the cytoplasm,[84] thereby removing it from its site of action, and recruits the histone deacetylase HDAC1 to deacetylate key residues of p53 thus making them available for ubiquitination.[85]

Mdm2 and p53 are linked to each other through an autoregulatory negative feedback loop aimed at maintaining low cellular p53 levels in the absence of stress. In normal cells, p53 activity is kept low by Mdm2, p53 stimulates the expression of Mdm2; Mdm2, in turn, inhibits p53 activity because it stimulates its degradation in the nucleus and the cytoplasm, blocks its transcriptional activity, and promotes its nuclear export (Fig. 7).

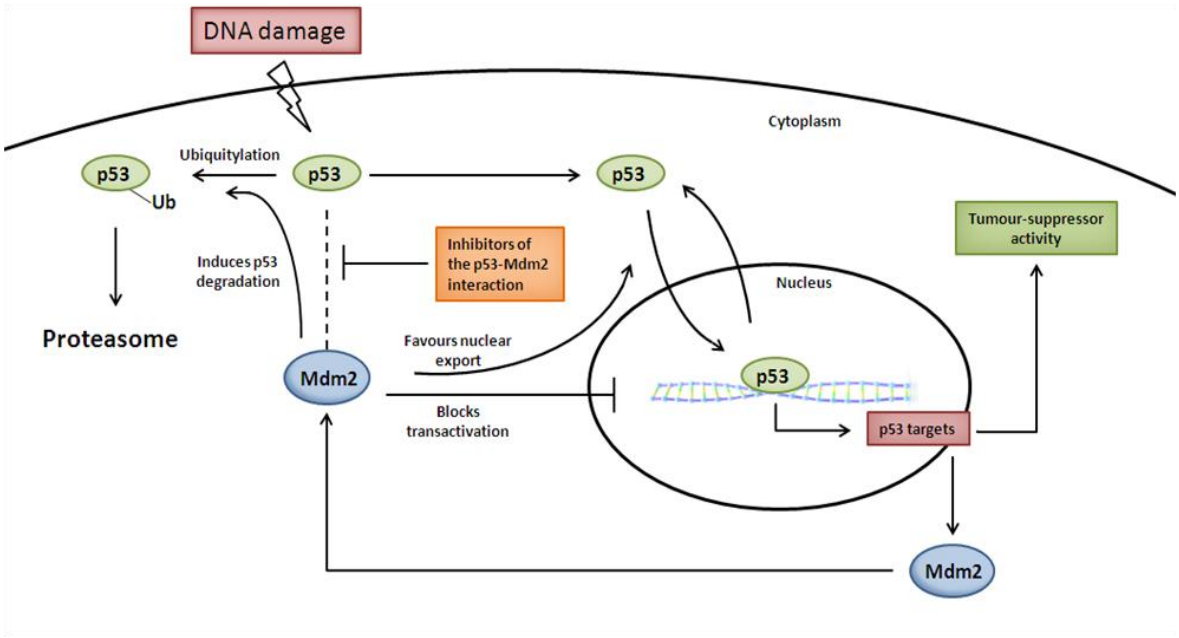



**Fig. 7**

In response to DNA damage, p53 is activated by disrupting Mdm2 association and stabilized against Mdm2-dependent degradation. p53 activation and stabilization likely are achieved by post-translational modifications; known modifications to p53 include phosphorylation, acetylation, and ubiquitination.

Crystallographic data showed that the N-terminal domain of Mdm2 forms a deep hydrophobic cleft into which the transactivation domain of p53 binds (amino acids 19-26), thereby concealing itself from interaction with the transcriptional machinery.[86] This suggest that amino acids 16–24 of Mdm2 can form a “flexible lid” that folds over and stabilizes the Mdm2 structure.[87] The key to the interface is a triad of hydrophobic and aromatic amino

acids of p53 — Phe19, Trp23, and Leu26 — which inserts deep into the Mdm2 cleft. The interface relies extensively on van der Waals contacts and the steric complementarity between the Mdm2 cleft and the hydrophobic face of the p53 helix as these interactions are augmented by only two intermolecular hydrogen bonds (Fig. 8).

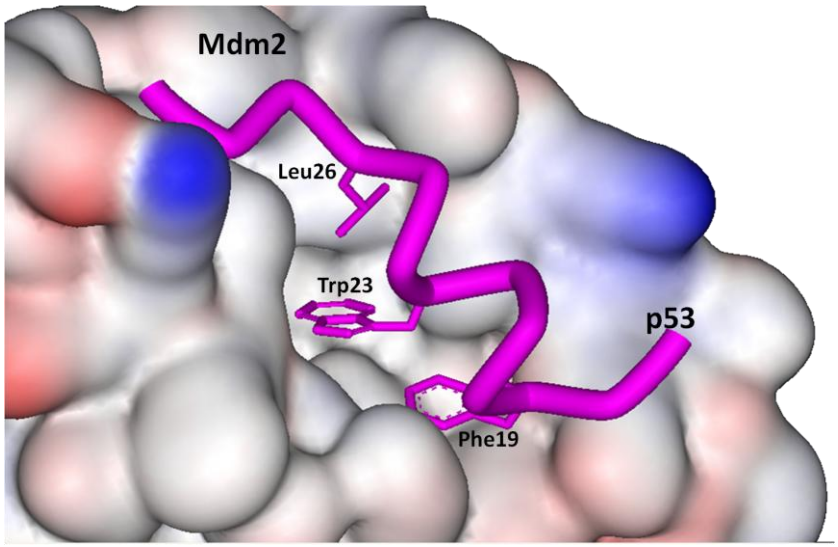



**Fig. 8**

The Mdm2 cleft is lined with 11 hydrophobic and aromatic amino acids that make multiple van der Waals contacts to p53 (Leu54, Leu57, Ile61, and Met62 from the α2 helix, Tyr67, Val75, Phe91, and Val93 from the middle β sheet, and Ile99, and Ile103 from the α2′ helix). Because the hydrophobic p53-Mdm2 interaction is structurally and biologically well understood, the design of small lipophilic molecules that disrupt or prevent the interaction is currently a therapeutic strategy. Only the Mdm2 partner has structurally well-defined binding sites. This implies that inhibitors should mimic p53 rather than Mdm2. Another reason is that only the p53 interface is composed of a single short contiguous stretch of amino acids. All these features favor the possibility that a small inhibitory molecule might work. An increasing number of small-molecule p53-Mdm2 binding inhibitors have been discovered and published in recent years, but only few compounds have acceptable cellular potency and selectivity for their molecular target and might represent viable leads for development of therapeutic agents.[88] In order to design an effective p53 mimic as inhibitor of human Mdm2, it is important to understand the p53-Mdm2 interaction at the atomic level. Recently, in different molecular dynamics (MD) studies, the p53-Mdm2 system was investigated to explore the binding interface, and the effect of mutating key residues in the human p53-Mdm2 complex. The first calculation was a 400 ps molecular dynamics simulations by Massova and Kollman.[89] Other studies were published from 2005 to date, and they were always referred to Mdm2 and the endogen ligand.[90–93] Carlson et al.[90] have applied a 2ns MD simulation to examine the binding interface in the human p53-Mdm2 complex in order to design a potent p53 mimic. This study suggested that an additional hydrofobic pocket interior of Mdm2 should possibly be used to design new inhibitors.

Recently Verma et al.[94] have applied MD simulations to investigate the binding of p53 peptide and nutlin to Mdm2 and MdmX. Simulations reveal that p53 has a higher affinity for Mdm2 than MdmX, driven by stronger electrostatic interactions. The differences are more pronounced for nutlin because it is a small molecule whose binding is driven by short range van der Waals interactions and lacks the long range electrostatics that mediate interactions with p53, supporting findings of how the Mdm2 surface (and MdmX) modulates and is modulated by ligands. In work carried out on the X-ray structure of MdmX bound to a single-domain antibody by Fersht et al.[95], the authors studied structural changes to a common conformation on removal of the ligand. The binding pocket converged to a common conformation after removal of the ligands, indicating that the differences are due to induced fit. However, the residues that comprise the Mdm2 lid are not conserved in MdmX; and also crystal structure of nutlin complexed to Mdm2 (1RV1) used for the simulation does not contain information concerning the lid.

Based on available structural data and computational studies, we can classify the conformational states of Mdm2 into the following three broad groups: open state, closed state and apo state.[96] These broad changes in its conformation could be an important starting point for designing new inhibitors.

# 4. Adenosine receptor A3

The adenosine class of heterotrimeric guanine nucleotide-binding protein (G-protein) coupled receptors (GPCRs) mediates the important role of extracellular adenosine in many physiological processes. Adenosine is an ubiquitous neuromodulator that acts by stimulating four cell surface receptors (A1, A2A, A2B, A3), all being part of the huge family of the GPCRs. These receptors are widely distributed throughout the tissues. A2A and A2B receptors are coupled to adenylate cyclase activity, and their stimulation increases the intracellular cyclic adenosine monophosphate (cAMP) concentration, while A1 and A3 receptor stimulation decreases cAMP concentration and raises intracellular $Ca^{2+}$ levels by a pathway involving phospholipase C (PLC) activation (Fig. 9).[97,98]

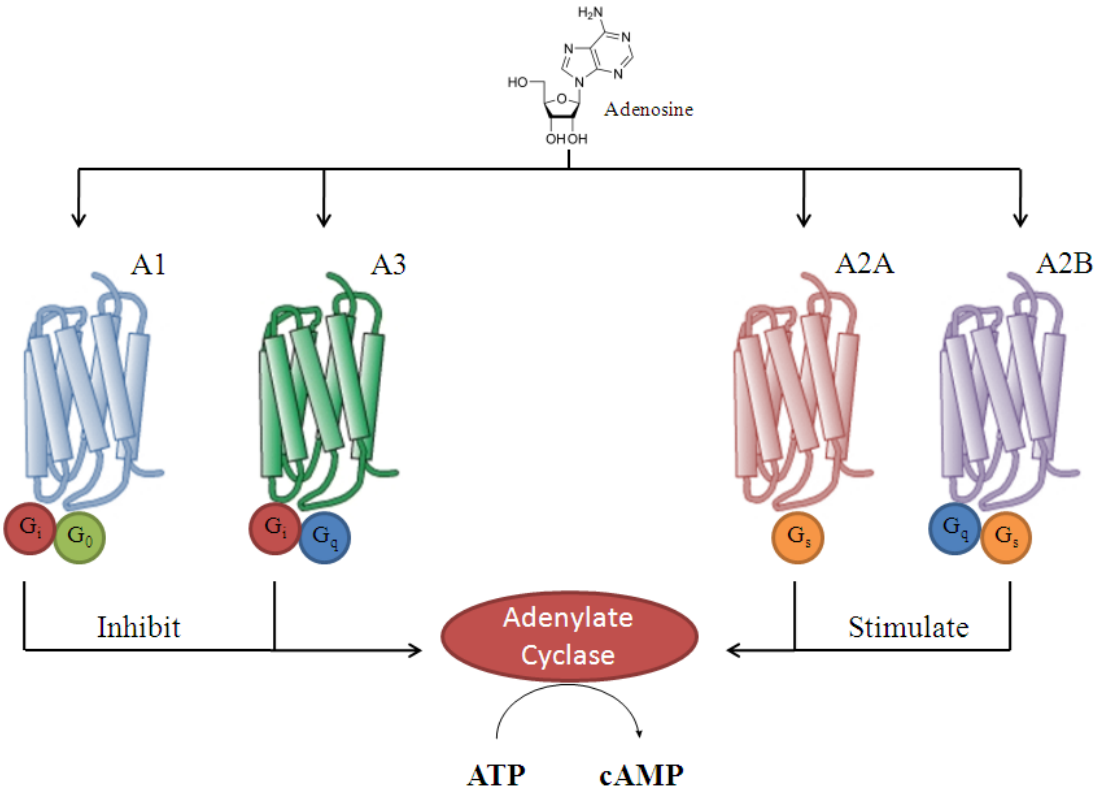



**Fig. 9: Adenosine class of heterotrimeric guanine nucleotide-binding protein coupled receptors.**

The A3R is involved in the control of the cell cycle and inhibition of tumour growth both *in vitro* and *in vivo*.[99] In fact adenosine A3 receptors have been demonstrated to be more highly expressed in tumours than in healthy cells, suggesting a role for A3R as a tumour marker.[33] In accordance with this notion, A3R antagonists are potential therapeutic agent and provide an opportunity for the generation of novel compounds that can be used as antitumor drugs.

All adenosine receptors have in common a core domain consisting of seven transmembrane helices (TM1 to TM7), with each TM composed of 20–27 amino acids, connected by three intracellular (IL1 to IL3) and three extracellular loops (EL1 to EL3). The N-terminus is located at the extracellular side of the cell and often contains one or more glycosylation sites. The C-terminus is located intracellularly and contains phosphorylation and palmitoylation sites, which are involved in regulation of receptor desensitization and internalization.[100] All adenosine receptors, with the exception of the A2AR, contain a palmitoylation site near the C-terminus. The A2AR is the only subtype with an extraordinary long C-terminus.[101] All the adenosine receptors are glycosylated on the second extracellular loop, although glycosylation does not appear to influence ligand binding. The third intracellular loop and/or the C-terminus are involved in coupling the adenosine receptors to G-proteins. The A1, A2B, and A3 receptors are very similar in regard to the number of amino acids composing their primary structure, the human homologs consist of 326, 328, and 318 amino acid residues, respectively. Conversely, the human A2A is composed by 409 amino acids. The human A1R and human A3R display 46.5% overall sequence identity at the amino acid level, while the human A2AR and human A2BR are 46.6% identical (Table 2).

| | **A2A** | **A2B** | **A3** |
|---|---|---|---|
| **A1** | 38.3 | 44.0 | 46.5 |
| **A2A** | | 46.6 | 31.0 |
| **A2B** | | | 35.7 |

**Table 2: Percentage of sequence identity.**

The A3 adenosine receptor has 318 amino acids and contains 7 TM helices connected by three intracellular and three extracellular loops (Fig. 10-11).

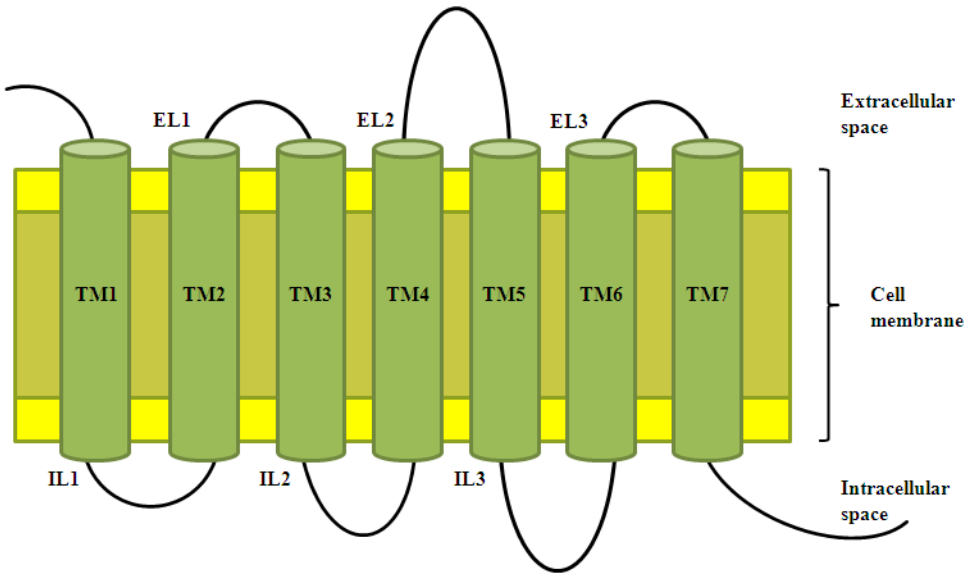



**Fig. 10: Secondary structure of A3R.**

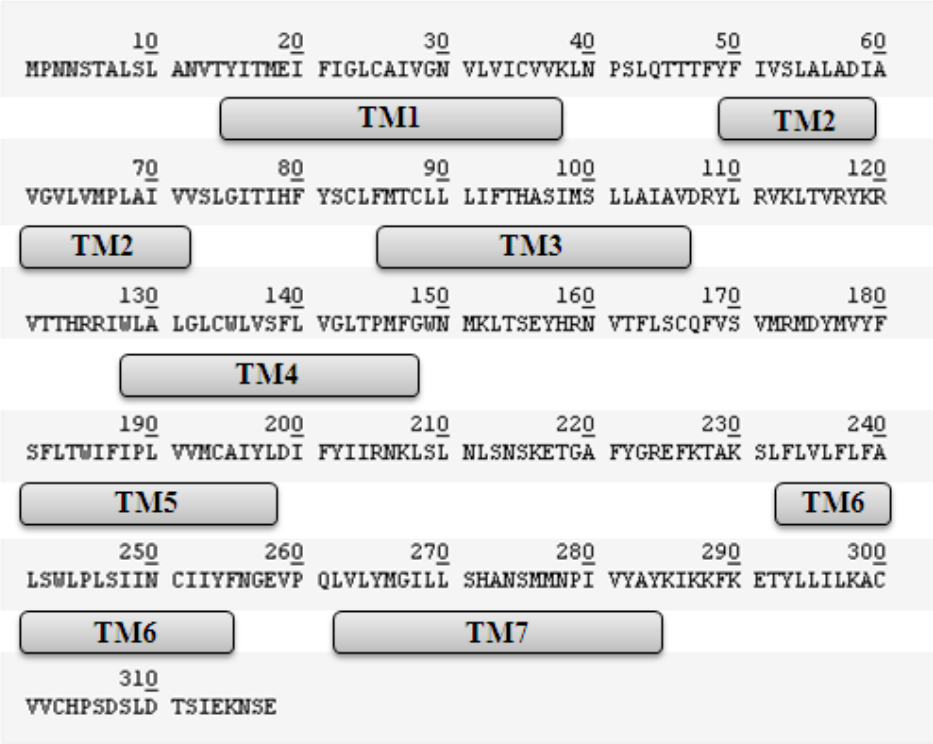



**Fig. 11: Schematic representation of the human A3 adenosine receptor.**

For many years, rhodopsin had represented the only structural information available for GPCRs. The first resolved structure of rhodopsin was published in 2000 by Palczewski et al.[102] and it had been broadly used as template.[103] In 2008 the crystal structure of the human A2A adenosine receptor in complex with a selective antagonist ZM241385 (PDB ID: 3EML) has been determined (Fig. 12).[104] Crystallographic model reveals features different from previously reported GPCR structures (Table 3).

| Protein or complex | PDB code | Resolution | Species |
|---|---|---|---|
| Rhodopsin | 1F88, 1GZM, 1HZX, 1L9H, **1U19**, 2G87, 2HPY, 2I35, 2I36, 2I37, 2J4Y, 2PED, 2X72, 3C9L, 3CM9, 3CAP, 3DQB, 3OAX, 3PQR, 3PXO | 2.2-4.15 Å | bovine |
| β1-Adrenergic receptor | **2VT4**, 2Y00, 2Y01, 2Y02, 2Y03, 2Y04 | 2.5-3.05 Å | turkey |
| β2-Adrenergic receptor | 2R4H, 2R4S, **2RH1**, 3D4S, 3KJ6, 3NY8, 3NY9, 3NYA, 3PDS | 2.4-3.5 Å | human |
| Adenosine A2A receptor | **3EML**, 3QAK | 2.6-2.7 Å | human |

**Table 3: Reported GPCR structures.**

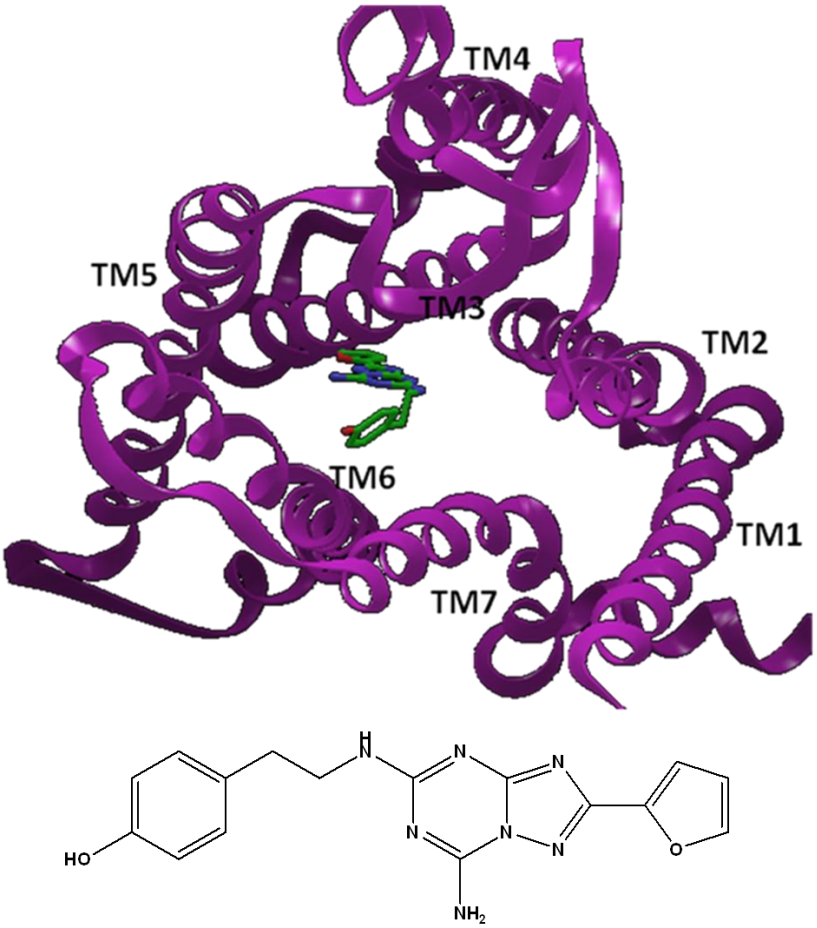



**Fig. 12: Crystal structure of the human A2A adenosine receptor in complex with ZM241385.**

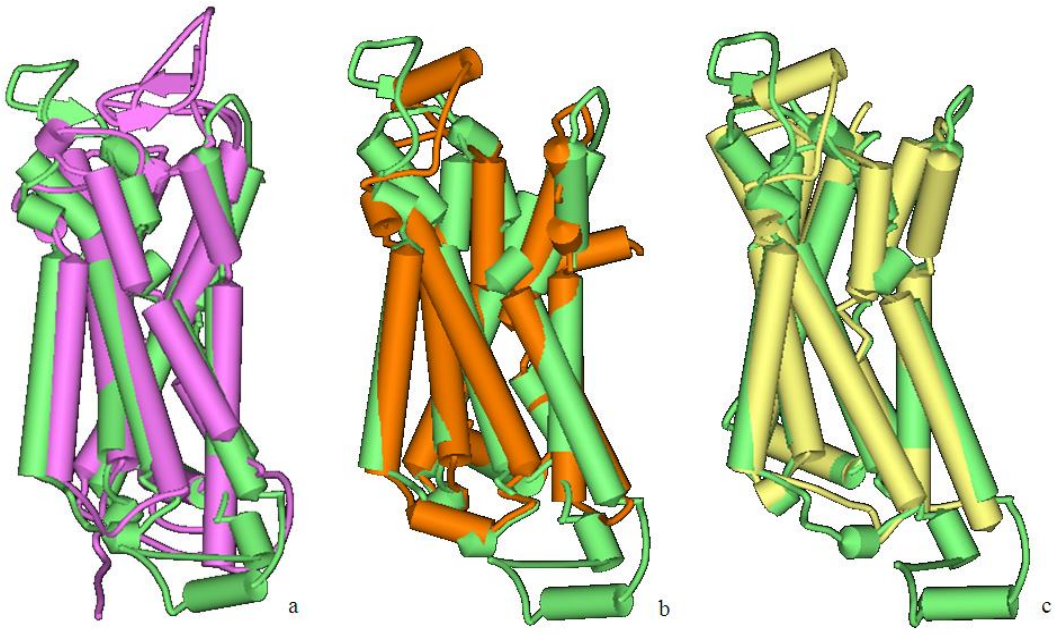



**Fig. 13: Superimposition of A2A adenosine receptor (green) with bovine rhodopsin (purple) (a), β1AR (orange) (b), β2AR (yellow) (c).**

The organization of the extracellular loops is markedly different from β1AR, β2AR and bovine rhodopsin (Fig.13).[105–107]

Furthermore ZM241385 binds A2A in conformation perpendicular to the plane of the membrane and co-linear with transmembrane helix VII, interacting with both EL2 and EL3. Finally, the binding pocket of the A2A adenosine receptor results closer to helices VI and VII and only limited interactions with helices III and V are allowed.

# 5. Aim of the work

Tumor hypoxia can be found in almost every solid tumor and it is now widely recognized as a cause of treatment failure for a wide variety of adult malignancies.[108,109] HIF-1 inhibition may represent a global strategy for targeting the hypoxic tumor microenvironment and there is an extensive effort involved in identifying new more potent and specific HIF-1 inhibitors. However, HIFs-independent pathways may bypass or overcome HIFs inhibition. Therefore, HIF-1 inhibitors may have to be combined with other targeted agents or conventional therapies to integrate hypoxia-targeting methods to get more reasonable results. Focusing research attention on these questions would thus not only be very beneficial for understanding the multifaceted roles of hypoxia on the hallmarks of human cancers but also facilitate the rational design of combination therapies to target hypoxia for cancer treatment.

Dysregulation of HIF-1 appears to play a central role for cancer therapy. Selection of the most appropriate point of therapeutic intervention to modulate HIF activity is also an important factor in pharmaceutical development. In this respect, selective inhibition of the HRE appears to be an attractive target. Our interest is focused on designing compounds that can form stable complexes with DNA and therefore we carried out docking studies on molecules which show a planar moiety, with the aim of performing a fast *in silico* screening of new potential DNA-interactive drugs. Another strategy to achieve inhibition of the hypoxic response in tumours is to target the binding interaction between the CH1 domain of p300 and the C-TAD of the HIF-1α subunit. The interaction of HIF-1α C-TAD and CBP/p300 CH1 domain controls expression of over seventy hypoxia-inducible genes[110] and inhibition of this complex should therefore downregulate multiple genes in a pathway-dependent manner.

Finally, given the important role of Mdm2 overexpression, amplification, or activation in the development of many tumors, the ability to inactivate Mdm2 function in such tumors would provide a potentially significant approach for therapy.

Also an interesting observation is the higher level of adenosine in hypoxia-related signaling. Adenosine is able to increase HIF-1α protein expression in response to hypoxia in a dose- and time-dependent manner.

In particular adenosine modulates a variety of cellular functions through occupancy of A3 receptor subtype, highly expressed in tumor cells. Therefore modulating A3 receptor overexpression may be a good target for the discovery of new candidates as tumor cell marker.

# 6. Results and discussion

## 6.1 HIF-1α: homology modeling and docking studies

The amino acid sequence (Q16665) of HIF-1α protein was obtained from SwissProt Protein Database.[111,112] The complete protein sequence (826 amino acids) of HIF-1α was screened against the BLAST Database in order to identify a template structure appropriate for modeling. The accuracy of comparative modeling is highly dependent on the sequence identity between the target sequence of interest and the template sequence. High accuracy comparative modeling can be achieved when the target and template proteins have sequence identity of more than 50%, while the accuracy drops when the identity of target and template sequences is less than 30%.[113]

The HIF-1α sequence have low identity percentage for all residues but ~73% identity in PAS domains. Therefore, automated homology modeling is likely to result in numerous errors. The major source of errors is from sequence misalignment,[114–116] which can be expected in areas of low sequence identity. Due to template dependent limitations of homology modeling, another computational biology approach, known as *de novo* protein structure prediction, was undertaken. *Ab initio* or *de novo* protein modeling works on the principle that all the information for a protein structure lies in its amino acid sequence. This method builds a 3D structure based on physical principles rather than on previously solved structures. Several online servers, grid services and offline standalone software applications have been developed for *de novo* protein modeling.

Amongst them, I-TASSER is one of the most widely used online servers for protein structure and function predictions. It works by using a combination of *ab initio* folding and threading methods.

Models are built based on multiple-threading alignments by LOMETS[117] and iterative I-TASSER simulations. Once the models were generated, they were subjected to structural assessment and validation using PROCHECK, DFIRE2 and the C-Score values from the I-TASSER. Ramachandran plots were generated by PROCHECK.

Additionally, the stereochemical qualities were assessed for each predicted model. As a result, there were 5 predicted HIF-1α models. Software I-TASSER shows a C-score as a parameter that indicates the quality of the resulting protein structure (Table 4).

| Model | C-Score | TM-Score | RMSD |
|---|---|---|---|
| 1 | -0.43 | 0.9051 | 1.67 |
| 2 | -2.46 | 0.7347 | 3.29 |
| 3 | -2.54 | 0.7311 | 3.89 |
| 4 | -2.58 | 0.6857 | 4.84 |
| 5 | -2.85 | 0.6135 | 5.05 |

**Table 4: C-score value in each model.**

A scoring function (C-score), based on the relative clustering structural density and the consensus significance score of multiple threading templates, is introduced to estimate the accuracy of the I-TASSER predictions. A large-scale benchmark test demonstrates a strong correlation between the C-score and the TM-score. However the correlation of RMSD with the C-score is not as strong as that of the TM-score.

In addition, the formation of the topography of the five models can be used as a comparison, where the model is the best result in the topography better than the other model, it is seen from the shape folding produced in each model (Fig. 14).

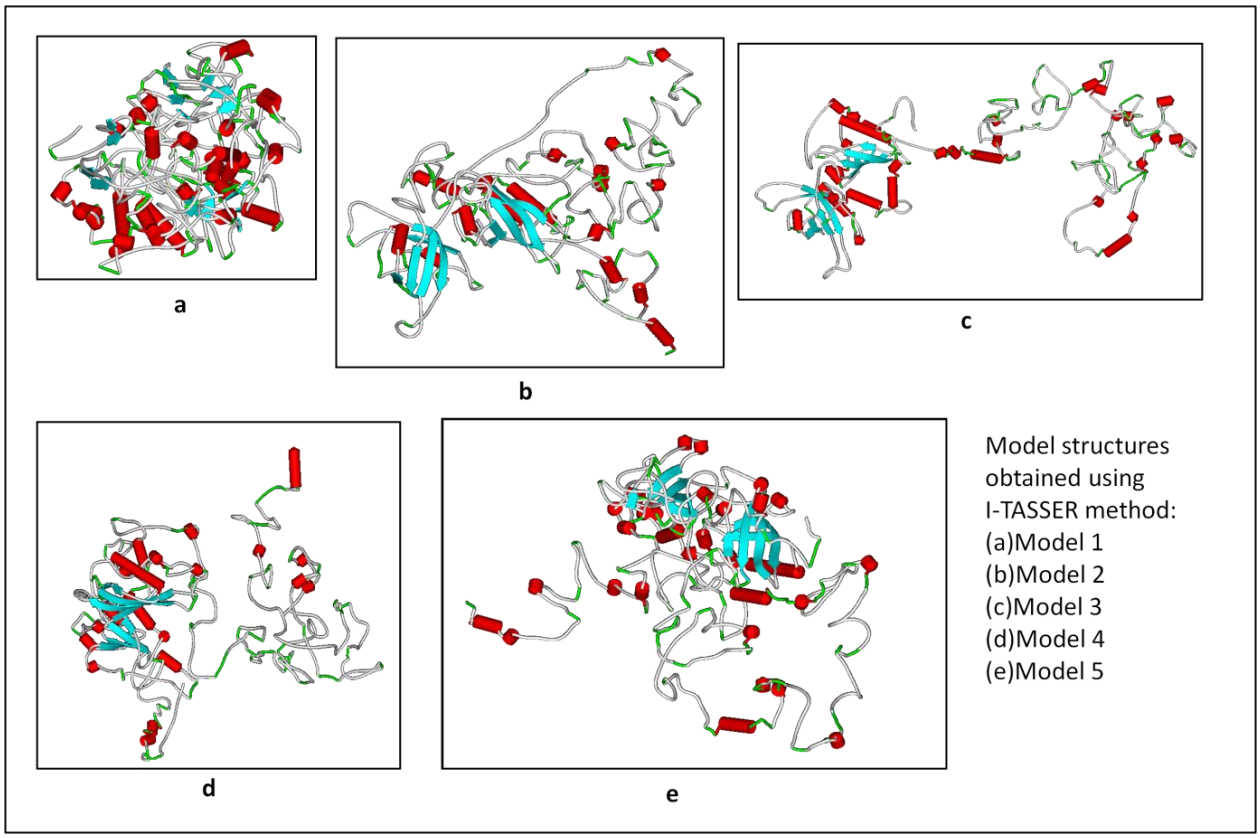



**Fig. 14**

On the basis of the obtained results, Model 2 structure gave a good C-score and the appearance of protein folding in tertiary structure showed typical PAS domains. Once generated, it was subjected to structural assessment and validation. Ramachandran plot was generated by RAMPAGE.[118] The plot showed that Model 2 contained 0.6% of residues in the disallowed region, 11.7% of residues in the generously allowed region and more than 87% of residues were in the most favoured regions (Fig. 15).

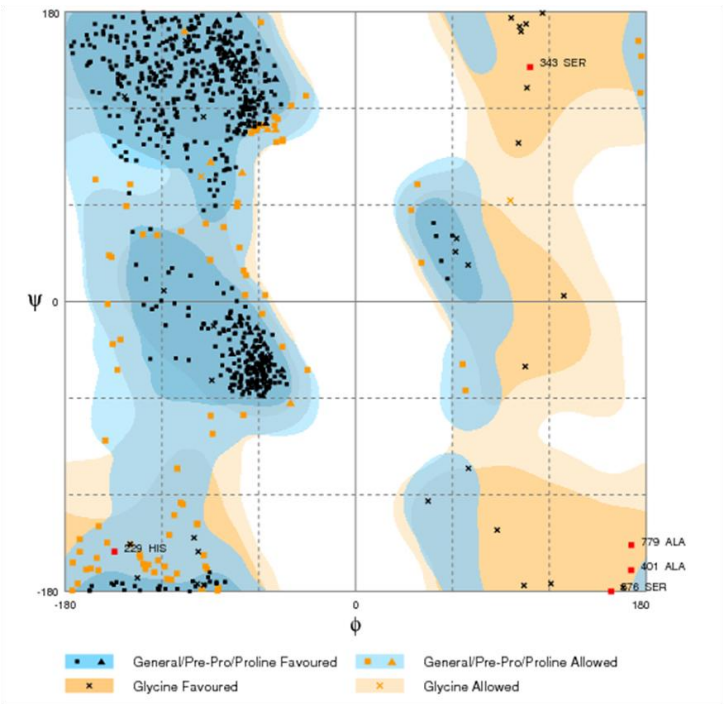
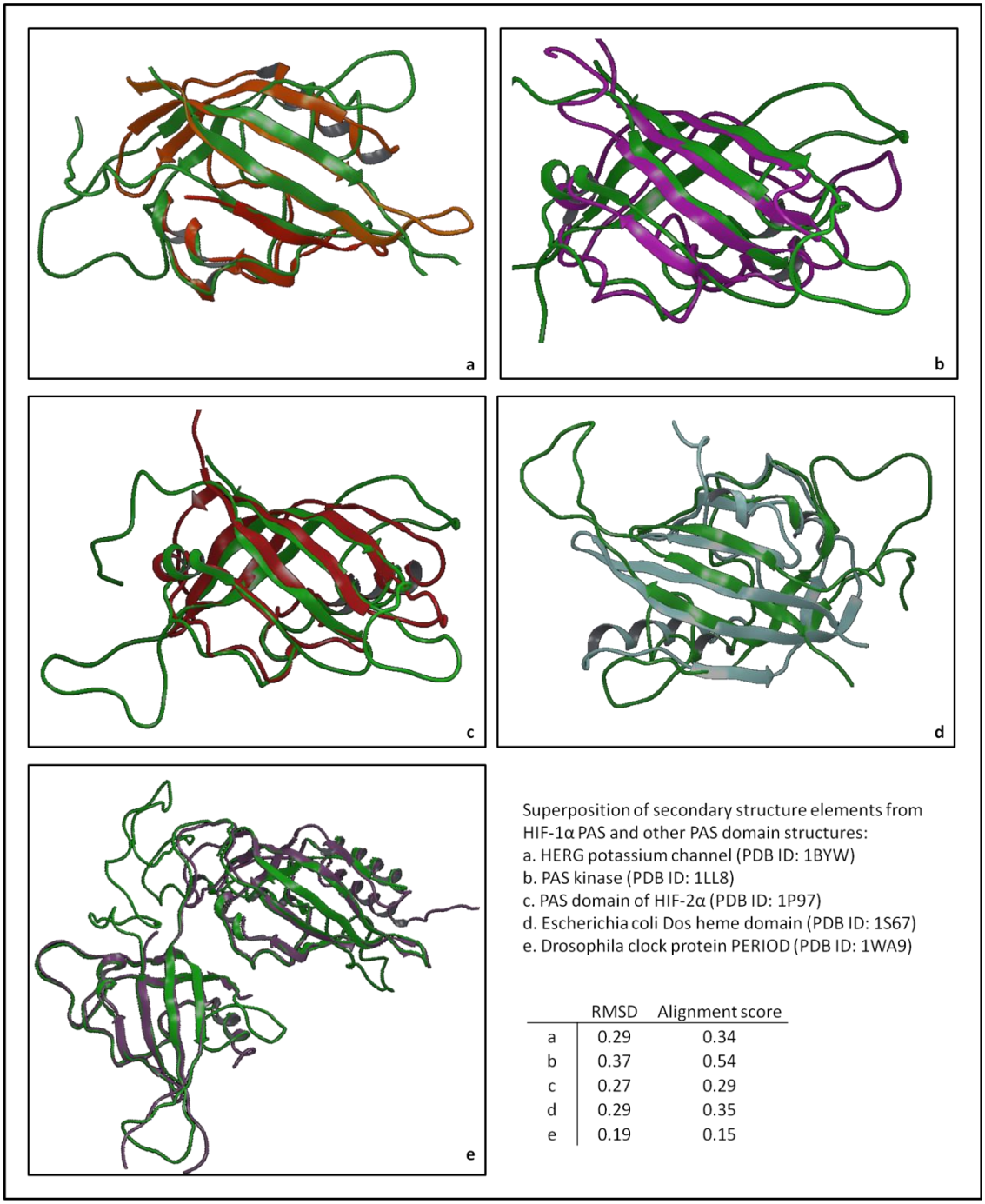



**Fig. 15: Ramachandran Plot**

From the calculated structure it is clear that model has standard PAS domains fold, including a five-stranded antiparallel β-sheet flanked by several α-helices that lie on one face of the protein. The representative structure of this ensemble is very similar to those of other PAS domains as evidenced by the low RMSD obtained by optimized superposition of secondary structure elements from HIF-1α PAS and other PAS domain structures (Fig. 16).




| | RMSD | Alignment score |
|---|---|---|
| a | 0.29 | 0.34 |
| b | 0.37 | 0.54 |
| c | 0.27 | 0.29 |
| d | 0.29 | 0.35 |
| e | 0.19 | 0.15 |

**Fig. 16: Superposition of secondary structure elements from HIF-1α PAS and other PAS domain structures.**

Some regions of protein (PAS domain) have been modeled with high accuracy, while others are more difficult to model especially when the sequence identity with templates is below 40%.

When stable, HIF-1α translocates to the nucleus, dimerises with HIF-1β and binds to hypoxia response elements in the regulatory regions of target genes. Dimerisation is an absolute prerequisite for DNA binding and is mediated by the bHLH and PAS domains of each subunit,

with the basic regions contiguous with the HLH motifs of both partners contacting the DNA (Fig. 17).[119]

In fact it has been shown that while the isolated bHLH domains of ARNT, AHR and HIF-1α can form heterodimeric complexes and bind E-box-like enhancer elements *in vitro*; the inclusion of the adjacent PAS domains enhances the affinity and specificity of the protein/DNA interaction.[120,121]

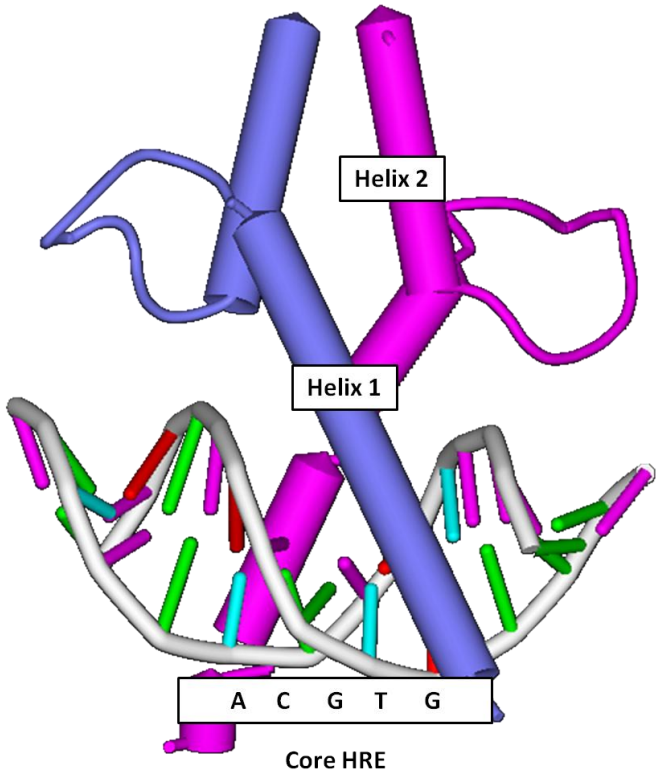



**Fig. 17: Complex between HIF-1 and its consensus DNA sequence (PDB ID: 1D7G).**

A model for the complex between HIF-1 and its consensus DNA sequence (PDB ID: 1D7G) was downloaded from Binding DataBase,[122] a public available database. This model, shown in Fig 17, predicted a pattern of interactions between amino acids and DNA bases which reflects for ARNT what is experimentally observed among different X-ray structures of other bHLH transcription factors possessing the H (His), E (Glu), R (Arg) triad, as ARNT does.[123]

This structure was used to screen possible inhibitors against HRE, through virtual screening of NCI (National Cancer Institute) Database. HIF-1 was removed from the complex and the DNA sequence was utilized for subsequent docking experiments with Glide.[124] Structure-based virtual screening is most commonly implemented as the prediction of binding modes and binding affinities of each compound in the dataset, by means of high-throughput docking to an X-ray structure or model of the target receptor.[125,126] Thus, from the entire NCI Database compounds which were docked, the best 10 docking solutions were selected (Fig. 18, Table 5). Complexes were further analyzed for their interaction.

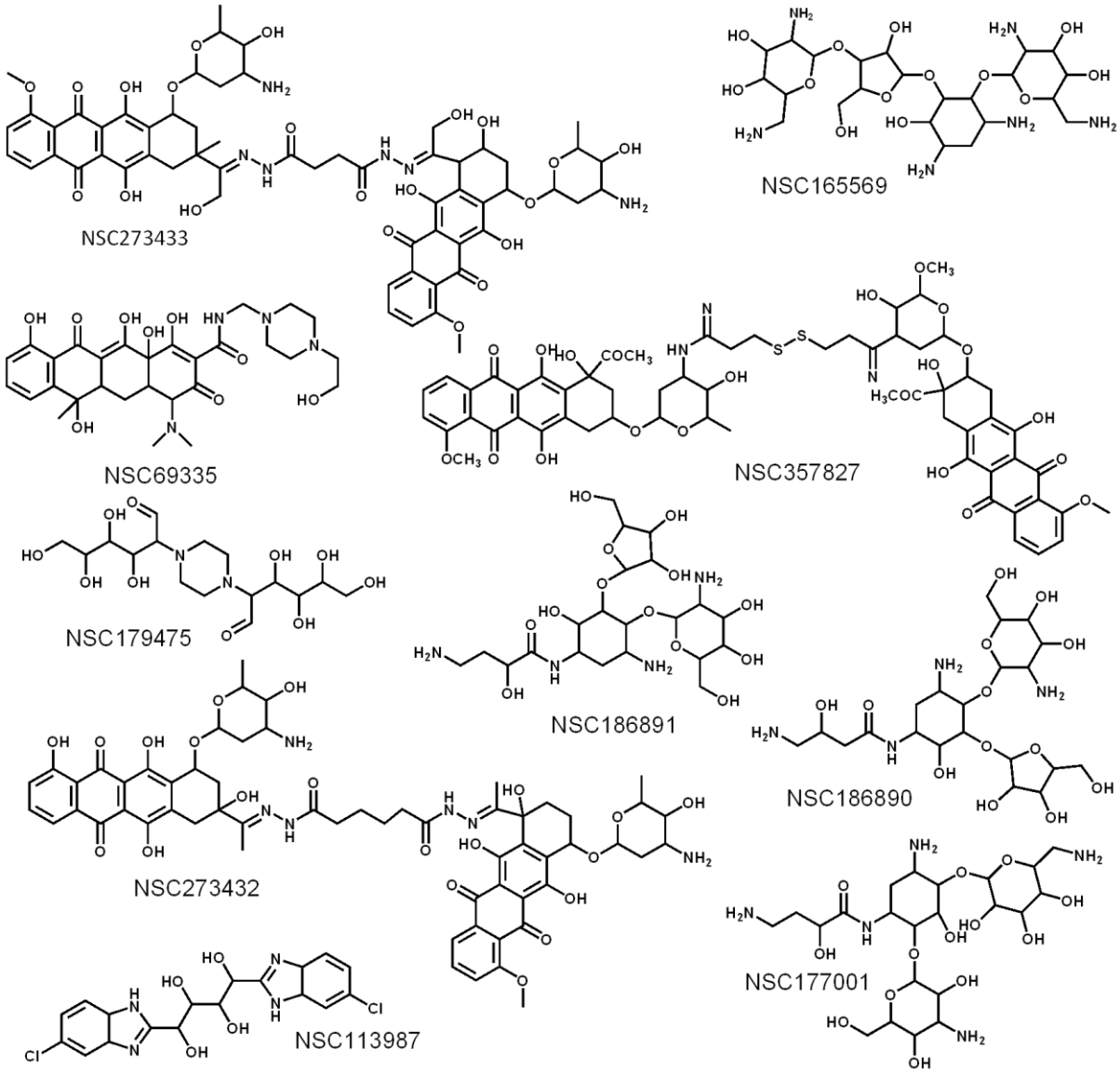



**Fig. 18**

| Ligand | Docking Score |
|---|---|
| NSC273433 | -11.587 |
| NSC165569 | -11.460 |
| NSC69335 | -11.260 |
| NSC357827 | -11.233 |
| NSC179475 | -11.205 |
| NSC186891 | -11.204 |
| NSC186890 | -11.152 |
| NSC273432 | -11.118 |
| NSC177001 | -11.064 |
| NSC113987 | -11.061 |

**Table 5: Docking results**

Many ligands having a planar structure can be effective pharmacophore moieties of DNA-interactive drugs because they can insert between the stacked base paired oligonucleotides. On the other hand, major groove-interacting drugs may act by directly blocking or inhibiting protein-DNA recognition. These ligands are mono- or bis-anthracyclines and exhibit frequently hydrogen bonding capabilities between the ligand groups and the base atoms situated in the DNA double helix. Examination of obtained complex conformations reveals common characteristics need to interact with DNA double helix (Fig. 19):

- presence of -OH group is especially important for binding interactions;
- also ligand size, with bulky ligands that form a higher number of interactions.

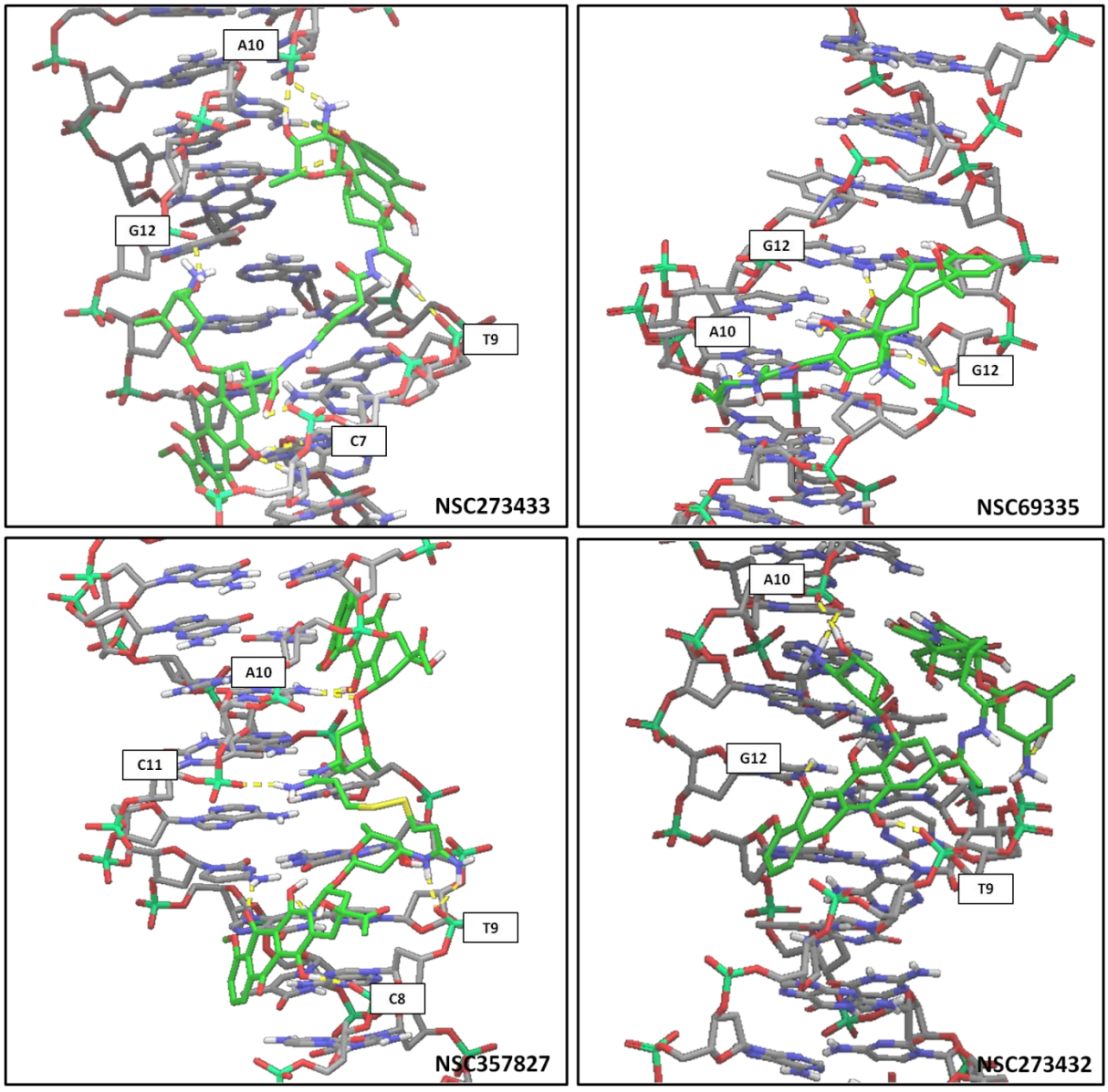



**Fig. 19: Mono- or bis- anthracycline ligands in complex with DNA.**

Once HIF-1 is assembled on the HRE, it must recruit transcriptional coactivators to form an intact initiation complex, a process mediated by distinct transactivation domains. HIF-1α contain two transactivation domains, an oxygen-regulated C-terminal transactivation domain (C-TAD, residues 786–826) and a more centrally located designated N-terminal one (N-TAD, residues 531–575).[127,128]

Both the N-TAD and C-TAD employ recruitment of the coactivators CBP/p300, SRC-1, and transcription intermediary factor 2 (TIF-2),[129–131] although direct interactions have only been demonstrated between the C-TAD and p300 or Creb-binding protein (CBP).[132] CBP and p300 are paralogous transcriptional coactivators that are essential for linking HIF and other transcription factors with coactivator complexes and the basal transcriptional machinery, and are thus indispensable for robust transcriptional activation.

NMR structure was available for the transactivation domain of HIF-1α (PDB ID: 1L3E). The p300 CH1 domain is composed of four α-helices and three $Zn^{2+}$-coordination sites. The three longer helices (designated α1, α2, and α3) pack across each other to form a roughly triangular structure (Fig. 20). The three $Zn^{2+}$ sites lie at the vertices of this triangle. The arrangement of helices α1, α2, and α3 exposes large areas of the hydrophobic core. These hydrophobic regions form the recognition surface for the bound HIF-1α C-TAD. The HIF-1α C-TAD includes four structural elements: an N-terminal extended region, two helices, αA and αB, and an intervening loop.

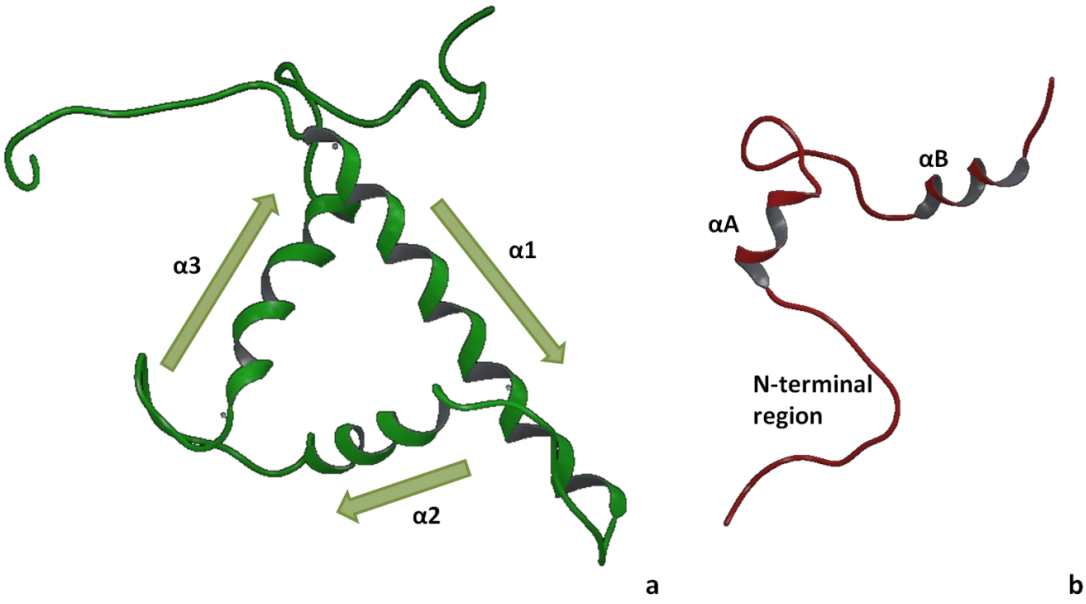



**Fig. 20: p300 CH1 domain (a) and HIF-1α C-TAD domain (b).**

In an effort to describe the interaction site and to explain conformational flexibility, a 20 ns molecular dynamics simulation was performed using the program DESMOND.[133] The structural changes and dynamic behavior were analyzed by calculating the RMSD of the HIF-1α and p300 backbone in function of time (ns) (Fig. 21).

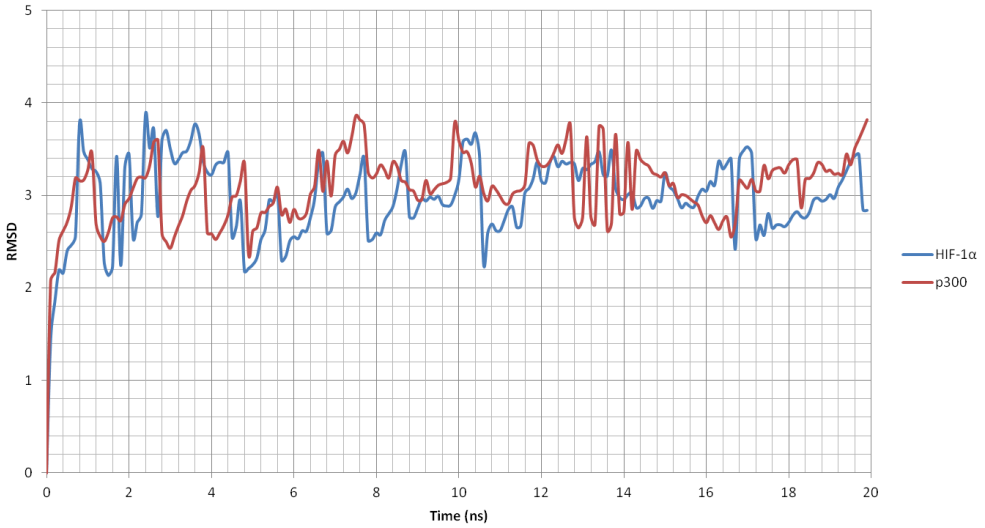



**Fig. 21**

The time evolution of the RMSD with respect to the NMR structure provides a measurement of the convergence of the dynamical properties of the protein. The trajectory of two systems (HIF-1α and p300) maintains quite stable throughout the 20 ns of the simulation with fluctuation of about 1.5 Å.

A series of snapshots from MD simulation was used to create an ensemble of protein configurations (Fig. 22).

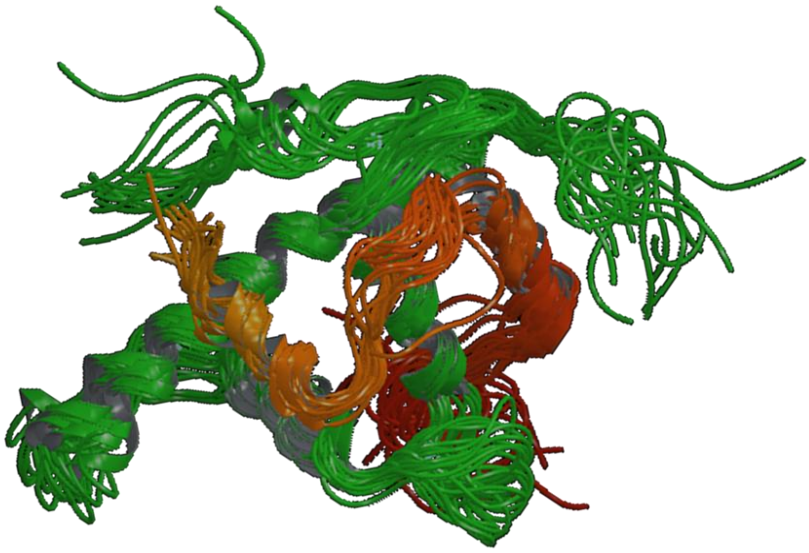


**Fig. 22: HIF-1α (red) and p300 (green).**

Those configurations describe the inherent flexibility of two proteins. Each of the four components of the HIF-1α C-TAD domain makes significant hydrophobic interactions. The N-terminal extended region is anchored by Leu792 and Leu795, which pack into the hydrophobic core of CH1 domain. The intermolecular interactions in the N-terminal region seem to be primarily hydrophobic. Residues 792–795 adopt an extended conformation and make numerous contacts with the surface of the CH1 domain. Hydrophobic contacts are

made by Leu792, which fits into a shallow depression at the $\alpha_1/\alpha_2$ interface, and by Pro793. Additional the side chain of Leu795 projects into a deep hydrophobic cavity (Fig. 23).

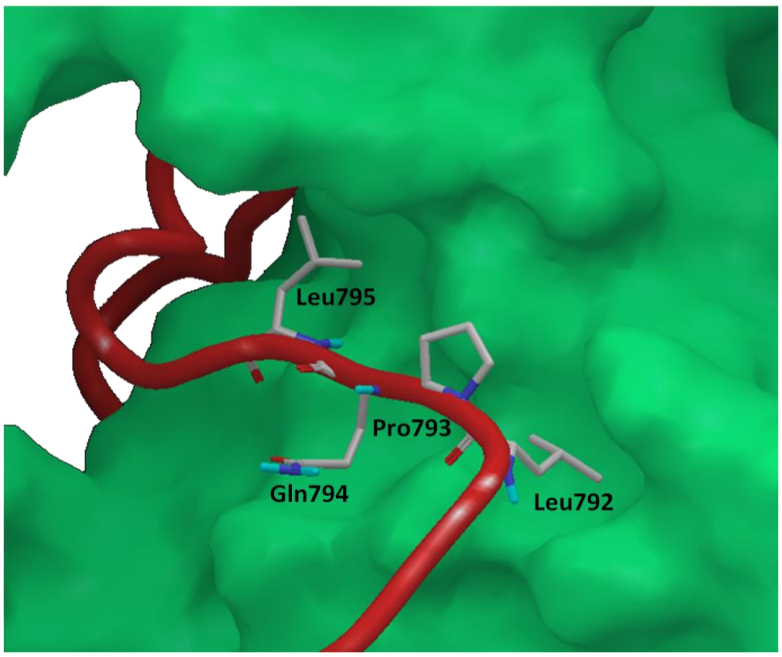



**Fig. 23**

During the 20ns simulation, an extensive hydrogen-bond network links Leu795 with His392 and Ser395 in the CH1 domain. Additional Pro793 in the HIF-1α C-TAD domain forms hydrogen bond with Gln341 (Fig. 24, Table 6). When analyzing the various complexes, the stabilization of the complexes was through hydrophobic interactions and reinforced by H bonds.

| Time | AMINO ACIDS INVOLVED IN HYDROPHOBIC INTERACTIONS | | AMINO ACIDS INVOLVED IN H-BOND | |
|---|---|---|---|---|
| | HIF-1α | p300 | HIF-1α | p300 |
| 0ns | L792, P793, Q794, L795 | Q341, H368, T371, A391, H392, S395 | L792, P793, L795 | Q341, H392, S395 |
| 5ns | S790, G791, L792, P793, Q794, L795, T796 | Q341, P367, H368, T371, M372, V375, H392, S395 | S790, P793, L795 | Q341, H368, H392 |
| 10ns | G791, L792, P793, Q794, L795, T796 | Q341, L345, T371, H392, S395 | P793, L795 | Q341, H392 |
| 15ns | L792, P793, Q794, L795, T796 | T371, M372, V375, V390, A391, H392, S395 | P793, L795 | Q341, S395 |
| 20ns | S790, L792, P793, Q794, L795 | T371, V375, V390, A391, H392, S395 | L795 | A391 |

**Table 6: Amino acids involved in hydrophobic and H-bond interactions for N-terminal extended region of HIF-1α C-TAD.**

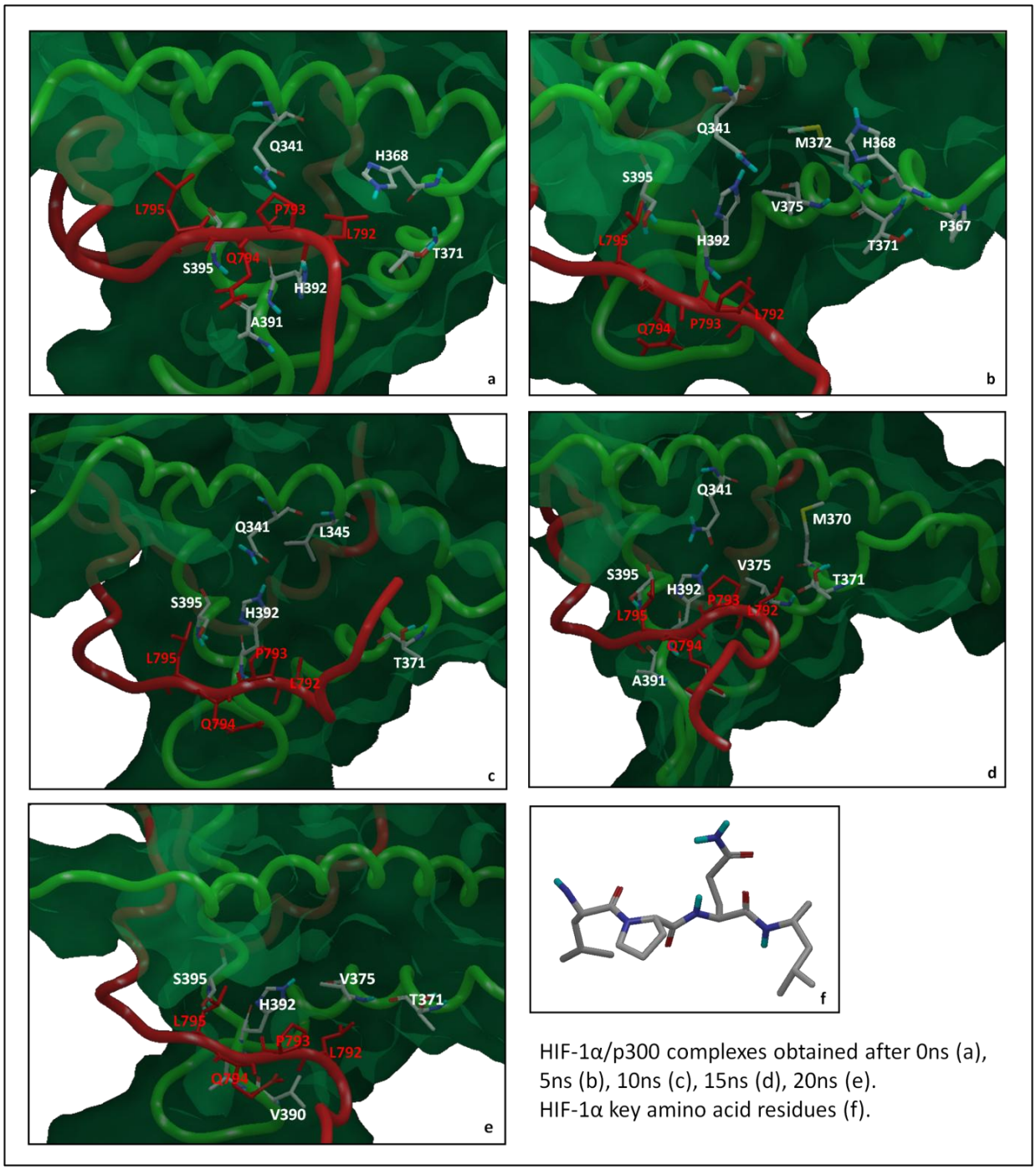



**Fig. 24: Some of the important hydrophobic contacts that define the topology of the interaction during the simulation.**

The HIF-1α key amino acid residues in the N-terminal extended region, shown in Fig. 24f, were used to search in ZINC database[134] small molecules that might be promising as potential inhibitors of the HIF-1α/p300 interaction. Identified structures (~40,000) were docked into the protein target, CH1 p300, using GLIDE *Virtual Screening Workflow*[124]. The chemical structures of ten compounds with the best XP-Docking score are presented in Fig. 25. The conformations of these hits, when bound to CH1 p300, were also analyzed to

determine hydrogen-bonding and hydrophobic interactions. The docking results for the best XP-docking score molecules are given in Table 7, whereas the binding modes (structures and interacting residues) are shown in Fig. 26.

| ZINC ID | XP GSCORE | AMINO ACIDS INVOLVED IN HYDROPHOBIC INTERACTIONS | AMINO ACIDS INVOLVED IN H-BOND |
|---|---|---|---|
| 8442279 | -6.594 | I338, I399 | S395, H402, I399 |
| 8442277 | -6.096 | I338, L344, M372, I399, V413 | Q341 |
| 8442281 | -5.879 | I338, I399, M372, P412, V413 | Q341 |
| 654240 | -5.838 | I338, L342, I399, P412, V413 | D331 |
| 8442211 | -5.791 | L337, I338, I399, P412, V413 | Q398 |
| 3901268 | -5.539 | I338, I399, V413 | H392, S395 |
| 8442187 | -5.293 | I338, A394, I399, P412, V413 | Q398 |
| 18141403 | -4.642 | I338, L342, I399, C406, V413 | H402 |
| 8442211 | -4.449 | L337, I338, I399, P412, V413 | - |
| 8442285 | -4.356 | L337, I338, I399, V413 | Q341, S395 |

**Table 7: Extra Precision (XP) Glide results for the ten lead molecules**

Ligands are associated to the target with hydrogen bonds and hydrophobic interactions. The top scored compounds showed that ligands form hydrogen interactions with Ser395, Gln398, Gln341, and His402. Additionally each of the ten molecules makes significant hydrophobic interactions with the p300 CH1 domain (Ile338, Ile399) that contribute to specific recognition. Also hydrophobic interactions contribute to the stability of the complex.

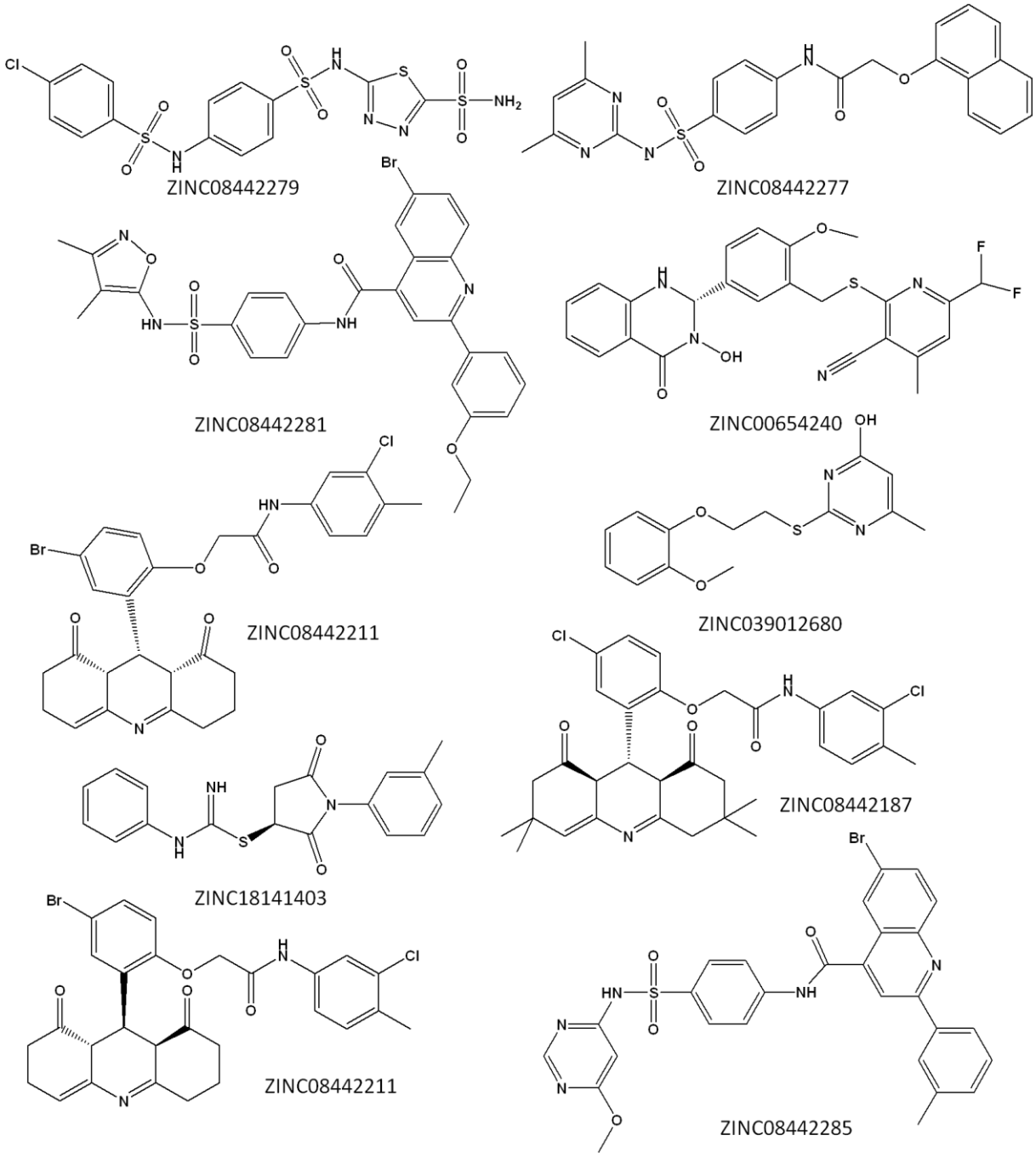

ZINC08442279
ZINC08442277
ZINC08442281
ZINC00654240
ZINC08442211
ZINC039012680
ZINC08442187
ZINC18141403
ZINC08442211
ZINC08442285

**Fig. 25**

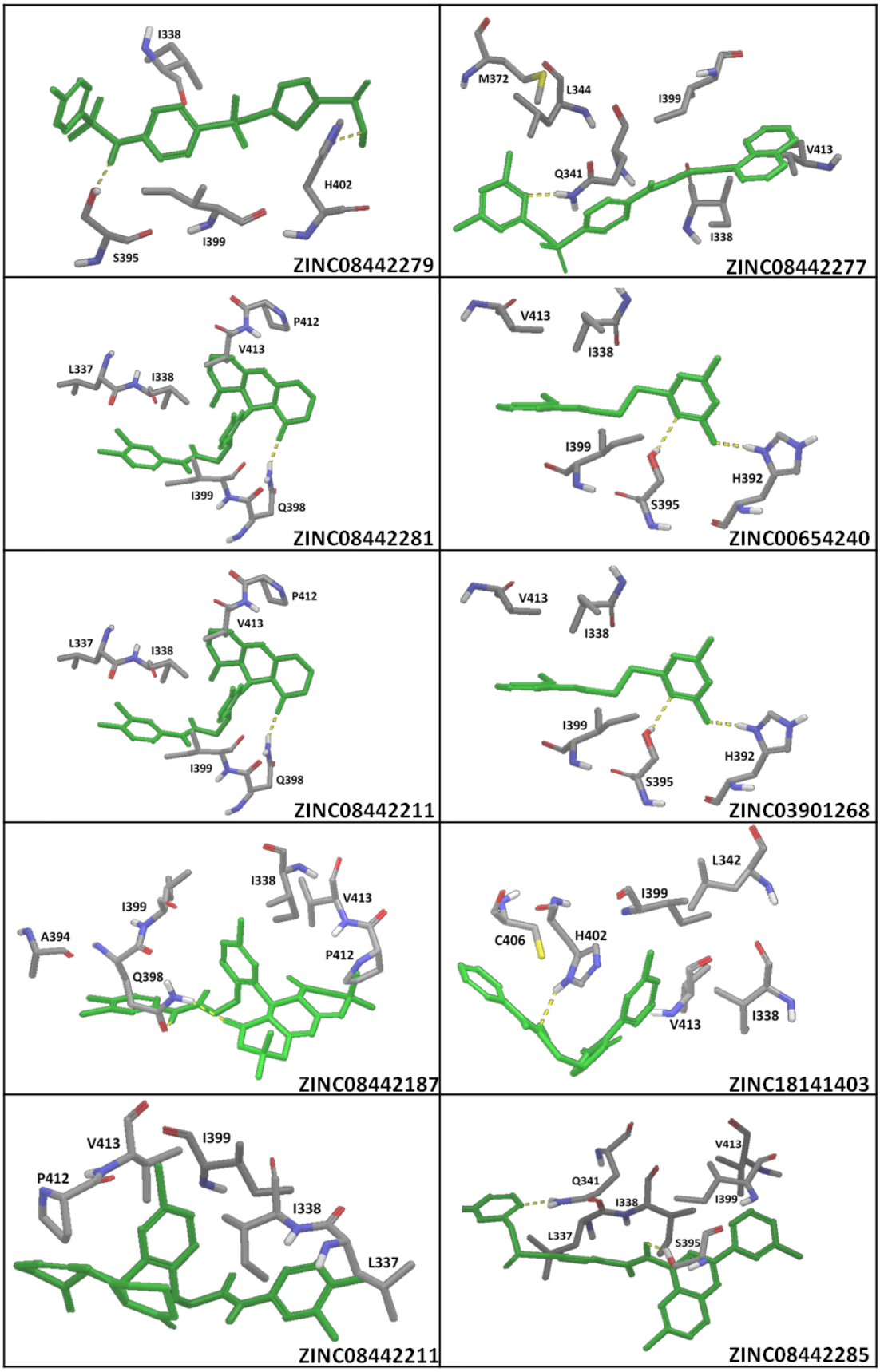

I338
H402
I399
S395
ZINC08442279
M372
L344
I399
V413
Q341
I338
ZINC08442277
P412
V413
L337
I338
I399
Q398
ZINC08442281
V413
I338
I399
H392
S395
ZINC00654240
P412
V413
L337
I338
I399
Q398
ZINC08442211
V413
I338
I399
H392
S395
ZINC03901268
I338
V413
I399
A394
P412
Q398
ZINC08442187
L342
I399
C406
H402
V413
I338
ZINC18141403
V413
I399
P412
I338
L337
ZINC08442211
V413
Q341
I338
I399
L337
S395
ZINC08442285

**Fig. 26**

## 6.2 Mdm2: molecular dynamics studies

The conformational evolution of apo-Mdm2 and bound Mdm2 was investigated by 35 ns MD simulations. Crystal structure of the complexed Mdm2 (PDB ID: 1T4E) was used as starting model. 1T4E, obtained by X-Ray crystallography, contains residues from 16 to 111, and is crystallized with the benzodiazepine derivative [BDZ], *(4-chlorophenyl)[3-(4-chlorophenyl)-7-iodo-2,5-dioxo-1,2,3,5-tetrahydro-4H-1,4-benzodiazepin-4-yl]acetic acid* (Fig. 27). The protein comprises a minimal N-terminal p53 binding domain (16-25), referred to as the lid. While structures of Mdm2 bound to various inhibitors have been solved by X-Ray crystallography and NMR,[135] no information is available concerning the lid behavior in apo-Mdm2. The limited amount of experimental data thereby precluded the accurate modeling of the lid and its precise binding mode to the cleft. In addition, it has been proposed that the binding event induces global conformational changes of Mdm2.

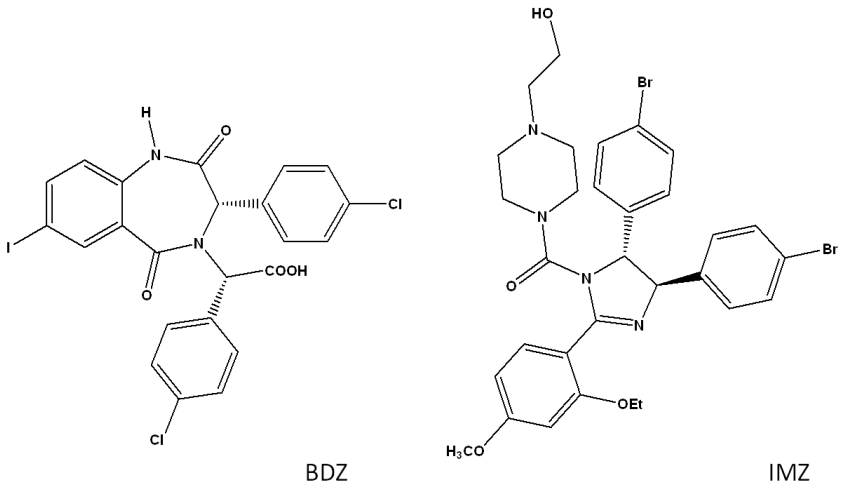



**Fig. 27: Structure of Mdm2 inhibitors.**

Here, we report the characterization of the behavior of human Mdm2 and of the lid in the apo state, and in complex with the inhibitors reported in Fig. 27. To study the apo state the ligand was removed from the protein and a 35 ns MD simulation of the free receptor was first carried out to extensively sample the protein conformations. Moreover a 35 ns MD simulation was carried out on the 1T4E.

Further by using the obtained apo state at 0 ns, a different inhibitor IMZ, *cis-[4,5-bis-(4-bromophenyl)-2-(2-ethoxy-4-methoxyphenyl)-4,5-dihydroimidazol-1-yl]-4-(2-hydroxyethyl) piperazin-1-yl]methanone*, was docked (Fig. 27). Multiple poses of the complex were generated with Induced Fit Docking (IFD) by Schrödinger.[136] Each complex was then ranked according to the IFD score which considers both the docking energy and solvation energy (IFDScore = GlideScore + 5% PrimeEnergy). The docking results are summarized in Table 8.

| IMZ poses | docking score | Glide energy | Prime Energy | IFDScore | RMSD |
|---|---|---|---|---|---|
| 1 | -9.27 | -46.87 | -4613.58 | -239.95 | 0.81093 |
| 2 | -7.36 | -43.55 | -4612.39 | -237.98 | 0.81018 |
| **3** | **-6.49** | **-47.74** | **-4613.70** | **-237.17** | **0.80996** |
| 4 | -1.10 | -42.54 | -4618.49 | -232.03 | 0.81063 |
| 5 | 0.50 | -45.13 | -4617.13 | -230.35 | 0.81249 |

**Table 8: Induced Fit Docking results.**

Visual inspection of the top-ranked poses showed that the IFD-generated model 3 can reproduce a conformation that is almost identical to the native pose of the crystal structure of the Mdm2 bound to IMZ as found in 1RV1 in which however the lid, residues 16-25, were not solved (Fig. 28). So the IFD protocol led to a receptor structure able to bind IMZ that can be used in 35 ns MD simulation.

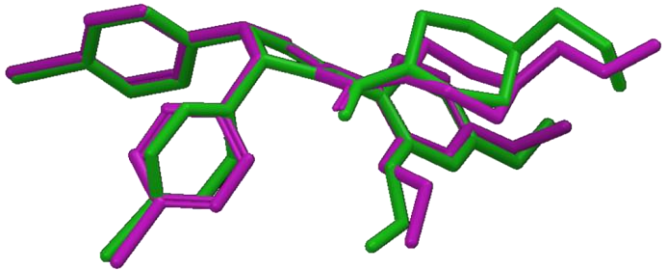


**Fig. 28: Superimposition of Mdm2 bound to IMZ (green) and IFD model (purple)**

We started our analysis by monitoring the structural changes occurring throughout the simulation of Mdm2. The structural changes resulting from the 35 ns simulation can be evaluated in terms of RMS Deviation. So the RMSD of the Cα atoms of the bound (1T4E+BDZ, 1T4E+IMZ) and unbound (1T4E) systems was computed and plotted (Fig. 29).

The trajectory of unbound system maintains quite stable up to about 10 ns and then increases by about 2 Å. After 20 ns increases again, then it remains stable throughout the 15 ns of the simulation, with fluctuation of about 0.5 Å. When the RMSD of the apo form is compared to that of the Mdm2 bound to IMZ, it was observed that during the first 3.3 ns RMSD was similar and comparable. After 3.3 ns it showed a significant increase, so this trend clearly showed major structural deviations in the 3.3–11.3 ns range. During the rest of the simulation dynamics of Mdm2 bound to IMZ remains very stable and was very similar to apo-Mdm2. The RMSD of the Mdm2 bound to BDZ appears different from the previous one, and showed a significant increase during the first 6 ns of simulation. After that point the dynamics of the protein remained relatively stable, with fluctuation of ≈ 0.5-0.6 Å. In this range, RMSD was comparable with that observed in the Mdm2 bound to IMZ. During the rest of the simulation it showed a substantial increase on the calculated RMSD, therefore major conformational changes on the protein take place. Analysis of evolution in dynamics

of RMSD for Cα atoms showed that the structures have important variation, but the Mdm2-BDZ is stabilized in the last 15 ns of simulation. In contrast to Mdm2-BDZ, a small fluctuation of RMSD values is observed for other two simulations. Similar behavior of apo-Mdm2 was found in simulations carried out by Carlson[137] and Pellicciari[138]. The trajectory of the p53-Mdm2 complex was very similar to that observed for Mdm2 bound to IMZ. p53-Mdm2 complexes were highly flexible in simulation, and IMZ and BDZ complexes appear to follow this trend.

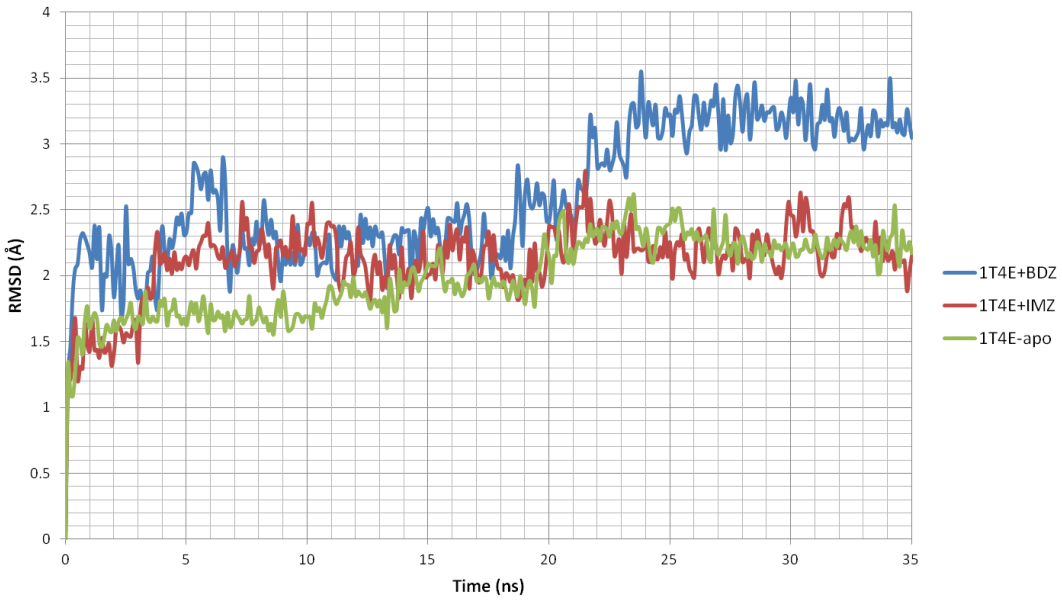



**Fig. 29: RMSD of 1T4E apo and bound forms (BDZ and IMZ)**

The RMSD of backbone atoms in the binding pocket (Leu54, Leu57, Ile61, Met62, Tyr67, Val75, Phe86, Phe91, Val93, Ile99, and Ile103) is presented in Fig. 30. In apo-Mdm2, the binding pocket was stable throughout the trajectory with fluctuation <0.5. Binding pocket dynamics of Mdm2 bound to IMZ was very similar to apo-Mdm2. Instead in presence of BDZ, the RMSD gradually increases. When the RMSD of the lid is compared for apo and bound Mdm2 (Fig. 31), it was observed that IMZ induced conformational changes in the lid. The trajectory of lid in the apo Mdm2 is quite stable over the timescale of simulation. As seen in the figure, the RMSD of Mdm2-IMZ lid was ≈ 2 Å during the first 5 ns, but a rise in RMSD of about 2 Å was observed in subsequent ns. A rise in RMSD was observed for Mdm2-BDZ lid between 1 and 24 ns. During the rest of the simulation RMSD showed a significant decrease: this suggest the probable interaction between BDZ and the lid that helps to stabilize lid conformation.

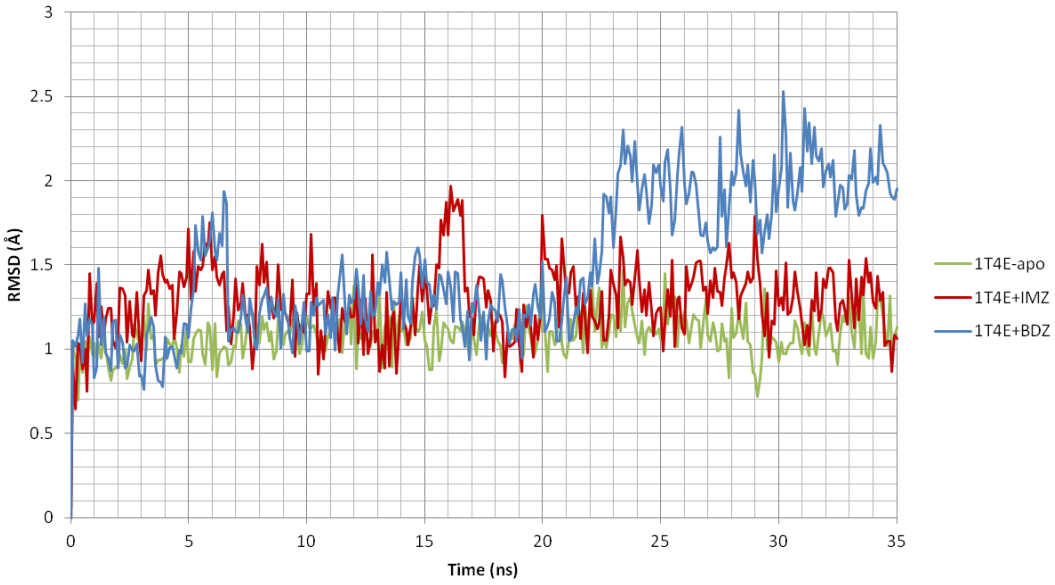



**Fig. 30: RMSD of binding pocket for 1T4E apo and bound forms.**

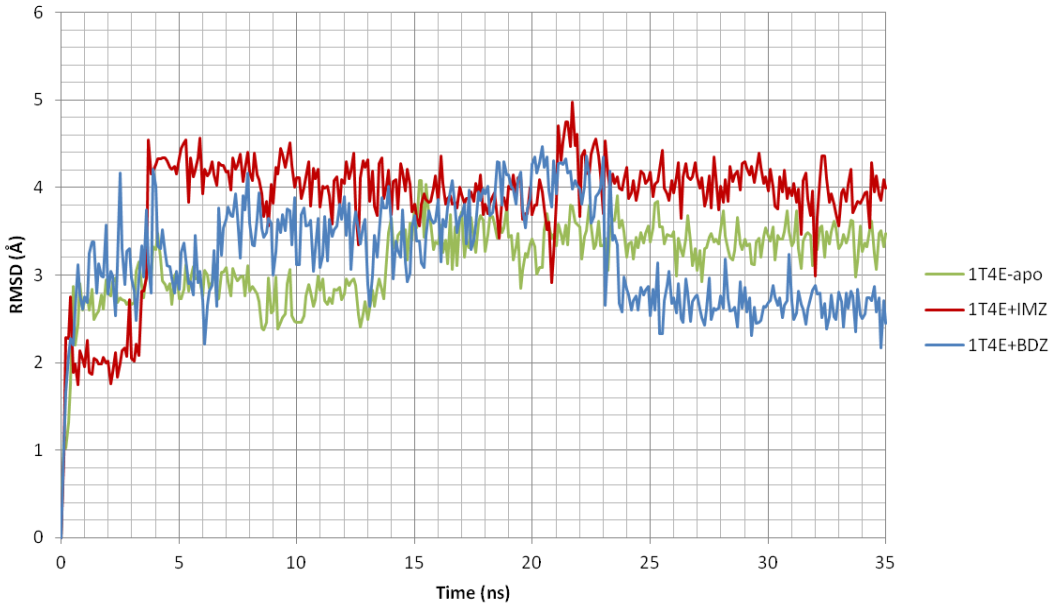



**Fig. 31: RMSD of lid for 1T4E apo and bound forms.**

The simulation findings indicate the structural stability of apo protein, while the RMSDs of bound systems showed structural mobility. The inhibitors in the binding pocket caused the displacement of the lid located above the binding cleft. Also the movement of binding cleft is adaptive, the inhibitors allow the binding cleft to be more flexible and to better adapt itself.

We used PCA as a guide to identify significant dynamic processes from this very large data set. The first step in PCA is the construction of the covariance matrix, which captures the degree of collinearity of atomic motions for each pair of atoms. The covariance matrix is subsequently diagonalized, yielding a matrix of eigenvectors and a diagonal matrix of eigenvalues. Each of the eigenvectors describes a collective motion of particles, where the

values of the vector indicate how much the corresponding atom participates in the motion. The sum of the eigenvalues is a measure of the total motility in the system. Compared to bound proteins (7.07 $nm^2$ for 1T4E-BDZ, 4.25 $nm^2$ for 1T4E-IMZ) the total sum of eigenvalues of apo system was relatively small (3.53 $nm^2$), indicating a more rigid structure.

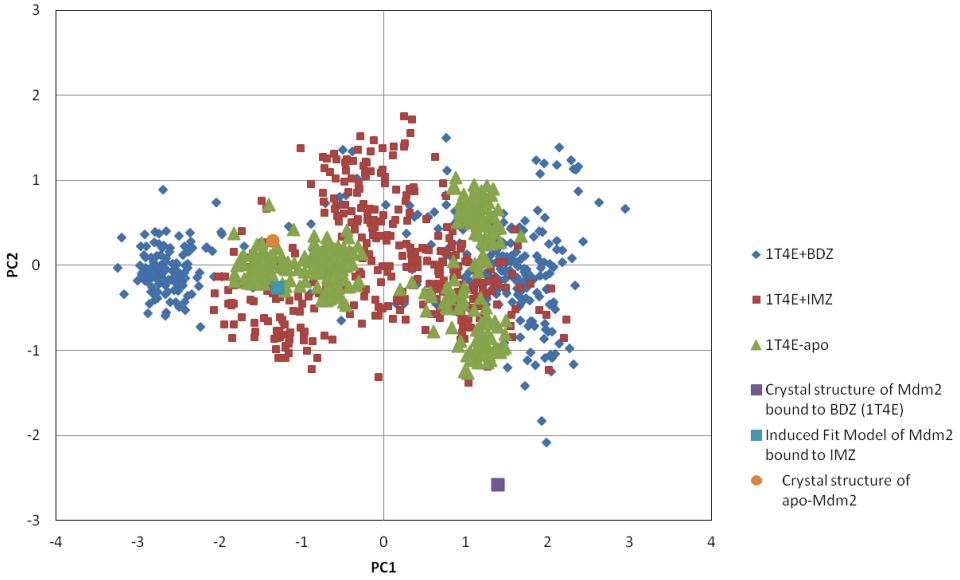



**Fig. 32: Principal components analysis computed for Mdm2 in various states: apo (green), complex with IMZ (red) and complex with BDZ (blue).**

PCA technique decomposes the intrinsic flexibility of a protein into motions of different frequencies of vibrations. These are then ordered such that the first component (PC1) characterizes the motion with the largest amplitude and lowest frequency. Principal components were computed for Mdm2 in various states: apo, complex with IMZ and complex with BDZ (Fig. 32).

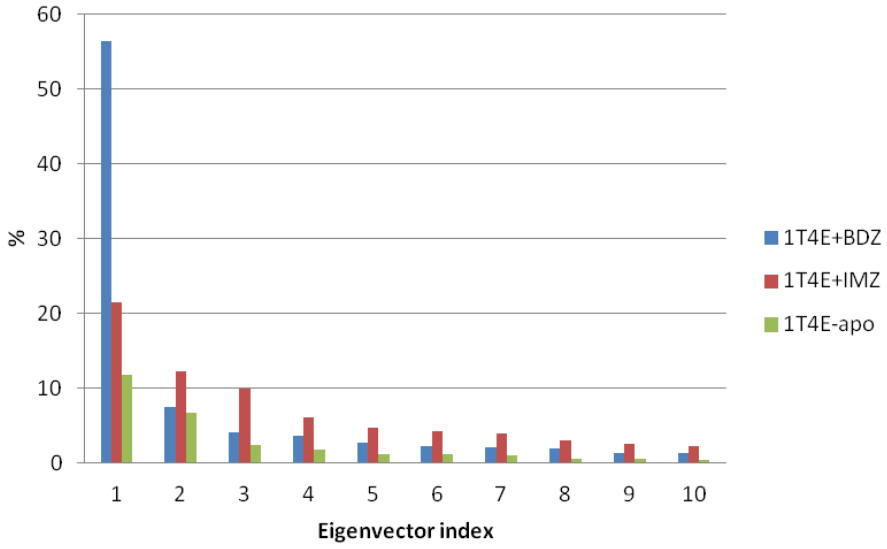



**Fig. 33: Contributions of first ten principal components.**

PC1 dominates the motions of Mdm2 bound to BDZ (covering 56% of overall motion). The motion is more distributed across PC1 and PC2 in Mdm2 bound to IMZ (21.5% PC1, 12.2% PC2) and apo-Mdm2 (11.7% PC1, 6% PC2). Histogram in Fig. 33 refers to the contributions of first ten individual principal component to the overall fluctuations.

The distribution of the structures in the phase space defined by principal components 1 (PC1) and 2 (PC2) shows that the BDZ and IMZ bound states are more flexible than apo one.

The evolution of the secondary structure of the three systems was also analyzed in order to determine if the dynamics of Mdm2 involves changes in the secondary structure pattern. In order to quantitatively measure the mean backbone mobility for each residue, the root mean square fluctuations (RMSF) relative the average structure of apo-Mdm2 were calculated (Fig. 34). It was observed that the mobility of the unbound system was relatively low for the regions including residues 26–76, excluding the N-terminus. In this region the RMSF was < 1 Å.

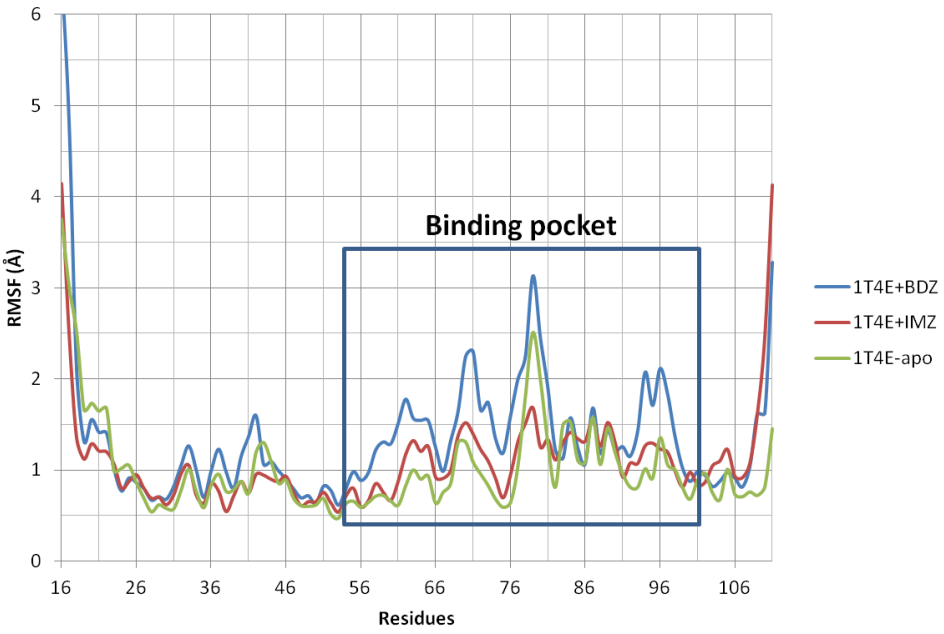



**Fig. 34: RMSF of 1T4E apo and bound forms (BDZ and IMZ).**

Few residues (42-44 and 69-71) displayed a mobility higher than 1 Å. Specifically, residues 69-71 form a hinge that connects helix α2 with the region formed by sheets β1'. There were a few residues that showed larger degree of flexibility throughout the 35 ns simulation (RMSF > 1.3 Å) in the region including amino acids 78-81. When comparing the mobility of apo-Mdm2 to that observed for Mdm2 bound to IMZ, a very similar trend of RMSF was observed. The mobility of Mdm2 bound to IMZ is lower than apo-Mdm2, except for residues between 61-74, in proximity of the amino acids which constitute the hydrophobic binding cleft (Met62, Tyr67, Val75). Excluding the N and C-terminus, few residues displayed a mobility higher than 1 Å. The residues that showed the highest flexibility were Lys70, Ser78, and Asn79 (RMSF ≈ 1.5 Å), located around the binding cleft. Residues which constitute

binding cleft showed larger degree of flexibility. The highest RMSFs highlight the ability to better adapt itself to ligand (BDZ, IMZ).

To qualitatively analyze the conformational changes experienced by the protein, the superimposition of three clusters of structures was analyzed (Fig. 35-37).

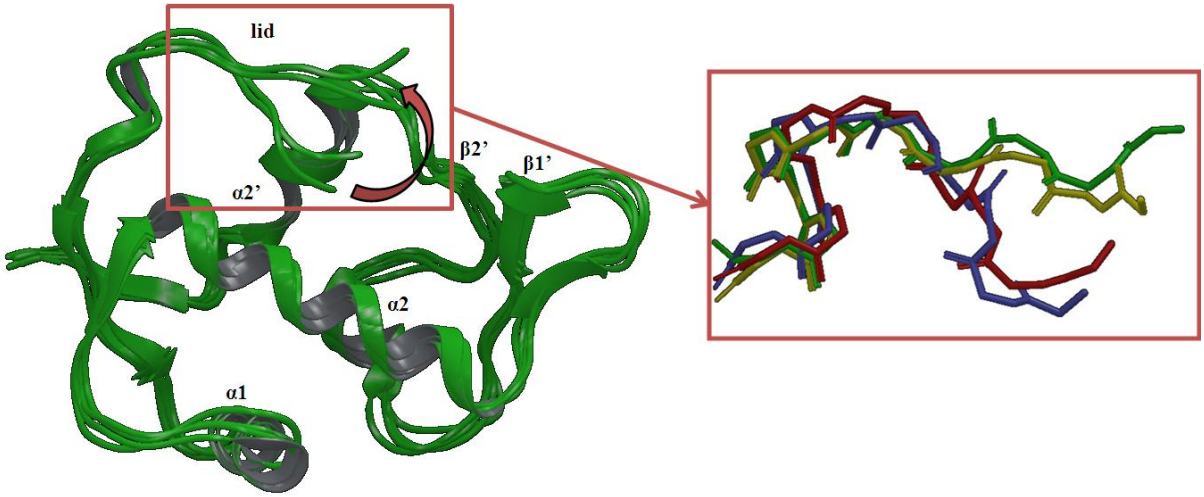



**Fig. 35: Superimpositions of 1T4E clusters apo forms. The backbone of apo-Mdm2 at 0 ns (blue), 10 ns (red), 20 ns (yellow) and 35 ns (green) are showed in square.**

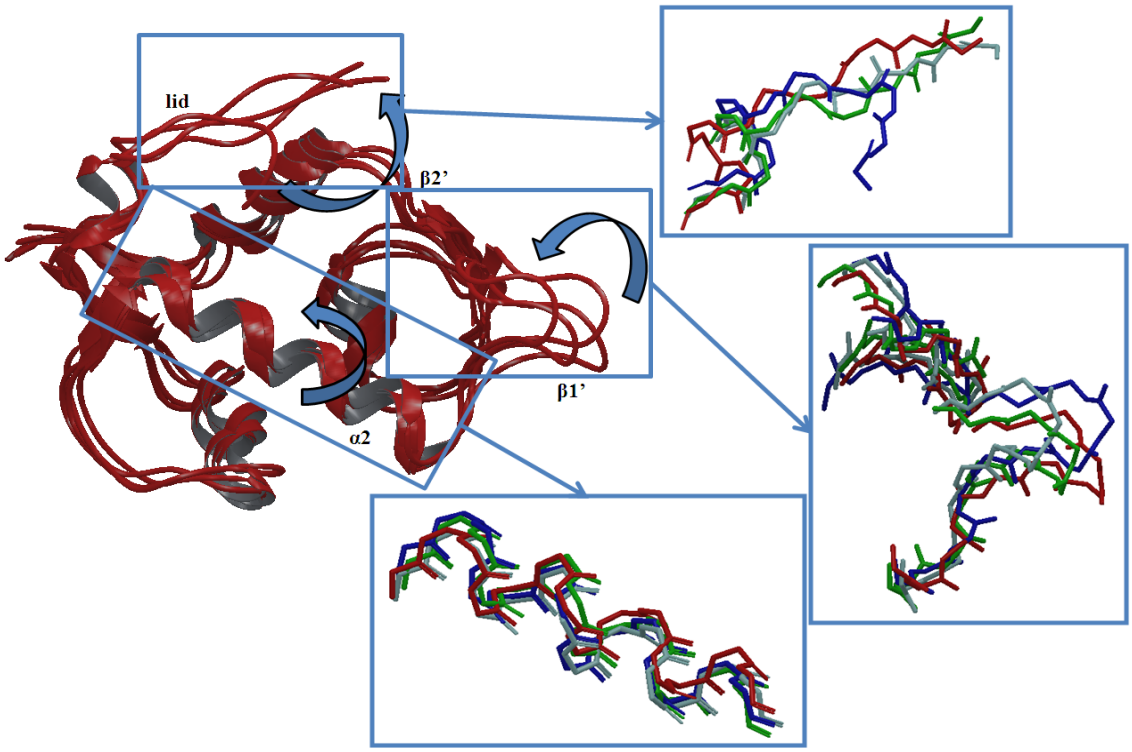



**Fig. 36: Superimpositions of 1T4E clusters bound to IMZ. The backbone of Mdm2-IMZ at 0 ns (blue), 10 ns (red), 20 ns (green) and 35 ns (turquoise) are showed in square.**

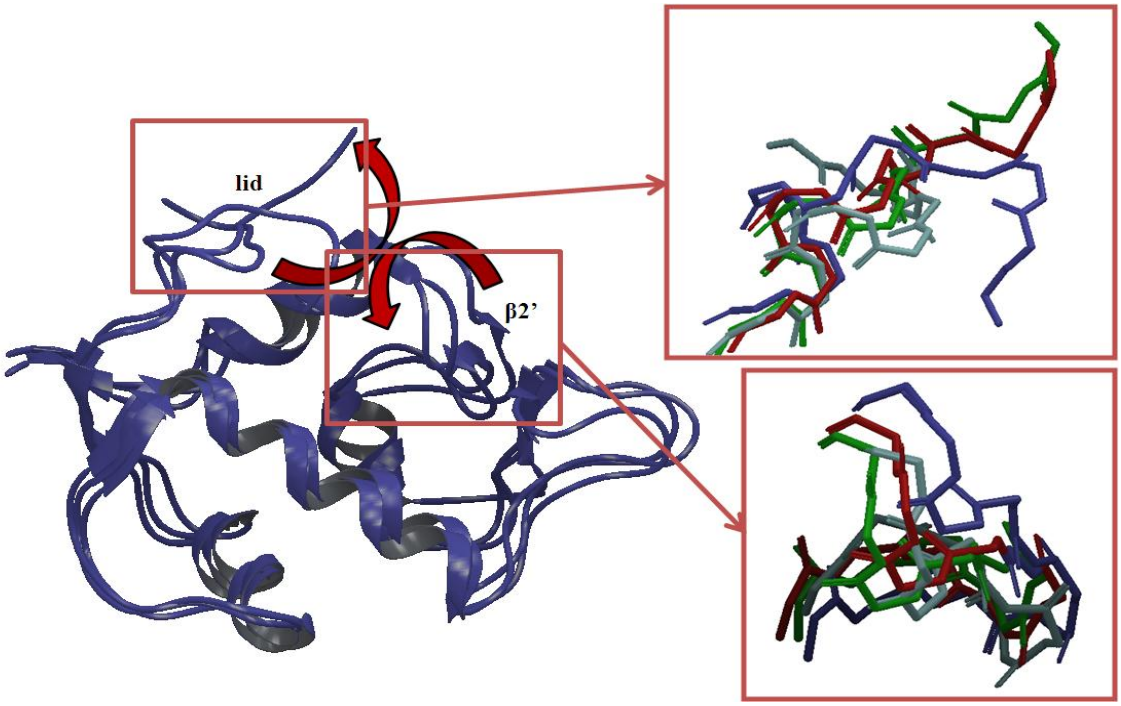



**Fig. 37: Superimpositions of 1T4E clusters bound to BDZ. The backbone of apo-Mdm2 at 0 ns (blue), 10 ns (red), 20 ns (green) and 35 ns (turquoise) are showed in square.**

Considering apo-Mdm2 superimpositions, we observed that the lid of the cleft was located just above the binding cleft in the first 13 ns. So residues 19–24 form a flexible lid burying a hydrophobic active site. Between 13 and 19 ns the lid displayed a movement outwards the cleft, opening the binding site. In Mdm2-IMZ system we observed that the lid of the cleft displayed a high degree of flexibility and a wide movement outwards the cleft. The helix α2 experienced an inward movement toward the center of the cleft. Furthermore, sheets β1'-β2' showed a movement toward the center of the binding cleft leading to a more closed conformation of the binding pocket. Similarly, in the case of Mdm2-BDZ superimposition, the helix α2' experienced an inward movement toward the center of mass of the cleft and the hinge that connected helix α2' with sheets β2', the floor of the cleft, showed a wide movement towards the center of the binding cleft. This inward movement was facilitated by the twisting motion of the hinge formed by sheets β1' and β2'. The movement of helices and hinges, which constitute the cleft, contributes to the global rearrangement of the cleft, as observed in our simulations. The shape of the binding cleft changes and the cleft's movement is influenced by the movement of hinges and helices located around. Also the binding site shape and size are defined by the ligand.

The superimposition of the backbones of apo-Mdm2 and bound Mdm2 was complemented with the quantitative measurement of the changes in the accessible area (Fig. 38).

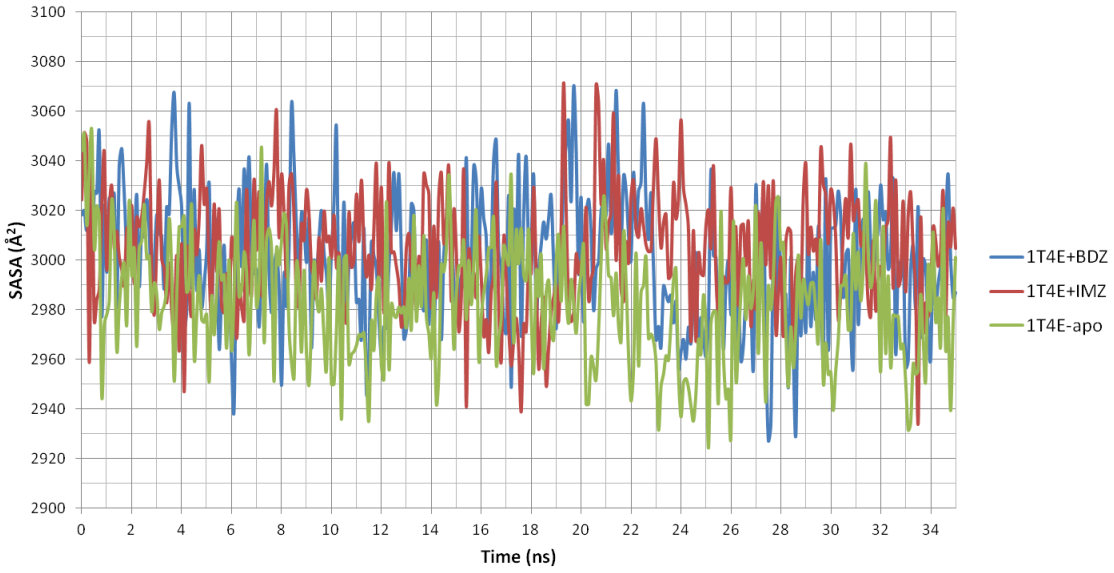



**Fig. 38: SASA of 1T4E apo and bound forms (BDZ and IMZ).**

Thus, the solvent-accessible surface area (SASA) was calculated for all atoms in Mdm2, using a probe radius of 1.4 Å. In the apo-Mdm2 simulations the value of SASA decreased in the first 13 ns to a minimum value of ≈ 2930 $Å^2$ and then increased between 13 and 19 ns, indicating that the structure assumes a less closed conformation as confirmed by RMSF. Between 19 and 23 ns the value of SASA decreased again, whereas in following ns the SASA value fluctuates around the average value of 2970 $Å^2$. It was observed that, in our Mdm2-IMZ simulations, the value of SASA decreased in the first 4 ns to a minimum value of ≈ 2946 $Å^2$ and then increased during the successive ns of simulation. Between 5 and 10 ns, the SASA value fluctuates around the average value of 3004 $Å^2$. After 15 ns the value of SASA decreased to a minimum value ≈ 2930 $Å^2$, then dramatically increased (≈ 3072 $Å^2$), indicating that the structure assumes a less closed conformation. For Mdm2-BDZ system, the value of SASA decreases until a minimum value of ≈ 2977 $Å^2$ in the first ns of simulation. In following ns the SASA value fluctuates around the average value of 3005 $Å^2$. After 22.5 ns the SASA value drastically decreases by about ≈ 100 $Å^2$ indicating that again the structure assumes a less closed conformation. After that point, it remained fairly stable during the rest of simulation, around the value of 2993 $Å^2$. A decrease of the SASA on the apo Mdm2 structure was observed, supporting the fact that the simulation converged to a closer structure compared to that obtained for bound systems. In fact the presence of the ligands influenced the cleft shape that assumes a less closed conformation. For the apo system the lid situated above the binding cleft displayed a lower degree of flexibility forming a less exposed cavity.
Another interesting feature observed in our simulations was the dynamics of the side chains of key residues at the binding pocket (χ angle), Leu54, Leu57, Ile61, Met62, Tyr67, Val75, Phe86, Phe91, Val93, Ile99, and Ile103. Leu57 of Mdm2 interacts with p53 *via* the formation of a hydrogen bond between the Trp23 ε-nitrogen of p53.

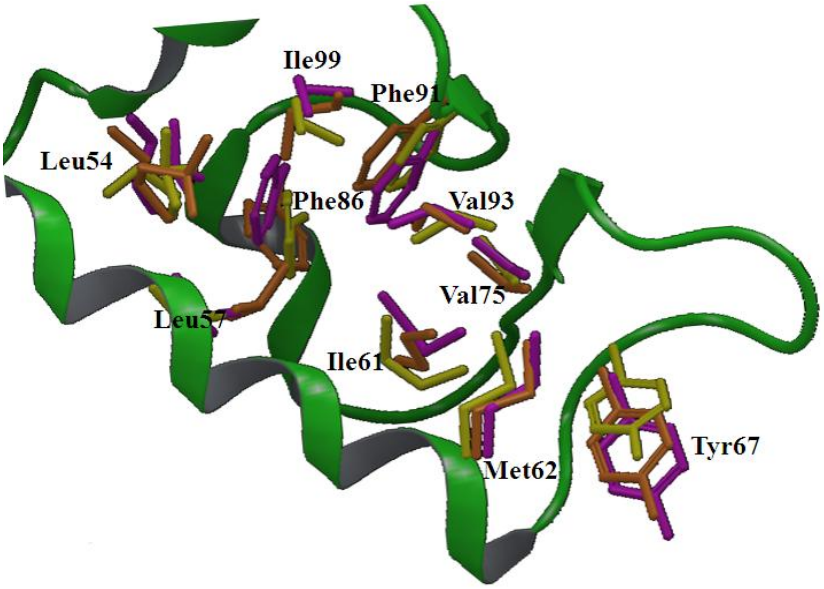



**Fig. 39: Side-chain motions of 1T4E apo forms at 10 ns (yellow), 20 ns (orange), 30 ns (purple).**

When Mdm2 is unbound (Fig. 39), Leu57 displayed rapid transitions going from -170° to +170°. No significant differences were observed in the type of motions displayed by Met62, Tyr67, Val75, Phe86, and Leu54. Also for Phe91 and Val93 no substantial differences in the type of motions were observed for the first 23 ns. After that point they displayed rapid transitions.

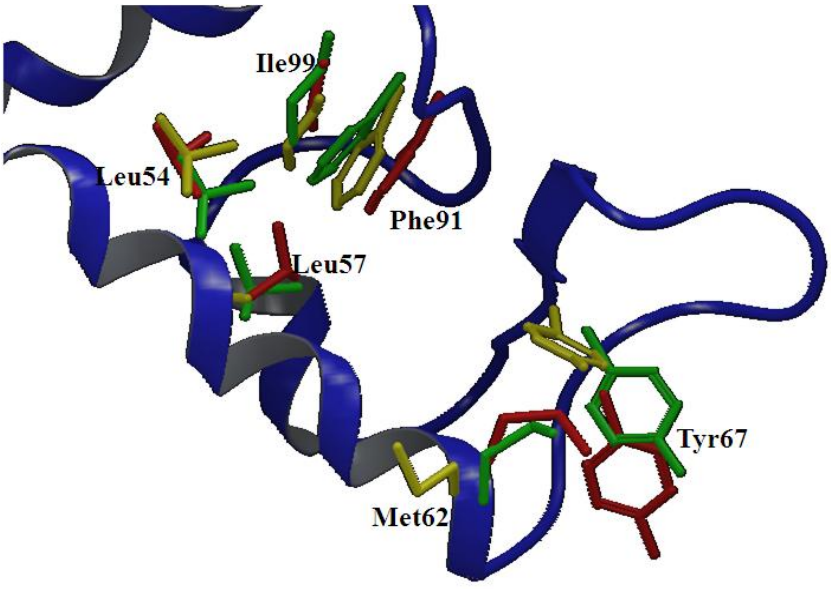



**Fig. 40: Side-chain motions of 1T4E bound to BDZ at 10 ns (yellow), 20 ns (green), 30 ns (red).**

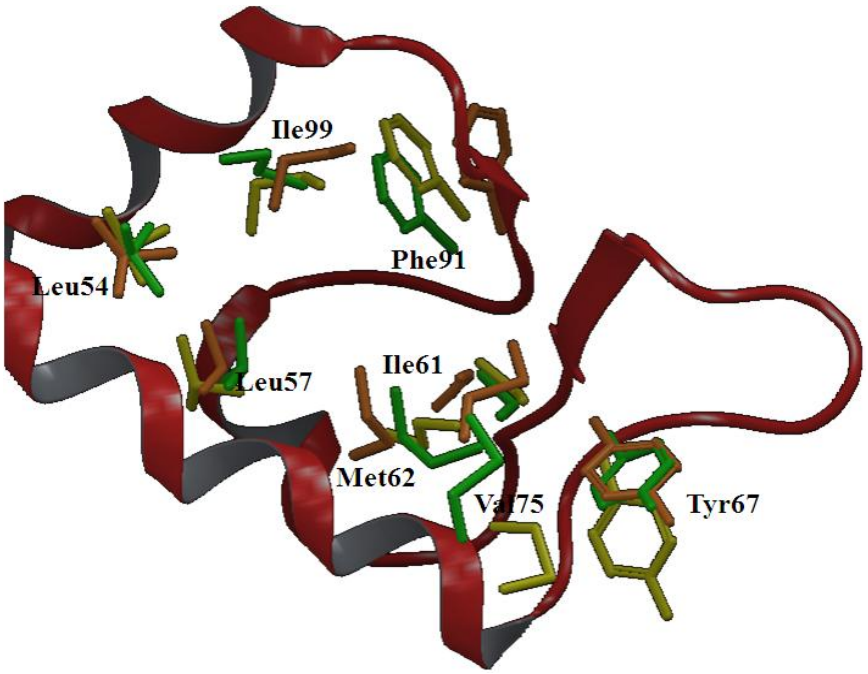



**Fig. 41: Side-chain motions of 1T4E bound to IMZ at 10 ns (yellow), 20 ns (orange), 30 ns (green).**

In the case of Mdmd2 bound to BDZ (Fig. 40), few differences were observed in the type of motions displayed by Leu54, except rapid transitions going from -60˚ to +164˚ and from -145˚ to +178˚ after 16 and 26 ns respectively. A similar situation was noticed for Leu57, which displayed transition movements in the range from -160˚ to +170˚, and for Met62. Also few differences were observed in the type of motions displayed by Tyr67 and Phe86. Val75 displayed rapid transitions going from -170˚ to +177˚ and a similar situation was noticed for Phe91 and Val93. Comparing the Mdm2-IMZ librational motions to the Mdm2-BDZ librational motions (Fig. 41), a slight different behaviour was observed. A similar situation was noticed for Leu54, Leu57, Val75 and Val93, which displayed transition movements in the range from -180˚ to +170˚. Similarly few differences were observed in the type of motions displayed by Tyr67, Phe86 and Phe91. The binding of two different ligands (IMZ and BDZ) to the cleft requires specific side-chain movements and local arrangements of the cleft. Side-chain fluctuations of the amino acids located at the binding cleft allow the binding cleft to be more flexible and to adapt itself to incoming ligand. The compounds studied herein tried to mimic the domain of p53 that binds to Mdm2.

To investigate the binding mode of these compounds, we analyzed the structure of the last snapshot of Mdm2 bound to the two different inhibitors (IMZ and BDZ) obtained after 35 ns simulations.

When the Mdm2–IMZ at 0 ns was analyzed, the structure verified that the inhibitor binds to the p53 binding site on Mdm2. The inhibitor mimics the interactions of the p53 peptide to a high degree, with one bromophenyl moiety sitting in the Trp pocket, the other bromophenyl group occupying the Leu pocket, and the ethyl ether side chain directed toward the Phe pocket.[139] In essence, the imidazoline scaffold replaces the helical backbone of the peptide and is able to direct, in a fairly rigid fashion, the projection of three groups into the pockets

normally occupied by Phe19, Trp23, and Leu26 of p53.[86] As shown in Fig. 42, the bound conformation of IMZ was similar to the p53 peptide; for the p53-bound Mdm2, we used the X-ray structure 1YCR.

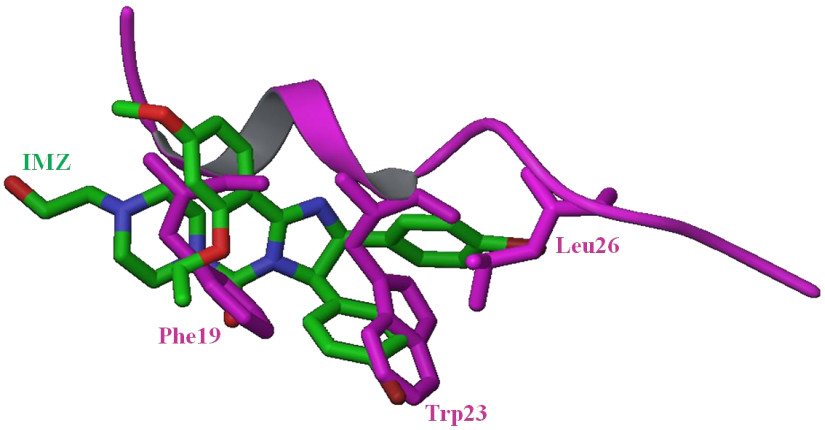



**Fig. 42: IMZ superimposed to bound p53 (PDB ID: 1YCR).**

IMZ interacts with Gln72 through the hydroxyethyl group of the IMZ, which donates a hydrogen bond to the side-chain oxygen of Gln72. We observed that one bromophenyl group established non polar interaction with Leu54, the other bromophenyl group with Leu57, Phe86, Phe91, Ile99 and Ile103. In addition the ethyl ether side chain of the IMZ established non polar interaction with Tyr67 and Ile61. When the protein-ihibitor complex, as obtained after the simulation, was analyzed, we found that bromophenyl groups established non polar interaction with Leu57, Gly58, Ile61, Val93 and His96. The hydroxyethyl group of the IMZ formed non polar interaction with Gln72, the piperazine group with Tyr67, and 2-ethoxy- 4-methoxyphenyl group with His73 and Ala43.
In the other investigated complex, 1T4E-BDZ, the inhibitor occupies the same pockets of the peptide p53 side chains Phe19, Trp23, and Leu26 (Fig. 43).

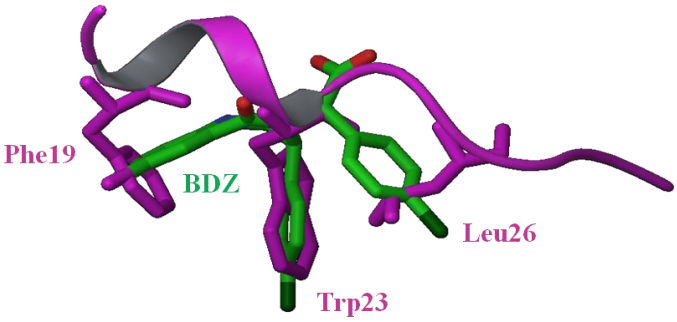



**Fig. 43: BDZ superimposed to bound p53 (PDB ID: 1YCR).**

As in the case of bound peptides, the Mdm2 interactions with the inhibitor were largely nonspecific van der Waals contacts. The BDZ pendant groups orient themselves to mimic the position of the hydrophobic side chains on one face of the helical p53 peptide ligand and thus act as an α-helix mimetic, with one chlorophenyl moiety sitting in the Leu pocket, the other occupying the Trp pocket, and the dioxo-tetrahydro-benzodiazepine moiety in the Phe pocket. When the protein-inhibitor complex, as obtained after 35 ns MD, was analyzed, we found that chlorophenyl groups established non polar interaction with Val93, Lys94, His96, Tyr100 and Ile103, the core benzodiazepine ring formed non polar interaction with Ile61, Tyr67, Lys94 and Gln72.

These findings are in partial agreement with the data reported in the paper by Pellicciari et al.[138] in which the authors performed a linear discriminant analysis (LDA) with the aim of identifying specific residues whose conformational changes are the marker of the apo and p53-bound states of Mdm2 and MdmX.[138] In IMZ-Mdm2 it was possible to identify two residues that are involved in the shape arrangement of the p53 binding site: Gln72 and His73. Gln72 and Tyr100 are identified in BDZ-Mdm2. Other residues, identified as involved in the shape arrangement of the binding site, were Ala43, Leu57, Gly58, Ile61, Tyr67, His73, Val93, Lys94, His96, and Ile103.

Therefore ligands make extensive van der Waals contacts with residues in the peptide binding cleft comprised of the structural elements α2, β1', and α2'. These residues are: Leu54, Leu57, Ile61 and Met62 in helix α2; Tyr67, Phe91, and Val93 in β1' sheet; and His96 and Tyr100 in helix α2'. These interactions support the observed changes in the backbone residues known to be in direct contact with the ligands. Both ligands caused similar changes in the peptide binding cleft. It was observed that, in our Mdm2-IMZ simulation, Leu57, Ile61 and Met62 (in helix α2) showed an inward movement toward the center of the binding cleft. On the contrary Phe91 showed an outward movement of about 1.5 Å. The larger difference was in Phe86, that showed an inward movement of about 3 Å. In Mdm2-BDZ simulation Ile61, Met62, Tyr67 and Phe91 showed an inward movement toward the center of the binding cleft. The larger differences were in His96 and Val93, that showed an inward movement of about 2 Å. The main conformational changes were found in linker turn regions or in the β sheets. These conformational changes seem to be sufficient to accommodate ligands that differ in their length: the changes alter the size of the cleft and indirectly affect the angle of the bottom α helices with respect to the cleft. The analysis of the results obtained by molecular dynamics of apo system showed a stable and less flexible structure, with the lid closing the binding cleft of the protein and leading to a closed conformation of the cleft (Fig. 44).

These findings are confirmed by SASA mean values that highlight the importance of the ligand affecting the binding cleft by assuming a different conformation in Mdm2 with respect to the apo form. Mdm2 shows an high solvent exposition in the bound form with respect to the apo form, with this behavior being ascribed to changes in the size of the cleft and to a high degree of flexibility of the lid in bound systems.

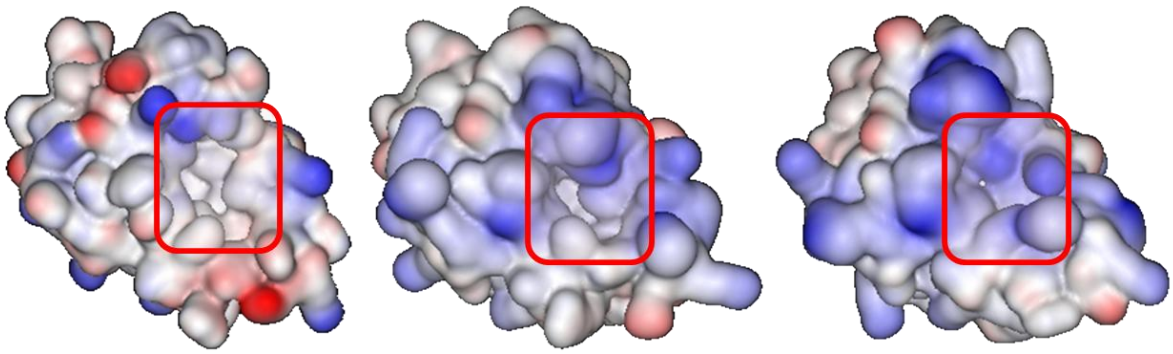


**Fig. 44: apo-Mdm2 snapshots; binding pocket highlighted in the box.**

On the contrary in the bound systems the lid displayed a high degree of flexibility and a wide movement outwards the cleft, forming a wider and more exposed cavity (Fig. 45-46).

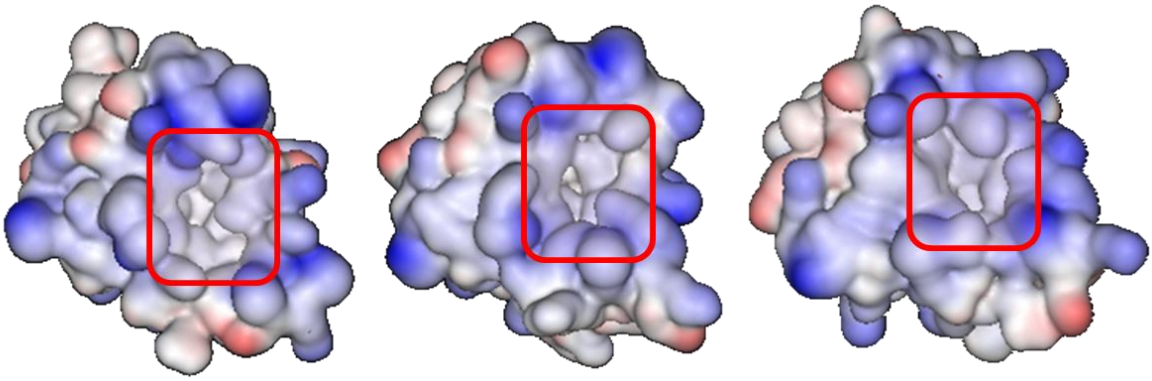


**Fig. 45: Mdm2-IMZ snapshots; binding pocket highlighted in the box.**

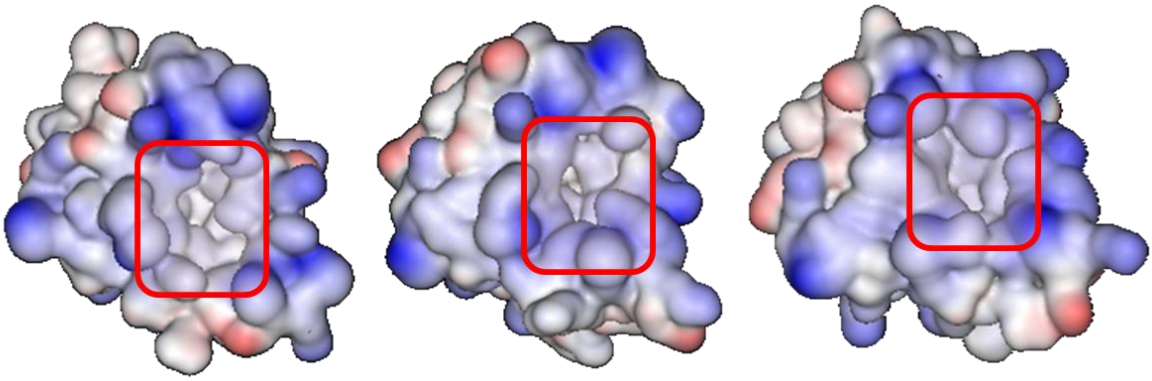


**Figura 46: Mdm2-BDZ snapshots; binding pocket highlighted in the box.**

In summary, in this study, the influence on Mdm2 structure upon binding of different ligands was analyzed. A rearrangement and an outward expansion of the Mdm2 helices, surrounding the binding cleft in bound systems, was observed. Ligand IMZ caused the most prominent changes in the β-sheets surrounding the binding cleft (β2'), whereas BDZ produced its most significant changes in the antiparallel β-sheets, the linker regions between the β-sheets and the α-helices that form the bottom and side walls of the cleft. It seems that the Mdm2 domain has an intrinsic flexibility that enables it to adapt its conformation to

ligands. The conformational changes alter the size of the cleft and were mainly in the linker regions suggesting that the overall dynamic nature of Mdm2 is related to dynamic movements in these regions. NMR spectroscopy studies[96] confirm the observation that changes in chemical shift were mainly in the linker regions between the secondary structure elements and suggests that the overall dynamic nature of Mdm2 is related to dynamic movements in these regions, which can occur without any loss of the secondary structure scaffold. Also the MD studies based on 1T4E X-ray structure showed that in the apo-state the lid exists in a dominant closed form, whereas in bound system exists in a open form in which the lid extends away from the binding cleft, allowing a full access of ligands to the binding pocket. This study concluded that the binding cleft is very adaptable and that different ligands might induce global conformational changes. The opening up of a ligand-binding pocket suggests that compounds inserted at this position might raise further conformational change and hence increase the pocket plasticity.

## 6.3 A3 receptor: homology modeling and 3D-QSAR studies

The A3 protein sequence was collected from the Swiss-Prot Protein Database (P33765).[140,141] The A2A adenosine receptor can be considered the best template for homology modeling according to the percentage identity of the aligned sequence. Identity increases from a comparison of bovine rhodopsin to hA2A adenosine receptor. The increase is even higher when comparing only TM regions and if N- and C-terminus are not taken into consideration (Table 9).

| | | Rhodopsin | hA2A |
|---|---|---|---|
| **All** | hA1 | 15.3 | 38.3 |
| | hA2A | 14.5 | 100 |
| | hA2B | 19.5 | 46.6 |
| | hA3 | **14.4** | **31** |
| **TM regions** | hA1 | 15.7 | 60.8 |
| | hA2A | 23.2 | 100 |
| | hA2B | 22 | 68.6 |
| | hA3 | **19.3** | **50.3** |
| **All except N and C-terminus** | hA1 | 16.4 | 51.7 |
| | hA2A | 21.2 | 100 |
| | hA2B | 22.9 | 61.6 |
| | hA3 | **16** | **41.9** |

**Table 9: Percentage identity of aligned sequences.**

Two different models of hA3R were generated by LOMETS[117] using the bovine rhodopsin (PDB ID: 1U19) and the hA2A adenosine receptor (PDB ID: 3EML) as templates (Fig. 47).

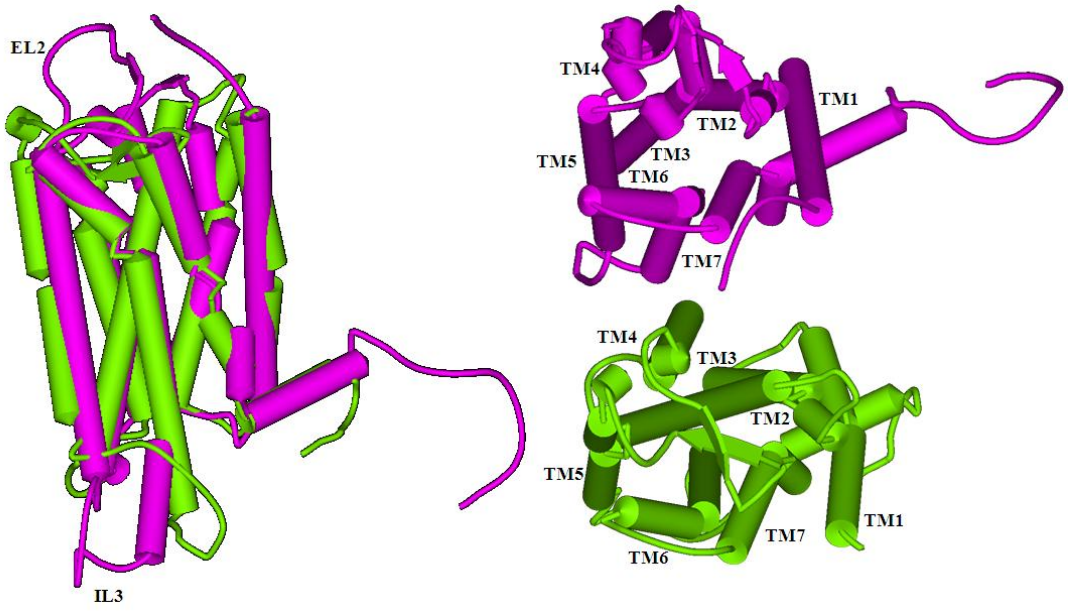



**Fig. 47: Superposition of hA3 models built using rhodopsin (green) and hA2A (purple) as templates.**

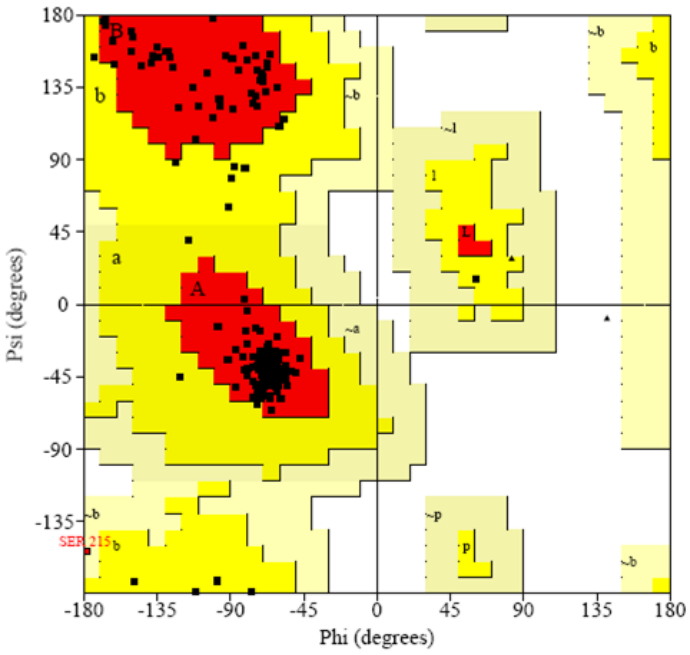
As it is seen from the RMSD of the aligned models, the main differences among the two models are found within EL2 and IL3 loops (Table 10). In particular EL2 belongs to the binding pocket and interacts with ligand. In the model built using hA2A as template, the binding pocket is open to extracellular side and closer to TM6 and TM7.

| | **RMSD** |
|---|---|
| Backbone | 7.08 |
| IL1 | 1.18 |
| IL2 | 3.01 |
| **IL3** | **6.26** |
| EL1 | 3.47 |
| **EL2** | **6.74** |
| EL3 | 1.36 |
| TM1 | 0.70 |
| TM2 | 1.37 |
| TM3 | 0.98 |
| TM4 | 1.39 |
| TM5 | 1.03 |
| TM6 | 0.75 |
| TM7 | 0.93 |

**Table 10: RMSD of the aligned hA3 adenosine receptor models built using A2A and rhodopsin.**

Then new structure of hA2A adenosine receptor improved modelization of A3 adenosine receptor. Obtained model was refined using MoodLoop,[142] and was ranked on QMEAN

server.[143] The model created using template 3EML had Qmean score = 0.47, indicating good quality of model produced, while the one generated using template 1U19 had Qmean score = 0.40. The overall stereochemical quality of the model was assessed by PROCHECK[144] and the validation of the structure was performed by inspecting the psi/phi Ramachandran plot (Fig. 48).




**Plot Statistics**

| | | |
|---|---|---|
| Residues in most favoured regions [A,B,L] | 279 | 94.3% |
| Residue in additional allowed regions [a,b,l,p] | 16 | 5.4% |
| Residue in generously allowed regions [~a,~b,~l,~p] | 1 | 0.3% |
| Residue in disallowed regions | 0 | 0.0% |
| Number of non-glycine and non-proline residues | 296 | 100.0% |
| Number of end-residues (excl. Gly and Pro) | 2 | |
| Number of glycine residues (shown as triangles) | 11 | |
| Number of Proline residues | 9 | |
| Total number of residues | 318 | |

**Fig. 48: Ramachandran plot of obtained model.**

The Ramachandran plot showed 94.3% of the residues in the most favorable region, 5.4% in the allowed region, 0.3% in the generously allowed region. This result revealed that the obtained model is reliable and of good quality.

Homology models represent a rigid conformation of a protein, but proteins are dynamic and show rapid, small-scale structural fluctuation.[145] Obtained model supposed to be antagonist-like state of A3 receptor and it can exist in more than one conformational state. Therefore a 20 ns molecular dynamics simulation in a lipid bilayer was performed to explain conformational flexibility and structural stability, using the program DESMOND.[133]

The structural changes and dynamic behavior were analyzed by calculating the RMSD of the TMs (Fig. 49) and loops (Fig. 50) backbone in function of time (ns).

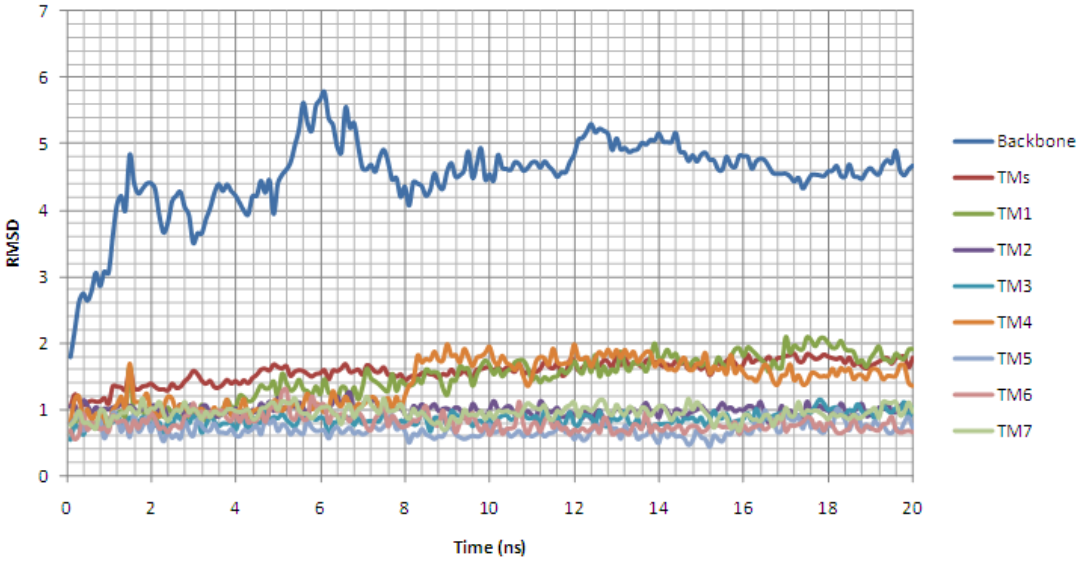



**Fig. 49: RMSD of the TMs backbone in function of time.**

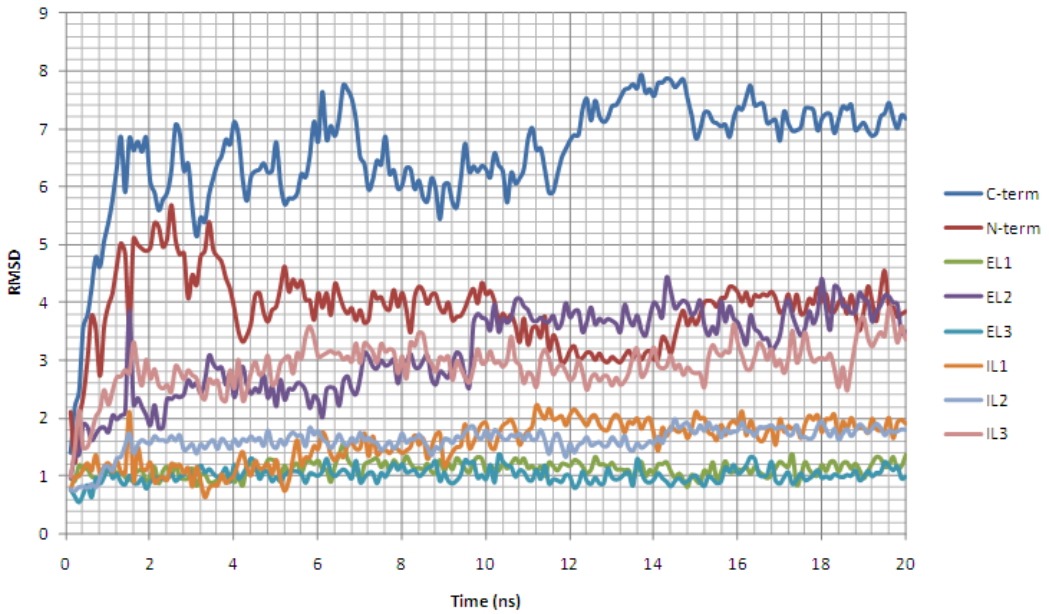



**Fig. 50: RMSD of the loops backbone in function of time.**

Very short loop, such as IL1, EL1, and EL3, have low values of RMSD, together with bigger loop like IL2. N-term and C-term are very flexible, probably due to the fact that are more exposed. Bigger loop like IL3 and EL2 are more flexible (Fig. 51). TMs show that the conformation is more stable through 20 ns of simulation. The RMSD values of the backbone atoms in the system tend to converge after 7 ns, showing fluctuations of around 1 Å. The low

RMSD value during the simulation time indicated that the 3D structural model represents a stable folding conformation.

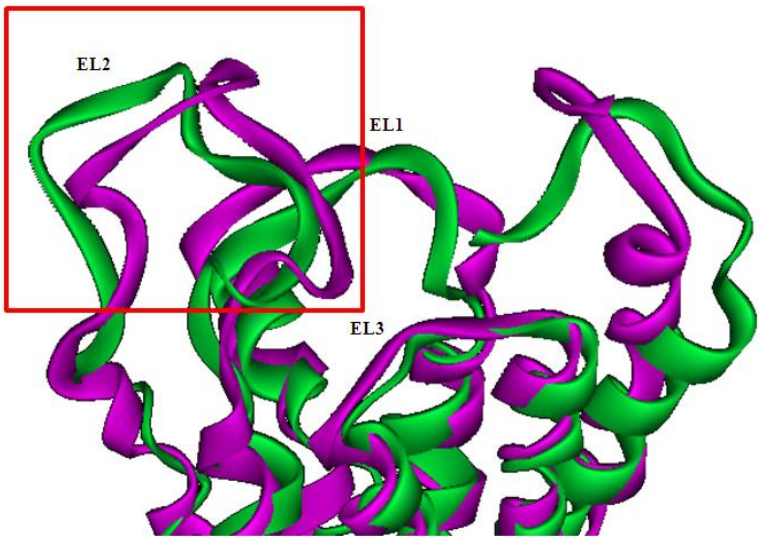



**Fig. 51: EL2 of the hA3 adenosine receptor before (purple) and after (green) 20ns of molecular dynamics.**

A number of activity data for A3 adenosine receptor inhibitors, that belong to diverse chemical classes, are currently avalaible. Biological activity data of 121 hA3R antagonists, represented as pKi, were collected from different literature papers.[146–152] All antagonist structures were docked into TM binding site of the hA3 model using GLIDE[124] (Fig. 52).

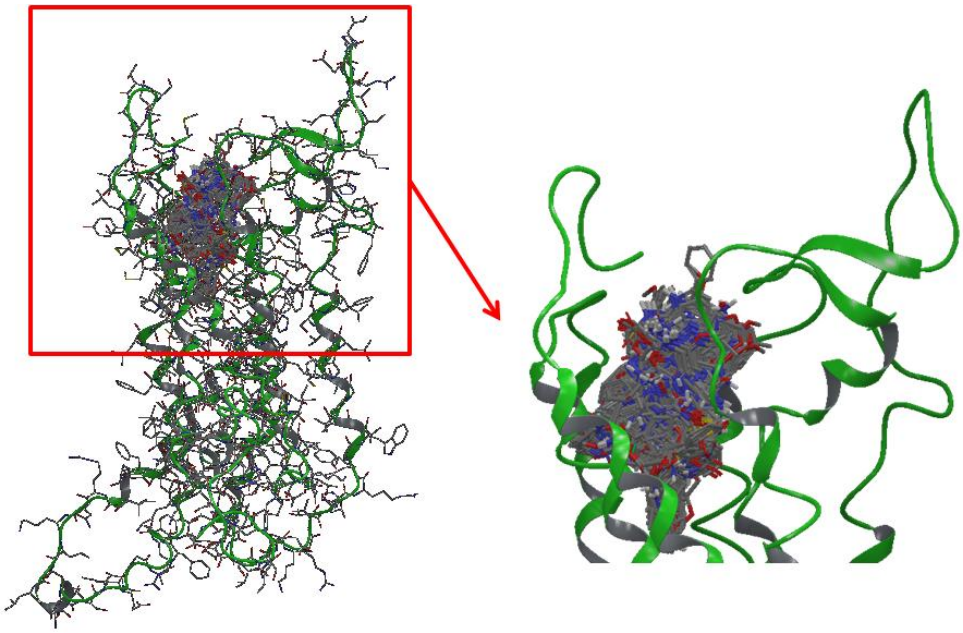


**Fig. 52: Antagonists structures docked into TM binding site of hA3.**

The resulting docked complexes were imported in PHASE module[153] for the development of 3D Quantitative Structure–Activity Relationship (QSAR) Pharmacophore model. All reported

A3 receptor antagonists have been clustered into four groups: triazolo-quinoxaline, (Group 1) (Fig. 53a); pyrazolo-quinoline, (Group 2) (Fig. 53b,c); annelated triazolo-pyrazine, (Group 3) (Fig. 53d-g); other heterocyclic derivatives, (Group 4) (Fig. 53h-n). The prepared ligands were then used to generate common feature pharmacophore models, which, in turn, were utilized to generate 3D-QSAR models.

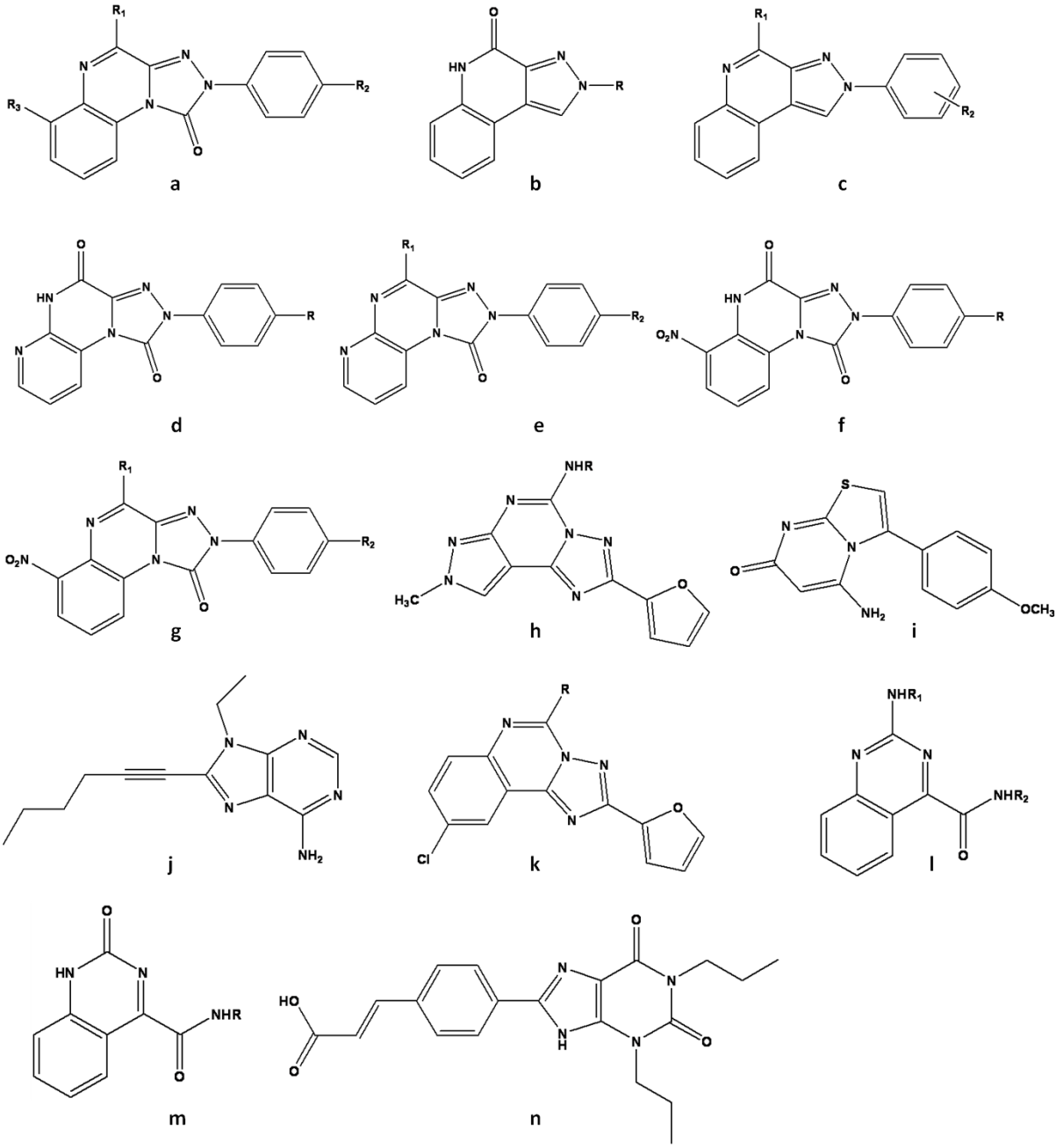



**Fig. 53: A3 adenosine receptor inhibitors: triazolo-quinoxaline derivatives (a), pyrazolo-quinoline derivatives (b-c), annelated triazolo-pyrazine derivatives (d-g), heterocyclic derivatives (h-n).**

The datasets of compounds were divided into training and test sets. Two different approaches for splitting the datasets were used. In the first one, training sets were constructed by choosing a percentage (80%) of the total number of compounds in each bin

randomly. The other approach is based on Kohonen map-artificial neural network, now more widely called self organizing maps (SOM).[154,155] Descriptors calculation, and autoscaling of descriptors matrix were the starting point to perform Kohonen clustering approach. The selected test set members are characterized by the minimal distance from the centroid of each cell in the top map. Common pharmacophore features were identified and scored. The generated pharmacophore hypotheses for these groups consisted of four features: one acceptor site (A), and three aromatic ring sites (R). The alignments of all the ligands to the top-ranked pharmacophores were used to generate QSAR models, setting the number of partial least square (PLS) factors from 1 to 6 (Fig. 54).

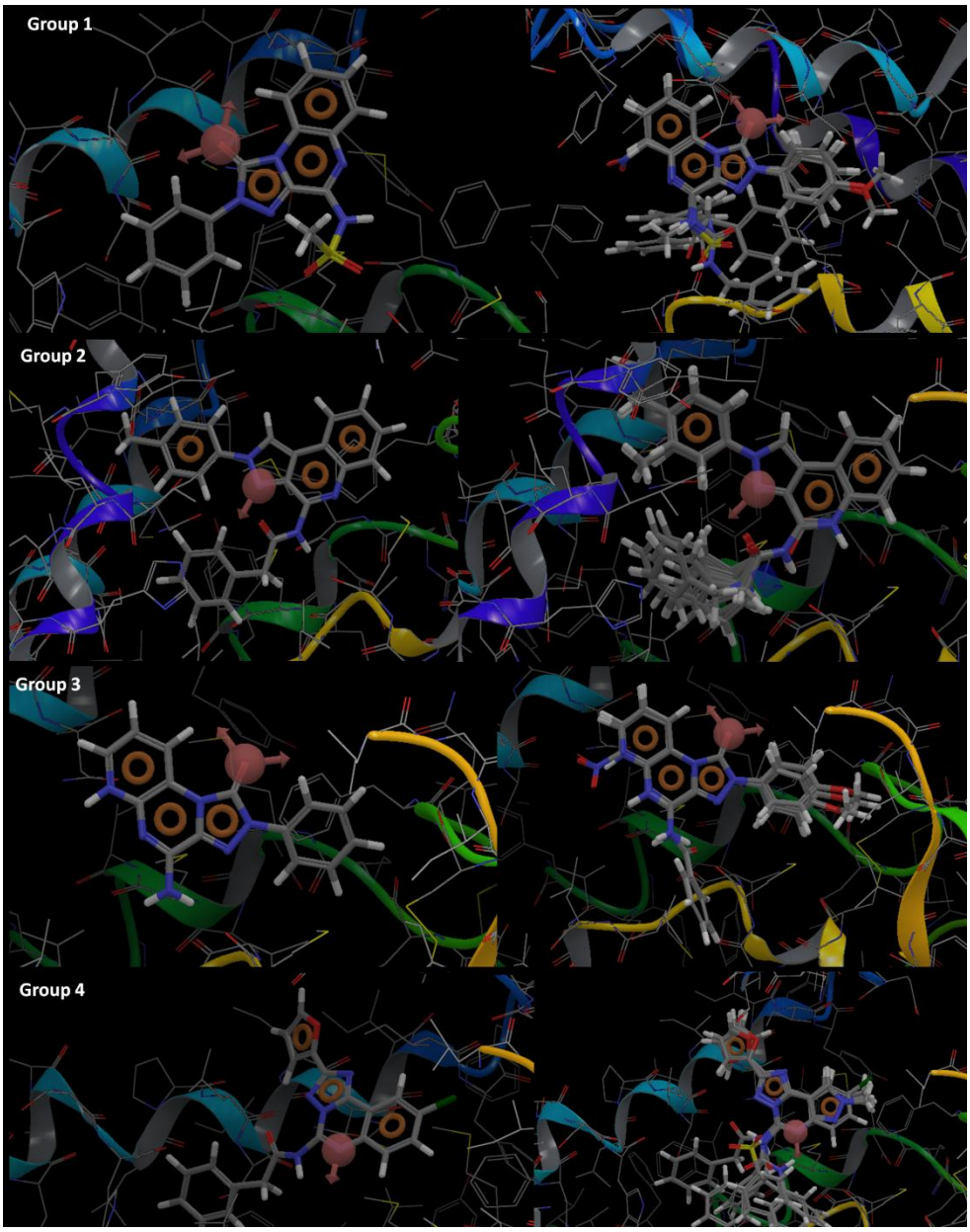



**Fig. 54: Pharmacophore mapping of the most active compounds (left); superimposition of active compounds with the pharmacophore.**

Predicted activity of test set compounds was plotted against their experimental activity, and the relevant statistics were computed. Phase statistical analysis for the four groups, labelled as Groups 1-4, for each of test set selection method is shown in Tables 11 and 12.

**Group 1**

| Factors | SD | $R^2$ | RMSE | $R^2_{pred}$ | R-Pearson | RMSE/SD | Δ | Opt. Model |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.62 | 0.66 | 0.52 | 0.19 | 0.61 | 0.84 | | - |
| 2 | 0.46 | 0.82 | 0.65 | 0.64 | 0.85 | 1.15 | 0.30 | - |
| 3 | 0.31 | 0.92 | 0.38 | 0.57 | 0.76 | 1.21 | 0.06 | √ |
| 4 | 0.20 | 0.97 | 0.44 | 0.42 | 0.70 | 2.17 | 0.96 | - |
| 5 | 0.13 | 0.99 | 0.38 | 0.57 | 0.79 | 2.97 | 0.80 | - |
| 6 | 0.08 | 0.99 | 0.38 | 0.57 | 0.78 | 4.52 | 1.55 | - |

**Group 2**

| Factors | SD | $R^2$ | RMSE | $R^2_{pred}$ | R-Pearson | RMSE/SD | Δ | Opt. Model |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.56 | 0.41 | 0.32 | 0.69 | 0.91 | 0.57 | | - |
| 2 | 0.33 | 0.80 | 0.34 | 0.65 | 0.85 | 1.01 | 0.44 | - |
| 3 | 0.26 | 0.88 | 0.32 | 0.69 | 0.89 | 1.22 | 0.21 | - |
| 4 | 0.21 | 0.93 | 0.33 | 0.67 | 0.88 | 1.58 | 0.36 | √ |
| 5 | 0.15 | 0.96 | 0.29 | 0.75 | 0.95 | 1.87 | 0.29 | - |

**Group 3**

| Factors | SD | $R^2$ | RMSE | $R^2_{pred}$ | R-Pearson | RMSE/SD | Δ | Opt. Model |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.61 | 0.48 | 0.74 | 0.10 | 0.64 | 1.21 | | - |
| 2 | 0.46 | 0.72 | 0.58 | 0.45 | 0.83 | 1.26 | 0.04 | - |
| 3 | 0.35 | 0.84 | 0.46 | 0.65 | 0.88 | 1.31 | 0.05 | - |
| 4 | 0.29 | 0.90 | 0.45 | 0.67 | 0.91 | 1.56 | 0.25 | √ |

**Group 4**

| Factors | SD | $R^2$ | RMSE | $R^2_{pred}$ | R-Pearson | RMSE/SD | Δ | Opt. Model |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.44 | 0.69 | 1.03 | 0.31 | 0.96 | 2.34 | | - |
| 2 | 0.26 | 0.90 | 0.00 | 0.34 | 0.96 | 0.01 | -2.33 | √ |
| 3 | 0.16 | 0.96 | 0.98 | 0.37 | 0.97 | 6.01 | 6.00 | - |

**Table 11: 3D-QSAR Results Summary for random selection of test set (Factors =number of factors in the partial least squares regression model; SD = standard deviation; RMSE = root-mean-square error; $R^2_{pred}$ = value of the predicted activities for test set; R-Pearson = Pearson R value for the correlation between the predicted and observed activity for the test set; Δ = difference from preceding).**

**Group 1**

| Factors | SD | $R^2$ | RMSE | $R^2_{pred}$ | R-Pearson | RMSE/SD | Δ | Opt. Model |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.55 | 0.71 | 0.99 | -0.85 | -0.08 | 1.80 | - | - |
| 2 | 0.38 | 0.86 | 1.08 | -1.24 | -0.20 | 2.83 | 1.02 | - |
| 3 | 0.25 | 0.94 | 1.02 | -1.00 | -0.18 | 4.10 | 1.27 | √ |
| 4 | 0.13 | 0.99 | 1.12 | -1.37 | -0.23 | 8.70 | 4.60 | - |
| 5 | 0.08 | 0.99 | 1.09 | -1.28 | -0.19 | 14.24 | 5.54 | - |
| 6 | 0.06 | 1.00 | 1.09 | -1.27 | -0.20 | 18.36 | 4.12 | - |

**Group 2**

| Factors | SD | $R^2$ | RMSE | $R^2_{pred}$ | R-Pearson | RMSE/SD | Δ | Opt. Model |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.56 | 0.42 | 0.47 | 0.33 | 0.60 | 0.83 | - | - |
| 2 | 0.38 | 0.75 | 0.44 | 0.40 | 0.70 | 1.16 | 0.33 | - |
| 3 | 0.28 | 0.87 | 0.28 | 0.75 | 0.89 | 1.02 | -0.14 | - |
| 4 | 0.23 | 0.91 | 0.27 | 0.77 | 0.89 | 1.17 | 0.15 | √ |
| 5 | 0.18 | 0.95 | 0.25 | 0.81 | 0.90 | 1.39 | 0.22 | - |

**Group 3**

| Factors | SD | $R^2$ | RMSE | $R^2_{pred}$ | R-Pearson | RMSE/SD | Δ | Opt. Model |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.63 | 0.46 | 0.94 | -0.19 | -0.26 | 1.48 | - | - |
| 2 | 0.40 | 0.80 | 0.83 | -0.18 | -0.07 | 2.10 | 0.62 | - |
| 3 | 0.29 | 0.90 | 0.75 | 0.03 | 0.25 | 2.57 | 0.47 | - |
| 4 | 0.19 | 0.96 | 0.84 | -0.20 | 0.08 | 4.51 | 1.94 | √ |

**Group 4**

| Factors | SD | $R^2$ | RMSE | $R^2_{pred}$ | R-Pearson | RMSE/SD | Δ | Opt. Model |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.40 | 0.82 | 0.62 | 0.09 | 0.59 | 1.55 | - | - |
| 2 | 0.21 | 0.95 | 0.69 | -0.12 | 0.54 | 3.27 | 1.72 | √ |

**Table 12: 3D-QSAR Results Summary for SOM selection of test set (Factors =number of factors in the partial least squares regression model; SD = standard deviation; RMSE = root-mean-square error; $R^2_{pred}$ = value of the predicted activities for test set; R-Pearson = Pearson R value for the correlation between the predicted and observed activity for the test set; Δ = difference from preceding).**

A statistical analysis which included the $R^2$ *versus* RMSE/SD plot was employed to choose the best PLS model for the different set selection methods. The best model was chosen on the basis of PLS factor model minimum observed in RMSE/SD value, with $R^2$ value still higher than 0.9. Only models with good statistical parameters for the training set and exhibiting good predictive ability against a test set were chosen. Group 4 did not exhibit excellent statistical prediction in both of test set selection method was not taken into account.

Group 2 exhibited comparatively better PLS statistical qualities and excellent prediction of the external test set compounds in both of test set selection method; it showed correlation coefficient ($R^2$) of 0.96 for random selection of test set, test set prediction ($R^2_{pred}$) of 0.75 and R-Pearson of 0.95. The low standard deviation (SD) and root-mean-squared error (RMSE) contributes significantly to the model.

For Groups 1-3, the most significant models were obtained with Random selection approach, because $R^2_{pred}$ for SOM resulted as negative value. Visualizing the hypotheses and various ligands in the context of the 3D-QSAR models, the Fig. 55 illustrates the most significant favorable and unfavorable interactions that arise when the 3D-QSAR model is applied to the reference ligand, chosen as the most active compound in the training set. In these representations, the blue cubes indicate favorable regions while red cubes indicate unfavorable ones for biological activity. The blue regions in the vicinity of 3′-position of the 4'-methyl-phenyl ring of Group 2 suggested that the substitution in this area may enhance the activity, whereas red regions surround the *para*-position.

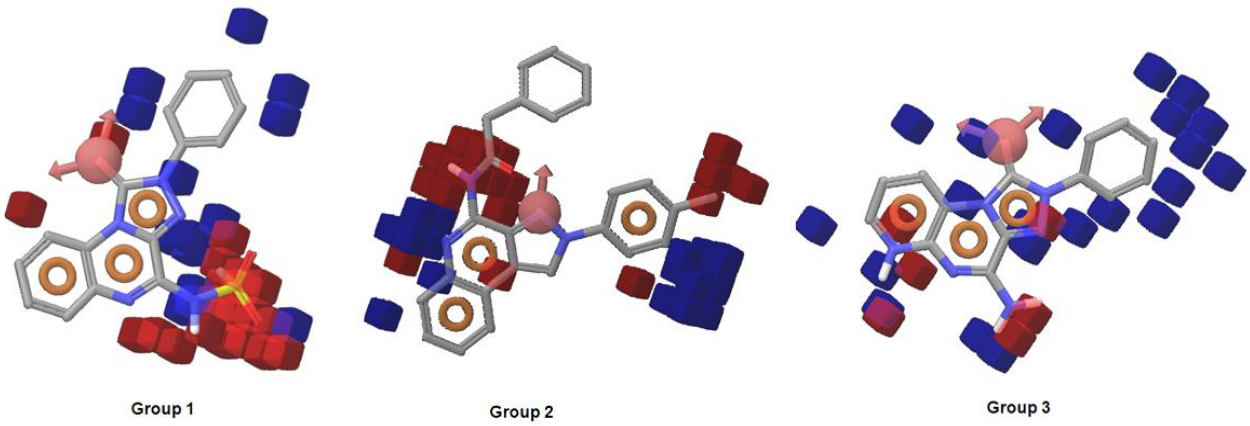



**Fig. 55: Most significant favorable and unfavorable interactions in QSAR models.**

The scatter plot for the training (a) and test set (b) indicates a reasonably good correlation between the predicted and experimental activities (Fig.56).

The best obtained pharmacophore models were used to retrieve new potential inhibitors from "lead like" ZINC database (3,027,619 compounds) using virtual parallel screening.

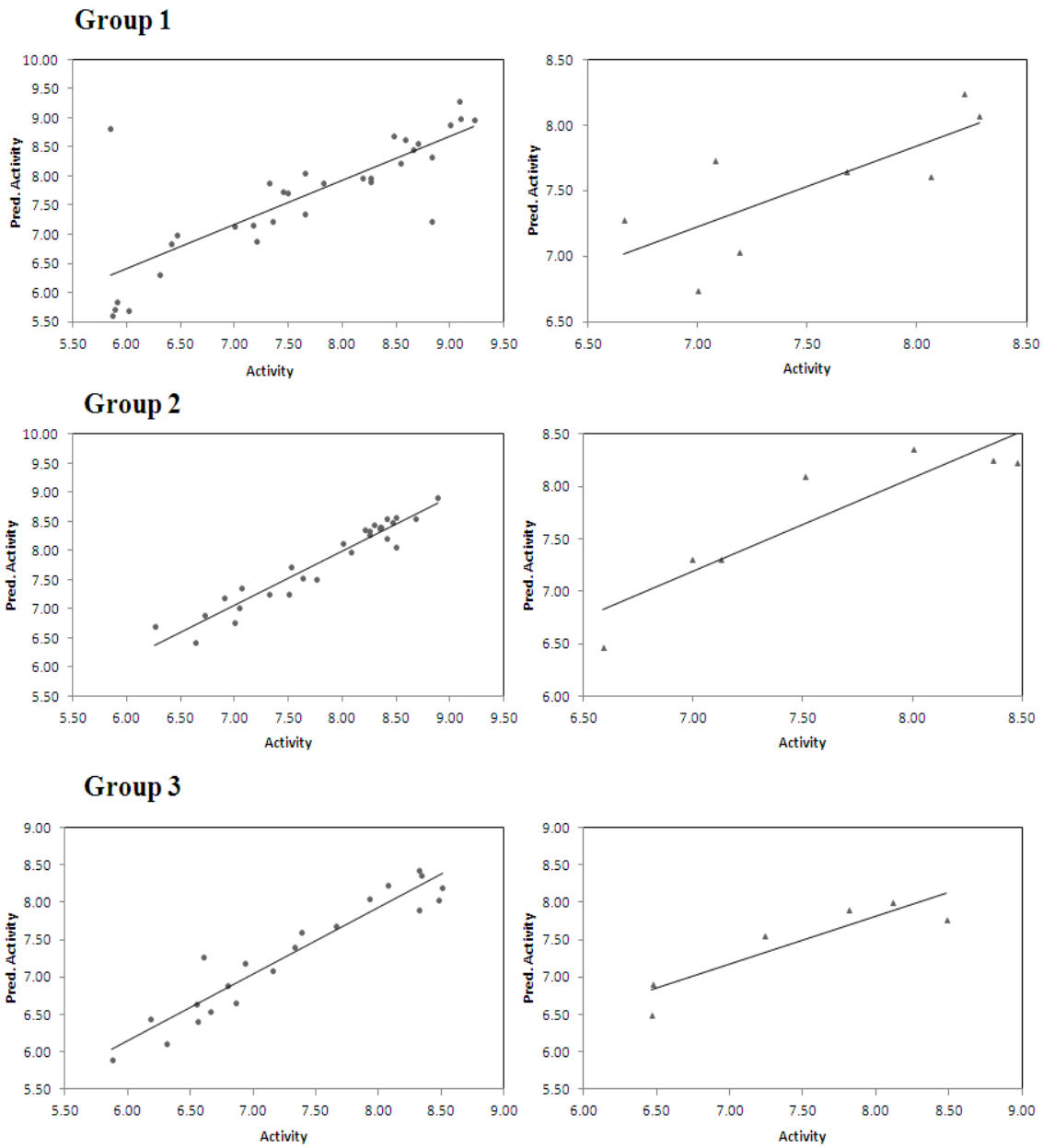



**Fig. 56: Scatter plots for the hA3R QSAR models applied to the training set (a) and the test set (b). Line indicates the hypothetical "best fit" line between the predicted and experimental values.**

Screening molecules were required to match all the hypotheses features. Database hits were ranked in order of their Fitness score, a measure of how well the aligned ligand conformer matches the hypothesis based on site matching, vector alignments and volume terms. As a result, 18189 potential ligand hits, that match with all the hypotheses, were identified. Here, we report the ten compounds with the best Fitness score (Table 13); the chemical structures of these hit molecules are presented in Fig. 57.

| | Model Group 1 | | Model Group 2 | | Model Group 3 | |
|---|---|---|---|---|---|---|
| **ZINC ID** | **FITNESS** | **PRED pKi** | **FITNESS** | **PRED pKi** | **FITNESS** | **PRED pKi** |
| 978790 | 2.268 | 7.75 | 1.439 | 7.09 | 1.495 | 6.89 |
| 3192107 | 2.455 | 7.59 | 1.802 | 7.44 | 2.514 | 7.25 |
| 6624316 | 2.458 | 7.80 | 1.695 | 8.00 | 2.536 | 7.37 |
| 6624541 | 2.459 | 7.77 | 1.696 | 8.04 | 2.540 | 7.53 |
| 8578947 | 2.460 | 7.63 | 1.719 | 7.25 | 2.463 | 7.07 |
| 13363145 | 2.473 | 7.67 | 1.705 | 7.37 | 2.423 | 7.29 |
| 20085946 | 2.493 | 7.60 | 1.696 | 7.41 | 2.577 | 7.03 |
| 21774670 | 2.463 | 8.24 | 1.574 | 7.51 | 2.476 | 7.42 |
| 21774675 | 2.447 | 7.84 | 1.540 | 7.73 | 2.507 | 7.47 |
| 21774724 | 2.468 | 7.66 | 1.551 | 7.48 | 2.515 | 7.33 |

**Table 13: Fitness score of the identified lead**

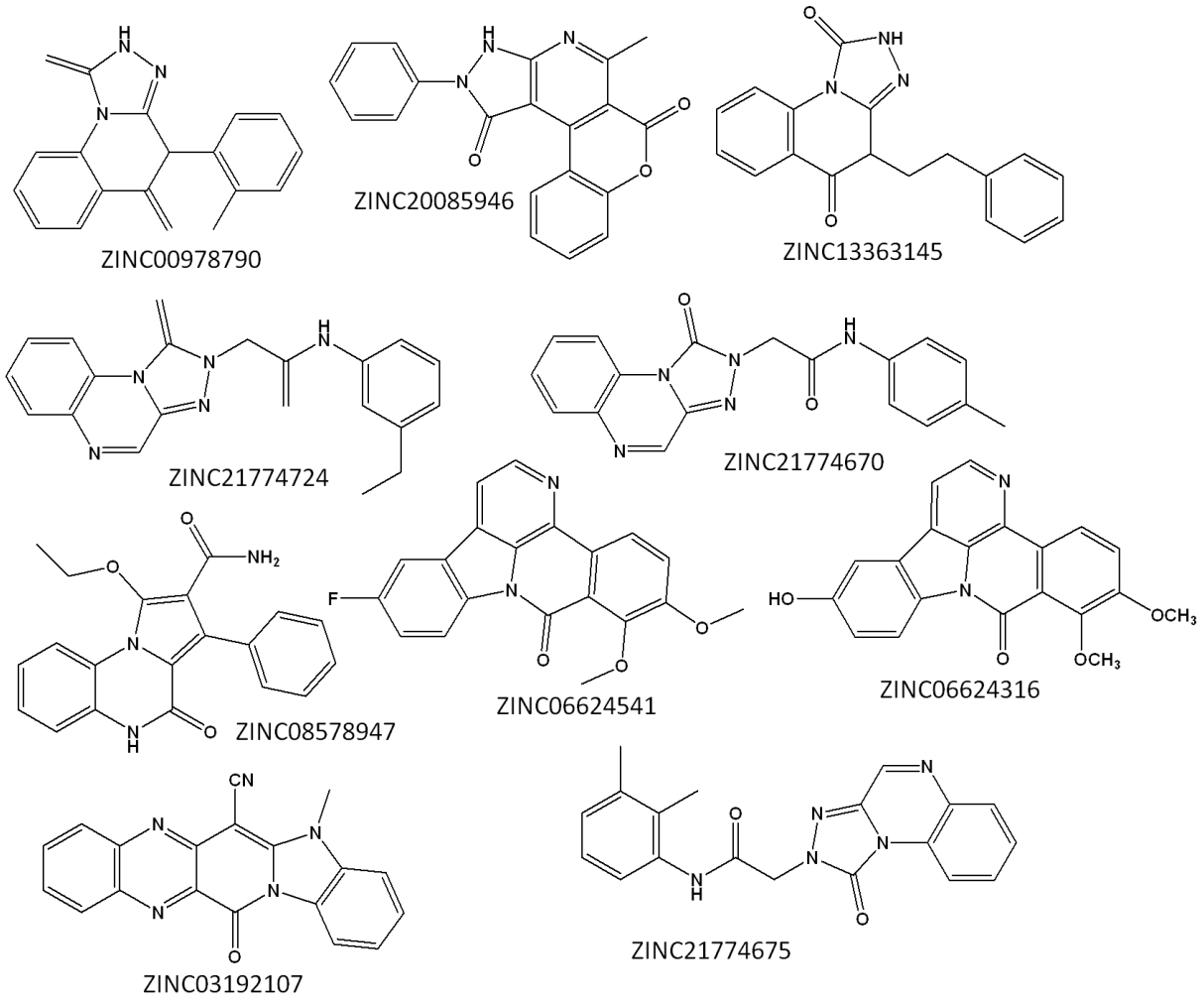



**Fig. 57: Chemical structures of ten compounds with the best Fitness score.**

Therefore, these selected compounds were docked into the binding site of the modeled protein. The conformations of these hits bound to the modeled A3R were also analyzed to determine hydrogen-bonding and hydrophobic interactions. The docking results for the final hit molecules are given in Table 14, whereas the binding modes (structures and interacting residues) are shown in Fig. 58.

| ZINC ID | FITNESS | XP GSCORE | AMINO ACIDS INVOLVED IN HYDROPHOBIC INTERACTIONS | AMINO ACIDS INVOLVED IN H-BOND |
|---|---|---|---|---|
| 978790 | 2.518 | -6.549 | V72, F168, M177, L264, I268 | M86 |
| 3192107 | 2.455 | -6.606 | V72, M86, F168, I253, V263, L264, I268 | |
| 6624316 | 2.458 | -7.434 | V72, L89, L90, L91, M177, L246, I253, V259, I263, L264, I268 | T94 |
| 6624541 | 2.459 | -7.39 | V72, M86, F168, M177, L246, I253, V263, L264 | D250 |
| 8578947 | 2.46 | -7.298 | V72, M86, F168, I253, V263, L264 | S73 |
| 13363145 | 2.473 | -6.927 | V72, L90, L91, M177, L246, I253, V259, L264, I268 | M86, L90 |
| 20085946 | 2.493 | -7.415 | V72, M86, L90, F168, M177, V259, L264, I268 | M86 |
| 21774670 | 2.463 | -8.427 | V72, M86, L90, L91, P168, M177, L246, I268 | M86, D250 |
| 21774675 | 2.447 | -8.461 | V72, M86, L90, L91, F168, M177, L246, I268 | M86 |
| 21774724 | 2.468 | -8.983 | V72, M86, L89, L90, L91, C166, F168, M177, L246, I268 | M86 |

**Table 14: Extra Precision (XP) Glide results for the ten lead molecules.**

The docked models indicate that the ligands maintain key interactions with the TM2, TM3, TM6 and TM7. Another important interaction with EL2 seems to be fundamental.

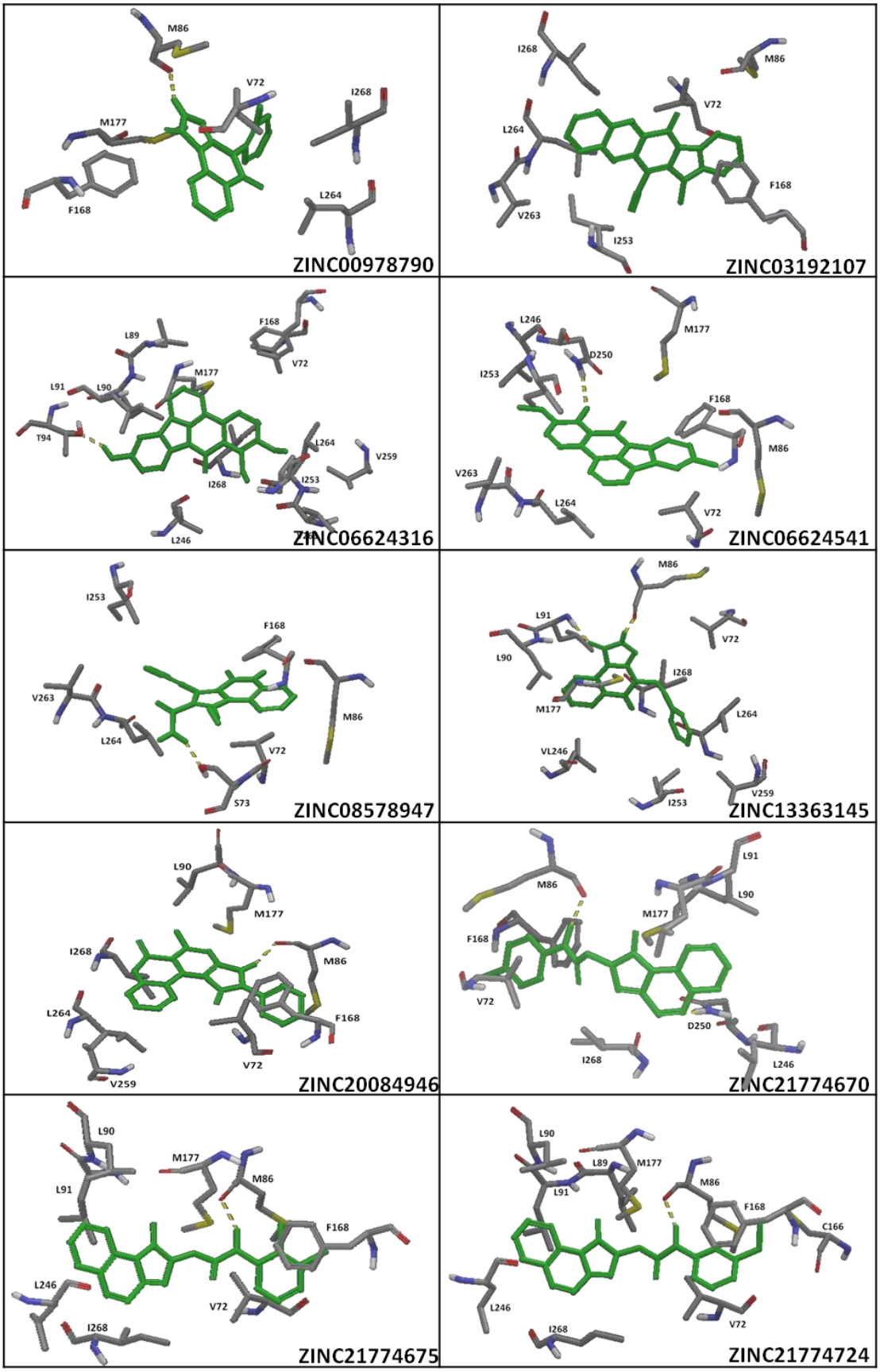



**Fig. 58: Binding poses of the ten lead molecules. Hydrogen bonds are shown as dotted yellow lines.**

The drug-like character of the lead compounds was assessed by evaluating their physicochemical properties using QikProp.[156] Their molecular weights were <500 Da; they had <5 hydrogen bond donors and <10 hydrogen bond acceptors, and logP values of <5 (Table 15). These properties are all well within the acceptable range of Lipinski's rule of five. The pharmacokinetic parameters of the molecules (ADMET) showed that the partition coefficient ($QPlogP_{o/w}$) and water solubility (QPlogS) values, which are crucial for estimating the absorption and distribution of drugs within the body, ranged approximately from 1.80 to 3.32 and −5.08 to −3.56, respectively.

| ZINC ID | MW | HBD | HBA | $QPlogP_{o/w}$ | QPlogS | Percent Human Oral Absorption | QPP Caco | QPP MDCK |
|---|---|---|---|---|---|---|---|---|
| 978790 | 292.29 | 1 | 6.00 | 1.80 | -3.62 | 85.55 | 483.06 | 225.31 |
| 3192107 | 325.32 | 0 | 6.50 | 2.07 | -4.25 | 89.99 | 697.98 | 335.40 |
| 6624316 | 346.34 | 1 | 6.25 | 2.56 | -3.58 | 96.61 | 1128.89 | 563.97 |
| 6624541 | 348.33 | 0 | 5.50 | 3.32 | -3.89 | 100.00 | 3874.95 | 3860.24 |
| 8578947 | 347.37 | 3 | 5.00 | 2.80 | -4.51 | 90.16 | 412.81 | 190.12 |
| 13363145 | 306.32 | 1 | 6.00 | 2.16 | -3.56 | 87.81 | 493.54 | 230.60 |
| 20085946 | 343.34 | 1 | 7.50 | 2.00 | -3.81 | 90.84 | 822.44 | 400.49 |
| 21774670 | 333.34 | 1 | 6.50 | 2.68 | -4.54 | 94.12 | 751.26 | 363.16 |
| 21774675 | 347.37 | 1 | 6.50 | 3.06 | -5.08 | 96.97 | 814.51 | 396.31 |
| 21774724 | 347.37 | 1 | 6.50 | 2.99 | -4.83 | 95.97 | 754.89 | 365.05 |

**Table 15: QikProp properties of the identified hits (MW = Molecular weight; HBD = Hydrogen bond donors by solute to water molecule; HBA = Hydrogen bond acceptors by solute from water molecule; $QPlogP_{o/w}$ = Predicted octanol/water partition coefficient logP; QPlogS = Predicted aqueous solubility S in mol/L; QPPCaco = Predicted apparent Caco-2 cell permeability in nm/sec; QPPMDC = Predicted apparent MDCK cell permeability in nm/sec)**

Thus, the pharmacokinetic parameters are well within the acceptable range defined for human use, thereby indicating the potential drug-likeness of these molecules. The structure selected by our proposed screening procedure could serve as potential specific inhibitors against the A3 human receptor and therefore can be proposed for therapeutic treatments of cancer.

# 7. Conclusions

In conclusion hypoxia plays an important role in tumor progression and metastasis. Tumor growth leads to hypoxia, which in turn results in decreased therapeutic efficacy, stimulation of angiogenesis, and tumor progression. Overexpression of HIF-1α has been observed in many common human cancers.

Currently, no empirical (X-ray, crystallographic) 3D structure for HIF-1α is available. In view of this, a protein model was constructed using homology modeling. Because the sequence identity between templates and the target is very low there may be significant errors. Regions of the model that were constructed without a template, by loop modeling, are generally much less accurate than the rest of the model.

Several approaches have been used to inhibit HIF-1α expression or activity: inhibition of proteins that modulate HIF-1 activity, signal transduction pathways involved in HIF-1α activation. Selective inhibitors target protein-protein interaction, protein-DNA binding, and transcriptional activity, whereas non-selective ones affect downstream signaling and other indirect pathways. Several domains of HIF-1 might be suitable targets for the development of selective HIF-1 inhibitors: domains involved in the recruitment of coactivators, which are required for maximal transcriptional activity, or binding to DNA, which is dependent on sequence specificity.

Domains of HIF-1α that mediate specific functions, such as DNA binding as well as transcriptional activity (C-TAD) might lead to the identification of more “selective” inhibitors. The identification of small molecules that inhibit the sequence-specific binding of transcription factors to DNA is an attractive approach for modulating HIF-1-dependent gene expression. The docking studies reported herein provided reliable information on the capability of new ligands to interact with DNA sequence.

Another approach has been to interfere with the interaction between the C-TAD of HIF-1α and the CH1 domain of p300.

The structure of the complex of the HIF-1α C-TAD with the CH1 domain of p300 is of interest as a potential target for design of antitumor agents. A docking study was performed on CH1 domain with the purpose of identifing new potential inhibitors for this protein.

Furthermore HIF-1α has been implicated to be involved in p53 stabilization. Mdm2 may act as a bridge and mediate the indirect interaction between HIF-1α and p53 in cells.

Protein flexibility is an essential property of biomolecules. Thus, knowledge about flexibility is important for the rational design of new drugs. Although there is an abundance of literature available for Mdm2, no information is available concerning the lid behavior in apo-Mdm2. In this study, we used 35 ns MD simulations in order to analyze the different conformational aspects that regulate apo form and the binding of ligands in Mdm2. Starting from a X-ray structure of human Mdm2 presenting the lid residue, we simulated the behaviour in the presence of different inhibitors. To study the influence of ligands binding on

the Mdm2 dynamics, backbone and librational variations for bound and apo-Mdm2 were analyzed. Our findings suggest that the structure of the binding domain of Mdm2 adapts itself to the ligands. Also the N-terminal lid of Mdm2 slowly interconverts between a closed state and an open state which is highly flexible. While apo-Mdm2 predominantly populates the closed state, the binding with inhibitors shifts the equilibrium toward the open state.

As expected different ligands could elicit different changes in structure and thus mediate a variety of biological functions. The *in silico* tools are probably able to simulate the behaviour of Mdm2 in the presence of any kind of ligand. A better understanding of these complex effects was achieved by using these computational techniques, which revealed the dynamic behavior of the protein-ligand complexes and the contribution of N-terminal domain flexibility, that has to be taken into account during the design of new inhibitors.

Levels of both HIF-1 and adenosine are elevated within the hypoxic environment of solid tumors. In fact under hypoxic conditions, adenosine upregulates HIF-1α protein expression exclusively through the A3 receptor subtype. The new structure of hA2A solved in 2008 provided a new starting point for homology modeling and was used as template to built homology model of hA3 adenosine receptor. The inclusion of an explicitly lipid bilayer into the energy minimisation simulation may have helped in optimising the quality of the structure. Our theoretical model of hA3 adenosine receptor has been used to evaluate and quantify the structure-activity relationship of known antagonists. Finally, the 3D-QSAR model has been used with the purpose of identifying new potential inhibitors for A3R. The identified hits seem to be both potent and specific for the desired target receptor.

# 8. Methods

**I-TASSER**

I-TASSER[157] is an automated pipeline for protein tertiary structure prediction using multiple threading alignments and iterative structure assembly simulations. The target sequences are first threaded through a representative PDB structure library (with a pair-wise sequence identity cut-off of 70%) to search for the possible folds by four simple variants of Profile-Profile threading Alignment (PPA) methods. The continuous fragments are then excised from the threading aligned regions which are used to reassemble full-length models while the threading unaligned regions (mainly loops) are built by *ab initio* modeling.[158]

The conformational space is searched by replica-exchange Monte Carlo simulations.[159] The structure trajectories are clustered by SPICKER[160] and the cluster centroids are obtained by averaging the coordinates of all clustered structures. To rule out the steric clashes on the centroid structures and to refine the models further, we implement the fragment assembly simulation again, which starts from the cluster centroid of the first round simulation. Spatial restraints are extracted from the centroids and the PDB structures searched by the structure alignment program TM-align,[161] which are used to guide the second round simulation. Finally, the structure decoys are clustered and the lowest energy structure in each cluster is selected, which has the $C_\alpha$ atoms and the side-chain centers of mass specified. Pulchra[162] is used to add backbone atoms (N, C, O) and Scwrl_3.0[163] to build side-chain rotamers. If any region with >80 residues has no aligned residues in at least two strong PPA alignments of Z-score > $Z_0$, the target will be judged as a multiple domain protein and domain boundaries are automatically assigned based on the borders of the large gaps. The final full-length models are generated by docking the model of domains together (Fig. 59). The domain docking is performed by a quick Metropolis Monte Carlo simulation where the energy is defined as the RMSD of domain models to the full-chain model plus the reciprocal of the number of steric clashes between domains. The goal of the docking is to find the domain orientation that is closest to the I-TASSER full-chain model but has the minimum steric clashes. This procedure does not influence the multiple domain proteins which have all domains completely aligned by the PPAs.

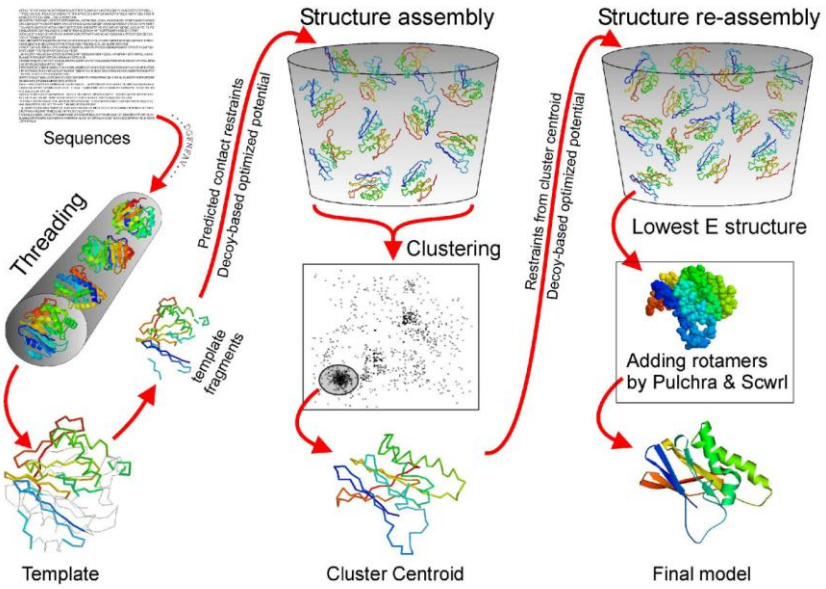



**Fig. 59: I-TASSER protocol for protein structure and function prediction.**

## GROMACS

The GROMACS 3.1.2 package[164,165] was used to perform molecular dynamics simulations, using the force-field parameter set 43A1. This package is a collection of programs and libraries for the MD simulations and subsequent analysis of trajectory data. The ligand was removed from the protein and 35 ns MD simulations of the free receptor were first performed, then 35 ns MD simulations of the Mdm2 complexed with the inhibitors were carried out. The systems were embedded in a water box, the simple point charge (SPC) water model, with margin of 9 Å between the protein and the boundaries of the periodic box. Chlorine counterions are added to produce a neutral charge on the system, and the simulations were performed in constant NPT ensemble. Lennard–Jones potentials were used to model the guest–host and guest–guest interactions, force calculations were truncated at a distance of 1.4 nm. For the calculation of long-range electrostatic forces, the particle-mesh Ewald (PME) method was used, with coulomb cut-off set to 0.9 nm. The temperature was fixed at 300 K using the Berendsen thermostat and the Berendsen pressure coupling algorithm was used to keep the pressure constant at 300 bar. A steepest-descent energy minimization of the systems was first performed to relax the solute-solvent contacts and in order to remove bad van der Waals contacts. The second step consisted in position-restrained MD, restraining the atom positions of the macromolecule while letting the solvent move in the simulation, to soak the water molecule into the protein. Finally, the third step consisted of the MD simulation. The trajectory files were analysed by using GROMACS utilities. Moreover VEGA[166] and VMD[167] programs were employed for trajectory analysis and to manipulate the simulation snapshot structures.

**PRINCIPLE COMPONENT ANALYSIS**

For Principle Component Analysis (PCA) a covariance matrix was constructed using the coordinates of all atoms from each of the 35 ns trajectories. The diagonalization of the covariance matrix generates a diagonal matrix of eigenvalues and a transformation matrix comprising eigenmodes. The modules *g_covar* and *g_anaeig* in GROMACS were used for the PCA calculation. The covariance matrix of the backbone atoms to a reference structure is diagonalized and the eigenvalues and eigenvectors are calculated. The eigenvectors of the covariance matrix represent the movements that have the largest impact on the overall movement (therefore principle components). The trajectories of the three systems were then projected on the first two eigenvectors, to give 2-D projections of the phase space. The eigenvalues revealed most of the structural changes.

**DESMOND**

Desmond[133] is a suite of computer programs for carrying out molecular dynamics simulations. Such simulations model the motion of a collection of atoms — a chemical system — over time, according to the laws of classical physics. The chemical system exists in a thermodynamic environment, which represents the conditions under which the simulation is carried out. This environment mimics the experimental conditions: whether the temperature or pressure is regulated, for example, or whether the system is isolated so that it can not exchange energy with its environment. The chemical system occupies a three-dimensional volume of space of a specified size, and each atom is generally represented by a particle at a specific position in that space. Motion is simulated in discrete time steps. From one step to the next, a tiny slice of time goes by, and atom positions update accordingly. The volume of space in which the simulation takes place is called the global cell. Desmond employs a technique known as periodic boundary conditions to wrap each face of the global cell to its opposite face. That is, particles that move leftwards out of the global cell appear to be moving in at a corresponding spot on the right-hand face, and *vice-versa*; particles that move out the top appear to enter at the bottom, and *vice-versa*; and finally, particles that move out the front appear at the back, and *vice-versa*.
A force field is a model of the potential energy of a chemical system. It is a set of functions and parameters used to model the potential energy of the system, and thereby to calculate the forces on each particle. To accurately simulate different kinds of systems, Desmond supports several variants of the Amber, CHARMM, and OPLS-AA force field models. The action of the force field on the particles is described by a differential equation that Desmond integrates — numerically solves— at every timestep, thus computing a new position and velocity for every particle in the system. The differential equation is based on the laws of Newtonian mechanics applied to particles in the system.

**GLIDE**

Glide[124] searches for favorable interactions between one or more ligand molecules and a receptor molecule, usually a protein. Each ligand must be a single molecule, while the receptor may include more than one molecule. Glide uses a hierarchical series of filters to search for possible locations of the ligand in the active-site region of the receptor. The shape and properties of the receptor are represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand poses. Conformational flexibility is handled in Glide by an extensive conformational search, augmented by a heuristic screen that rapidly eliminates unsuitable conformations, such as conformations that have long-range internal hydrogen bonds. Each ligand is divided into a core region and some number of rotamer groups. Each rotamer group is attached to the core by a rotatable bond, but does not contain additional rotatable bonds. During conformation generation, each core region is represented by a set of core conformations, the number of which depends on the number of rotatable bonds. For each core conformation an exhaustive search of possible locations and orientations is performed over the active site of the protein. The search begins with the selection of "site points" on an equally spaced 2 Å grid that covers the active site region (Fig. 60).

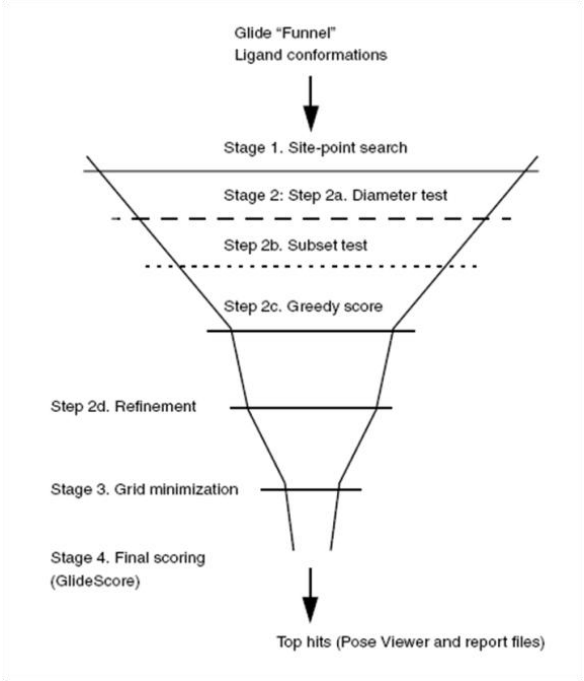



**Fig. 60: The Glide docking hierarchy.**

The second stage of the hierarchy begins by examining the placement of atoms that lie within a specified distance of the line drawn between the most widely separated atoms (the ligand diameter). Next rotation about the ligand diameter is considered, and the interactions of a subset consisting of all atoms capable of making hydrogen bonds or ligand-metal interactions with the receptor are scored (subset test). If this score is good enough, all

interactions with the receptor are scored. This stage is called “greedy scoring,” because the actual score for each atom depends not only on its position relative to the receptor but also on the best possible score it could get by moving ±1 Å in x, y, or z. The final step in Stage 2 is to re-score the top greedy-scoring poses *via* a “refinement” procedure, in which the ligand as a whole is allowed to move rigidly by ±1 Å in the Cartesian directions. Only a small number of the best refined poses is passed on to the third stage in the hierarchy-energy minimization on the pre-computed OPLS-AA van der Waals and electrostatic grids for the receptor. This energy minimization consists only of rigid-body translations and rotations when external conformations are docked. Finally, the minimized poses are re-scored using Schrödinger’s proprietary GlideScore scoring function. GlideScore is based on ChemScore, but includes a steric-clash term, adds buried polar terms devised by Schrödinger to penalize electrostatic mismatches, and has modifications to other terms:

*GScore = 0.065*vdW + 0.130*Coul + Lipo + Hbond + Metal + BuryP + RotB + Site*

| Component | Description |
|---|---|
| vdW | van der Waals energy. This term is calculated with reduced net ionic charges on groups with formal charges, such as metals, carboxylates, and guanidiniums. |
| Coul | Coulomb energy. This term is calculated with reduced net ionic charges on groups with formal charges, such as metals, carboxylates, and guanidiniums. |
| Lipo | Lipophilic contact term. Rewards favorable hydrophobic interactions. |
| HBond | Hydrogen-bonding term. This term is separated into differently weighted components that depend on whether the donor and acceptor are neutral, one is neutral and the other is charged, or both are charged. |
| Metal | Metal-binding term. Only the interactions with anionic acceptor atoms are included. If the net metal charge in the apo protein is positive, the preference for anionic ligands is included; if the net charge is zero, the preference is suppressed. |
| BuryP | Penalty for buried polar groups. |
| RotB | Penalty for freezing rotatable bonds. |
| Site | Polar interactions in the active site. Polar but non-hydrogen-bonding atoms in a hydrophobic region are rewarded. |

The choice of best-docked structure for each ligand is made using a model energy score (Emodel) that combines the energy grid score, the binding affinity predicted by GlideScore, and (for flexible docking) the internal strain energy for the model potential used to direct the conformational-search algorithm. This hierarchical search gives Glide exceptionally high accuracy in predicting the binding mode of the ligand.

The *virtual screening workflow* in the Maestro graphical user interface offers the user a unified interface for compound database processing and the submission of a series of large-scale docking runs using a hierarchy of Glide docking protocols. While this is a straightforward example of automation, other Maestro workflows offer more novel

functionality: for example, *induced fit docking* iterates between Glide docking and Prime homology modelling functionality in order to incorporate some degree of protein backbone and sidechain refinement;[168] *quantum-polarised ligand docking* iterates between Glide docking and Jaguar quantum chemical calculations in order to derive partial atomic charges that reflect polarisation of the ligand by the protein.[169]

**INDUCED FIT DOCKING**

The Induced Fit Protocol developed by Schrödinger, is a method for modeling the conformational changes induced by ligand binding. This protocol models induced fit docking of one or more ligands using the following steps:

1. Constrained minimization of the receptor (Glide protein preparation, refinement only) with an RMSD cut off of 0.18 Å.
2. Initial Glide docking of each ligand using a softened potential (van der Waals radii scaling).
3. One round of Prime side-chain prediction for each protein/ligand complex, on residues within a given distance of any ligand pose (default 5 Å).
4. Prime minimization of the same set of residues and the ligand for each protein/ligand complex pose. The receptor structure in each pose now reflects an induced fit to the ligand structure and conformation.
5. Glide redocking of each protein/ligand complex structure within a specified energy of the lowest-energy structure. The ligand is now rigorously docked, using XP Glide, into the induced-fit receptor structure
6. Estimation of the binding energy (IFDScore) for each output pose.

Each complex is then ranked according to the IFD score.

**PHASE**

Phase[153] is a versatile product for pharmacophore perception, structure alignment, activity prediction, and 3D database searching. Given a set of molecules with high affinity for a particular protein target, Phase uses fine-grained conformational sampling and a range of scoring techniques to identify common pharmacophore hypotheses, which convey characteristics of 3D chemical structures that are purported to be critical for binding.

Pharmacophores from all conformations of the ligands in the active set are examined, and those pharmacophores that contain identical sets of features, with very similar spatial arrangements, are grouped together. If a given group is found to contain at least one pharmacophore from each ligand, then this group gives rise to a common pharmacophore. Then common pharmacophores are examined, and a scoring procedure is applied to identify the pharmacophore from each surviving n-dimensional box that yields the best alignment of

the chosen actives. This pharmacophore provides a hypothesis to explain how the active molecules bind to the receptor. The scoring procedure provides a ranking of the different hypotheses, allowing to make rational choices about which hypotheses are most appropriate for further investigation. If the pharmacophore is an adequate hypothesis, it should discriminate between active and inactive molecules. Each hypothesis is accompanied by a set of aligned conformations that suggest the relative manner in which the molecules are likely to bind. The quality of alignment is measured in three ways: (1) the alignment score, which is the root-mean-squared deviation in the site-point positions; (2) the vector score, which is the average cosine of the angles formed by corresponding pairs of vector features (acceptors, donors, and aromatic rings) in the aligned structures; and (3) a volume score based on the overlap of van der Waals models of the non-hydrogen atoms in each pair of structures.

$$S_{Vol}(i) = \frac{V_{common}(i)}{V_{total}(i)}$$

where $V_{common}(i)$ is the common or overlapping volume between ligand $i$ and the reference ligand, while $V_{total}(i)$ is the total volume occupied by both ligands.

The site score, the vector score, and the volume score are combined with separate weights to yield a combined alignment score for each non-reference pharmacophore that has been aligned to the reference. Once the hypotheses have been scored on the basis of the alignment of the chosen actives, the score is adjusted by subtracting a multiple of the survival score of the inactives from the survival score of the actives.

Finally, hypothesis may be combined with known activity data to create a 3D QSAR model that identifies overall aspects of molecular structure that govern activity. This model may be used in conjunction with the hypothesis to mine a 3D database for molecules that are most likely to exhibit strong activity toward the target.

# References


1. J. Bajorath, R. Stenkamp, A. Aruffo, Knowledge-based model building of proteins: concepts and examples., *Protein Sci.* **2**, 1798-1810 (1993).

2. T. L. Blundell, B. L. Sibanda, M. J. E. Sternberg, J. M. Thornton, Knowledge-based prediction of protein structures and the design of novel molecules, *Nature* **326**, 347-352 (1987).

3. M. S. Johnson, N. Srinivasan, R. Sowdhamini, T. L. Blundell, Knowledge-based protein modeling, *Crit. Rev. Biochem. Mol. Biol.* **29**, 1-68 (1994).

4. Š. Andrej, Modelling mutations and homologous proteins, *Curr. Opin. Biotechnol.* **6**, 437-451 (1995).

5. R. Sánchez, A. Šali, Advances in comparative protein-structure modelling, *Curr. Opin. Struct. Biol.* **7**, 206-214 (1997).

6. C. Chothia, A. M. Lesk, The relation between the divergence of sequence and structure in proteins, *EMBO J.* **5**, 823-826 (1986).

7. P. Koehl, M. Levitt, A brighter future for protein structure prediction, *Nat. Struct. Biol.* **6**, 108-111 (1999).

8. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank, *Nucleic Acids Res.* **28**, 235-242 (2000).

9. D. Hanahan, R. A. Weinberg, The hallmarks of cancer, *Cell* **100**, 57-70 (2000).

10. S. Bao, G. Ouyang, X. Bai, Z. Huang, C. Ma, M. Liu, R. Shao, R. M. Anderson, J. N. Rich, X. F. Wang, Periostin potently promotes metastatic growth of colon cancer by augmenting cell survival *via* the Akt/PKB pathway, *Cancer Cell* **5**, 329-339 (2004).

11. C. D. Roskelley, M. J. Bissell, The dominance of the microenvironment in breast and ovarian cancer, *Semin. Cancer Biol.* **12**, 97-104 (2002).

12. M. W. Dewhirst, Concepts of oxygen transport at the microcirculatory level, *Semin. Radiat. Oncol.* **8**, 143-150 (1998).

13. S. E. Rademakers, P. N. Span, J. H. Kaanders, F. C. Sweep, A. J. van der Kogel, J. Bussink, Molecular aspects of tumour hypoxia, *Mol. Oncol.* **2**, 41-53 (2008).

14. A. L. Harris, Hypoxia-a key regulatory factor in tumour growth, *Nat. Rev. Cancer* **2**, 38-47 (2002).

15. A. J. Giaccia, M. B. Kastan, The complexity of p53 modulation: emerging patterns from divergent signals, *Gene Dev.* **12**, 2973-2983 (1998).

16. C. Koumenis, R. Alarcon, E. Hammond, P. Sutphin, W. Hoffman, M. Murphy, J. Derr, Y. Taya, S. W. Lowe, M. Kastan, A. Giaccia, Regulation of p53 by hypoxia: dissociation of transcriptional repression and apoptosis from p53-dependent transactivation, *Mol. Cell. Biol.* **21**, 1297-1310 (2001).

17. K. M. Ryan, A. C. Phillips, K. H. Vousden, Regulation and function of the p53 tumor suppressor protein, *Curr. Opin. Cell Biol.* **13**, 332-337 (2001).

18. W. G. An, M. Kanekal, M. C. Simon, E. Maltepe, M. V. Blagosklonny, L. M. Neckers, Stabilization of wild-type p53 by hypoxia-inducible factor 1α, *Nature* **392**, 405-408 (1998).

19. M. V. Blagosklonny, W. G. An, L. Y. Romanova, J. Trepel, T. Fojo, L. Neckers, p53 Inhibits Hypoxia-inducible Factor-stimulated Transcription, *J. Biol. Chem.* **273**, 11995-11998 (1998).

20. R. Ravi, B. Mookerjee B, Z. M. Bhujwalla, C. H. Sutter, D. Artemov, Q. Zeng, L. E. Dillehay, A. Madan, G. L. Semenza, A. Bedi, Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1α, *Genes Dev.* **14**, 34-44 (2000).

21. D. Chen, M. Li, J. Luo, W. Gu, Direct Interactions between HIF-1α and Mdm2 Modulate p53 Function, *J. Biol. Chem.* **278**, 13595-13598 (2003).

22. S. Latini, F. Pedata, Adenosine in the central nervous system: release mechanisms and extracellular concentrations, *J. Neurochem.* **79**, 463-484 (2001).

23. V. Ralevic, G. Burnstock, Receptors for purines and pyrimidines, *Pharmacol. Rev.* **50**, 413-492 (1998).

24. B. B. Fredholm, A. P. IJzerman, K. A. Jacobson, K. N. Klotz, J. Linden, International Union of Pharmacology. XXV. Nomenclature and classification of adenosine receptors, *Pharmacol. Rev.* **53**, 527-552 (2001).

25. S. Merighi, P. Mirandola, D. Milani, K. Varani, S. Gessi, K. N. Klotz, E. Leung, P. G. Baraldi, P. A. Borea, Adenosine receptors as mediators of both cell proliferation and cell death of cultured human melanoma cells, *J. Invest. Dermatol.* **119**, 923-933 (2002).

26. S. Merighi, P. G. Baraldi, S. Gessi, V. Iannotta, K. N. Klotz, E. Leung, P. Mirandola, M. A. Tabrizi, K. Varani, P. A. Borea, Adenosine receptors and human melanoma, *Drug Dev. Res.* **58**, 377-385 (2003).

27. Z. Gao, B. S. Li, Y. J. Day, J. Linden, A3 adenosine receptor activation triggers phosphorylation of protein kinase B and protects rat basophilic leukemia 2H3 mast cells from apoptosis, *Mol. Pharmacol.* **59**, 76-82 (2001).

28. L. Madi, S. Bar-Yehuda, F. Barer, E. Ardon, A. Ochaion, P. Fishman, A3 Adenosine receptor activation in melanoma cells, *J. Biol. Chem.* **278**, 42121-42130 (2003).

29. S. Gessi, K. Varani, S. Merighi, A. Morelli, D. Ferrari, E. Leung, P. G. Baraldi, G. Spalluto, P. A. Borea, Pharmacological and biochemical characterization of A3 adenosine receptors in Jurkat T cells, *Br. J. Pharmacol.* **134**, 116-126 (2001).

30. S. Merighi, K. Varani, S. Gessi, E. Cattabriga, V. Iannotta, C. Ulouglu, E. Leung, P. A. Borea, Pharmacological and biochemical characterization of adenosine receptors in the human malignant melanoma A375 cell line, *Br. J. Pharmacol.* **134**, 1215-1226 (2001).

31. B. Suh, T. Kim, J. Lee, J. Seong, K. Kim, Pharmacological characterization of adenosine receptors in PGT-β mouse pineal gland tumour cells, *Br. J. Pharmacol.* **134**, 132-142 (2001).

32. S. Gessi, K. Varani, S. Merighi, E. Cattabriga, V. Iannotta, E. Leung, P. G. Baraldi, P. A. Borea, A(3) adenosine receptors in human neutrophils and promyelocytic HL60 cells: a pharmacological and biochemical study, *Mol. Pharmacol.* **61**, 415-424 (2002).

33. S. Gessi, E. Cattabriga, A. Avitabile, R. Gafa', G. Lanza, L. Cavazzini, N. Bianchi, R. Gambari, C. Feo, A. Liboni, S. Gullini, E. Leung, S. Mac-Lennan, P. A. Borea, Elevated expression of A3 adenosine receptors in human colorectal cancer is reflected in peripheral blood cells, *Clin. Cancer Res.* **10**, 5895-5901 (2004).

34. S. Merighi, A. Benini, P. Mirandola, S. Gessi, K. Varani, E. Leung, S. Mac-Lennan, P. G. Baraldi, P. A. Borea, A3 adenosine receptors modulate hypoxia-inducible factor-1α expression in human A375 melanoma cells, *Neoplasia* **7**, 894-903 (2005).

35. G. L. Wang, B. H. Jiang, E. A. Rue, G. L. Semenza, Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5510-5514 (1995).

36. P. C. Mahon, K. Hirota, G. L. Semenza, FIH-1: a novel protein that interacts with HIF-1α and VHL to mediate repression of HIF-1 transcriptional activity, *Genes Dev.* **15**, 2675-2686 (2001).

37. P. Maxwell, M. S. Wiesener, G. W. Chang, S. C. Clifford, E. C. Vaux, M. E. Cockman, C. C. Wykoff, C. W. Pugh, E. R. Maher, P. J. Ratcliffe, The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis, *Nature* **399**, 271-275 (1999).

38. P. Jaakkola, D. R. Mole, Y. M. Tian, M. I. Wilson, J. Gielbert, S. J. Gaskell, A. Kriegsheim, H. F. Hebestreit, M. Mukherji, C. J. Schofield, P. H. Maxwell, C. W. Pugh, P. J. Ratcliffe, Targeting of HIF-α to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation, *Science* **292**, 468-472 (2001).

39. M. Ivan, K. Kondo, H. Yang, W. Kim, J. Valiando, M. Ohh, A. Salic, J. M. Asara, W. S. Lane, W. G. Jr Kaelin, HIF-α targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing, *Science* **292**, 464-468 (2001).

40. M. E. Cockman, N. Masson, D. R. Mole, P. Jaakkola, G. W. Chang, S. C. Clifford, E. R. Maher, C. W. Pugh, P. J. Ratcliffe, P. H. Maxwell, Hypoxia inducible factor-α binding and ubiquitylation by the von Hippel-Lindau tumor suppressor protein, *J. Biol. Chem.* **275**, 25733-25741 (2000).

41. R. K. Bruick, S. L. McKnight, A conserved family of prolyl-4-hydroxylases that modify HIF, *Science* **294**, 1337-1340 (2001).

42. A. C. Epstein, J. M. Gleadle, L. A. McNeill, K. S. Hewitson, J. O'Rourke, D. R. Mole, M. Mukherji, E. Metzen, M. I. Wilson, A. Dhanda,Y. M. Tian, N. Masson, D. L. Hamilton, P. Jaakkola, R. Barstead, J. Hodgkin, P. H. Maxwell, C. W. Pugh, C. J. Schofield, P. J. Ratcliffe, C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation, *Cell* **107**, 43-54 (2001).

43. V. Nizet, R. S. Johnson, Interdependence of hypoxic and innate immune responses, *Nat. Rev. Immunol.* **9**, 609-617 (2009).

44. D. C. Schwartz, M. Hochstrasser, A superfamily of protein tags: ubiquitin, SUMO and related modifiers, *Trends Biochem. Sci.* **28**, 321-328 (2003).

45. F. Melchior, M. Schergaut, A. Pichler, SUMO: ligases, isopeptidases and nuclear pores, *Trends Biochem. Sci.* **28**, 612-618 (2003).

46. M. Tojo, K. Matsuzaki, T. Minami, Y. Honda, H. Yasuda, T. Chiba, H. Saya, Y. Fujii-Kuriyama, M. Nakao, The aryl hydrocarbon receptor nuclear transporter is modulated by the SUMO-1 conjugation system, *J. Biol. Chem.* **277**, 46576-46585 (2002).

47. H. B. Newton, Molecular neuro-oncology and development of targeted therapeutic strategies for brain tumors. Part 2: PI3K/Akt/PTEN, mTOR, SHH/PTCH and angiogenesis, *Expert Rev. Anticancer Ther.* **4**, 105-128 (2004).

48. E. Laughner, P. Taghavi, K. Chiles, P. C. Mahon, G. L. Semenza, HER2 (neu) signaling increases the rate of hypoxia-inducible factor 1α (HIF-1α) synthesis: novel mechanism for HIF-1-mediated vascular endothelial growth factor expression, *Mol. Cell. Biol.* **21**, 3995-4004 (2001).

49. H. Zhong, K. Chiles, D. Feldser, E. Laughner, C. Hanrahan, M. M. Georgescu, J. W. Simons, G. S. Semenza, Modulation of hypoxia-inducible factor 1α expression by the epidermal growth factor/phosphatidylinositol 3-kinase/PTEN/AKT/FRAP pathway in human prostate cancer cells: implications for tumor angiogenesis and therapeutics, *Cancer Res.* **60**, 1541-1545 (2000).

50. F. S. Wang, C. J. Wang, Y. J. Chen, P. R. Chang, Y. T. Huang, Y. C. Sun, H. C. Huang, Y. J. Yang, K. D. Yang, Ras induction of superoxide activates ERK-dependent angiogenic transcription factor HIF-1α and VEGF-A expression in shock wave-stimulated osteoblasts, *J. Biol. Chem.* **279**, 10331-10337 (2004).

51. E. Hur, K. Y. Chang, E. Lee, S. K. Lee, H. Park, Mitogen-activated protein kinase kinase inhibitor PD98059 blocks the trans-activation but not the stabilization or DNA binding ability of hypoxia-inducible factor-1α, *Mol. Pharmacol.* **59**, 1216-1224 (2001).

52. E. B. Friedrich, E. Liu, S. Sinha, S. Cook, D. S. Milstone, C. A. MacRae, M. Mariotti, P. J. Kuhlencordt, T. Force, A. Rosenzweig, R. St-Arnaud, S. Dedhar, R. E. Gerszten, Integrin-linked kinase regulates endothelial cell survival and vascular development, *Mol. Cell. Biol.* **24**, 8134-8144 (2004).

53. C. Tan, S. Cruet-Hennequart, A. Troussard, L. Fazli, P. Costello, K. Sutton, J. Wheeler, M. Gleave, J. Sanghera, S. Dedhar, Regulation of tumor angiogenesis by integrin-linked kinase (ILK), *Cancer Cell* **5**, 79-90 (2004).

54. J. Zhou, T. Schmid, R. Frank, B. Brüne, PI3K/Akt Is Required for Heat Shock Proteins to Protect Hypoxia-inducible Factor 1α from pVHL-independent Degradation, *J. Biol. Chem.* **279**, 13506-13513 (2004).

55. P. Büchler, H. A. Reber, M. Büchler, S. Shrinkante, M. W. Büchler, H. Friess, G. L. Semenza, O. J. Hines, Hypoxia-inducible factor 1 regulates vascular endothelial growth factor expression in human pancreatic cancer, *Pancreas* **26**, 56-64 (2003).

56. D. Feldser, F. Agani, N. V. Iyer, B. Pak, G. Ferreira, G. L. Semenza, Reciprocal positive regulation of hypoxia-inducible factor 1α and insulin-like growth factor 2, *Cancer Res.* **59**, 3915-3918 (1999).

57. S. Kajimura, K. Aida, C. Duan, Insulin-like growth factor-binding protein-1 (IGFBP-1) mediates hypoxia-induced embryonic growth and developmental retardation, *Proc. Natl. Acad. Sci.* **102**, 1240-1245 (2005).

58. N. Ferrara, T. Davis-Smyth, The biology of vascular endothelial growth factor, *Endocr. Rev.* **18**, 4 -25 (1997).

59. T. Hellwig-Bürgel, K. Rutkowski, E. Metzen, J. Fandrey, W. Jelkmann, Interleukin-1β and tumor necrosis factor-α stimulate DNA binding of hypoxia-inducible factor-1, *Blood* **94**, 1561-1567 (1999).

60. M. B. Kastan, O. Onyekwere, D. Sidransky, B. Vogelstein, R. W. Craig, Participation of p53 protein in the cellular response to DNA damage, *Cancer Res.* **51**, 6304-6311 (1991).

61. X. Wu, A. J. Levine, p53 and E2F-1 cooperate to mediate apoptosis, *Proc. Natl. Acad. Sci. U S A* **91**, 3602-3606 (1994).

62. H. Hermeking, D. Eick, Mediation of c-Myc-induced apoptosis by p53, *Science* **265**, 2091-2093 (1994).

63. J. D. Oliner, J. A. Pietenpol, S. Thiagalingam, J. Gyuris, K. W. Kinzler, B. Vogelstein, Oncoprotein Mdm2 conceals the activation domain of tumour suppressor p53, *Nature* **362**, 857-860 (1993).

64. M. S. Greenblatt, W. P. Bennett, M. Hollstein, C. C. Harris, Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis, *Cancer Res.* **54**, 4855-4878 (1994).

65. J. D. Oliner, K. W. Kinzler, P. S. Meltzer, D. L. George, B. Vogelstein, Amplification of a gene encoding a p53-associated protein in human sarcomas, *Nature* **358**, 80-83 (1992).

66. C. Bueso-Ramos, Y. Yang, E. deLeon, P. McCown, S. A. Stass, M. Albitar, The human mdm2 oncogene is overexpressed in leukemias, *Blood* **82**, 2617 -2623 (1993).

67. C. Wasylyk, R. Salvi, M. Argentini, C. Dureuil, I. Delumeau, J. Abecassis, L. Debussche, B. Wasylyk, p53 mediated death of cells overexpressing Mdm2 by an inhibitor of Mdm2 interaction with p53, *Oncogene* **18**, 1921-1934 (1999).

68. A. Böttger, V. Böttger, C. Garcia-Echeverria, P. Chène, H. K. Hochkeppel, W. Sampson, K. Ang, S. F.Howard, S. M. Picksley, D. P. Lane, Molecular characterization of the hdm2-p53 interaction, *J. Mol. Biol.* **269**, 744-756 (1997).

69. C. García-Echeverría, P. Chène, M. J. Blommers, P. Furet, Discovery of potent antagonists of the interaction between human double minute 2 and tumor suppressor p53, *J. Med. Chem.* **43**, 3205-3208 (2000).

70. L. Cahilly-Snyder, T. Yang-Feng, U. Francke, D. L. George, Molecular analysis and chromosomal mapping of amplified genes isolated from a transformed mouse 3T3 cell line, *Somatic Cell Mol. Genet.* **13**, 235-244 (1987).

71. J. Momand, G. P. Zambetti, D. C. Olson, D. George, A. J. Levine, The mdm2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation, *Cell* **69**, 1237-1245 (1992).

72. Y. Haupt, R. Maya, A. Kazaz, M. Oren, Mdm2 promotes the rapid degradation of p53, *Nature* **387**, 296-299 (1997).

73. R. Honda, H. Tanaka, H. Yasuda, Oncoprotein Mdm2 is a ubiquitin ligase E3 for tumor suppressor p53, *FEBS Letters* **420**, 25-27 (1997).

74. M. H. Kubbutat, S. N. Jones, K. H. Vousden, Regulation of p53 stability by Mdm2, *Nature* **387**, 299-303 (1997).

75. J. C. Lee, M. E. Peter, Regulation of apoptosis by ubiquitination, *Immunol. Rev.* **193**, 39-47 (2003).

76. Y. Yang, X. Yu, Regulation of apoptosis: the ubiquitous way, *FASEB J.* **17**, 790-799 (2003).

77. S. Nakamura, J. A. Roth, T. Mukhopadhyay, Multiple Lysine Mutations in the C-terminal domain of p53 interfere with Mdm2-dependent protein degradation and ubiquitination, *Mol. Cell. Biol.* **20**, 9391-9398 (2000).

78. M. S. Rodriguez, J. M. P. Desterro, S. Lain, D. P. Lane, R. T. Hay, Multiple C-terminal lysine residues target p53 for ubiquitin-proteasome-mediated degradation, *Mol. Cell. Biol.* **20**, 8458-8467 (2000).

79. S. Fang, J. P. Jensen, R. L. Ludwig, K. H. Vousden, A. M. Weissman, Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself and p53, *J. Biol. Chem.* **275**, 8945-8951 (2000).

80. Honda, H. Yasuda, Activity of Mdm2, a ubiquitin ligase, toward p53 or itself is dependent on the RING finger domain of the ligase., *Oncogene* **19**, 1473-1476 (2000).

81. D. Michael, M. Oren, The p53-Mdm2 module and the ubiquitin system, *Semin. Cancer Biol.* **13**, 49-58 (2003).

82. S. R. Grossman, M. Perez, A. L. Kung, M. Joseph, C. Mansur, Z. X. Xiao, S. Kumar, P. M Howley, D. M. Livingston, p300/Mdm2 complexes participate in Mdm2-mediated p53 degradation, *Molecular Cell* **2**, 405-415 (1998).

83. S. R. Grossman, M. E. Deato, C. Brignone, H. M. Chan, A. L. Kung, H. Tagami, Y. Nakatani, D. M. Livingston, Polyubiquitination of p53 by a ubiquitin ligase activity of p300, *Science* **300**, 342-344 (2003).

84. J. Roth, M. Dobbelstein, D. A. Freedman, T. Shenk, A. J. Levine, Nucleo-cytoplasmic shuttling of the hdm2 oncoprotein regulates the levels of the p53 protein *via* a pathway used by the human immunodeficiency virus rev protein., *EMBO J.* **17**, 554-564 (1998).

85. A. Ito, Y. Kawaguchi, C. H. Lai, J. J. Kovacs, Y. Higashimoto, E. Appella, T. P. Yao, Mdm2-HDAC1-mediated deacetylation of p53 is required for its degradation, *EMBO J.* **21**, 6236-6245 (2002).

86. P. H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A. J. Levine, N. P. Pavletich, Structure of the Mdm2 oncoprotein bound to the p53 tumor suppressor transactivation domain, *Science* **274**, 948-953 (1996).

87. M. A. McCoy, J. J. Gesell, M. M. Senior, D. F. Wyss, Flexible lid to the p53-binding domain of human Mdm2: Implications for p53 regulation, *PNAS* **100**, 1645-1648 (2003).

88. L. T. Vassilev, Mdm2 inhibitors for cancer therapy, *Trends Mol. Med.* **13**, 23-31 (2007).

89. I. Massova, P. A. Kollman, Computational alanine scanning to probe protein−protein interactions: a novel approach to evaluate binding free energies, J. Am. Chem. Soc. 121, 8133-8143 (1999).

90. H. Zhong, H. A. Carlson, Computational studies and peptidomimetic design for the human p53–Mdm2 complex, *Proteins Struct. Funct. Bioinf.* **58**, 222-234 (2005).

91. L. M. Espinoza-Fonseca, J. G. Trujillo-Ferrara, Conformational changes of the p53-binding cleft of Mdm2 revealed by molecular dynamics simulations, *Biopolymers* **83**, 365-373 (2006).

92. A. Macchiarulo, N. Giacche, A. Carotti, M. Baroni, G. Cruciani, R. Pellicciari, Targeting the conformational transitions of Mdm2 and MdmX: insights into dissimilarities and similarities of p53 recognition, *J. Chem. Inf. Mod.* **48**, 1999-2009 (2008).

93. S. A. Showalter, L. Bruschweiler-Li, E. Johnson, F. Zhang, R. Brüschweiler, Quantitative lid dynamics of Mdm2 reveals differential ligand binding modes of the p53-binding cleft, *J. Am. Chem. Soc.* **130**, 6472-6478 (2008).

94. T. L. Joseph, A. Madhumalar, C. J. Brown, D. P. Lane, C. S. Verma, Differential binding of p53 and nutlin to Mdm2 and MdmX: computational studies, *Cell Cycle* **9**, 1167-1181 (2010).

95. G. W. Yu, M. Vaysburd, M. D. Allen, G. Settanni, A. R. Fersht, Structure of human Mdm4 N-terminal domain bound to a single-domain antibody, *J. Mol. Biol.* **385**, 1578-1589 (2009).

96. S. G. Dastidar, A. Madhumalar, G. Fuentes, D. P. Lane, C. S. Verma, Forces mediating protein–protein interactions: a computational study of p53 “approaching” Mdm2, *Theor. Chem. Acc.* **125**, 621-635 (2009).

97. M. P. Abbracchio, R. Brambilla, S. Ceruti, H. O. Kim, D. K. von Lubitz, K. A. Jacobson, F. Cattabeni, G protein-dependent activation of phospholipase C by adenosine A3 receptors in rat brain, *Mol. Pharmacol.* **48**, 1038-1045 (1995).

98. B. B. Fredholm, A. P. IJzerman, K. A. Jacobson, K. N. Klotz, J. Linden, International Union of Pharmacology. XXV. Nomenclature and classification of adenosine receptors, *Pharmacol. Rev.* **53**, 527-552 (2001).

99. R. Yaar, M. R. Jones, J.-F. Chen, K. Ravid, Animal models for the study of adenosine receptor function, *J. Cell. Physiol.* **202**, 9-20 (2005).

100. D. M. Perez, S. S. Karnik, Multiple signaling states of G-protein-coupled receptors, *Pharmacol. Rev.* **57**, 147-161 (2005).

101. M. Klinger, M. Freissmuth, C. Nanoff, Adenosine receptors: G protein-mediated signalling and the role of accessory proteins, *Cell. Signal.* **14**, 99-108 (2002).

102. K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, M. Miyano, Crystal structure of rhodopsin: A G protein-coupled receptor, *Science* **289**, 739-745 (2000).

103. F. Fanelli, P. G. De Benedetti, Inactive and active states and supramolecular organization of GPCRs: insights from computational modeling, *J. Comput. Aided Mol. Des.* **20**, 449-461 (2006).

104. V.-P. Jaakola, M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. Chien, J. R. Lane, A. P. Ijzerman, R. C. Stevens, The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist, *Science* **322**, 1211-1217 (2008).

105. D. M. Rosenbaum, V. Cherezov, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, X. J. Yao, W. I. Weis, R. C. Stevens, B. K. Kobilka, GPCR engineering yields high-resolution structural insights into β2-adrenergic receptor function, *Science* **318**, 1266-1273 (2007).

106. T. Warne, M. J. Serrano-Vega, J. G. Baker, R. Moukhametzianov, P. C. Edwards, R. Henderson, A. G. Leslie, C. G. Tate, G. F. Schertler, Structure of a β1-adrenergic G-protein-coupled receptor, *Nature* **454**, 486-491 (2008).

107. T. Okada, M. Sugihara, A. N. Bondar, M. Elstner, P. Entel, V. Buss, The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure, *J. Mol. Biol.* **342**, 571-583 (2004).

108. K. Lundgren, C. Holm, G. Landberg, Hypoxia and breast cancer: prognostic and therapeutic implications, *Cell. Mol. Life Sci.* **64**, 3233-3247 (2007).

109. Q.-T. Le, D. Courter, Clinical biomarkers for hypoxia targeting, *Cancer and Metastasis Rev.* **27**, 351-362 (2008).

110. G. L. Semenza, Targeting HIF-1 for cancer therapy, *Nat. Rev. Cancer* **3**, 721-732 (2003).

111. A. Bairoch, R. Apweiler, The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000, *Nucleic Acids Res.* **28**, 45-48 (2000).

112. N. V. Iyer, S. W. Leung, G. L. Semenza, The human hypoxia-inducible factor 1α gene: HIF-1α structure and evolutionary conservation, *Genomics* **52**, 159-165 (1998).

113. D. Baker, A. Sali, Protein Structure Prediction and Structural Genomics, *Science* **294**, 93-96 (2001).

114. S. Shacham, M. Topf, N. Avisar, F. Glaser, Y. Marantz, S. Bar-Haim, S. Noiman, Z. Naor, O. M. Becker, Modeling the 3D structure of GPCRs from sequence, *Med. Res. Rev.* **21**, 472-483 (2001).

115. N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali, Comparative protein structure modeling using Modeller, *Curr. Protocols in Bioinformatics*, **15**, 5.6.1–5.6.30 (2006).

116. B. John, A. Sali, Comparative protein structure modeling by iterative alignment, model building and model assessment, *Nucleic Acids Res.* **31**, 3982-3992 (2003).

117. S. Wu, Y. Zhang, LOMETS: a local meta-threading-server for protein structure prediction, *Nucleic Acids Res.* **35**, 3375-3382 (2007).

118. S. C. Lovell, I. W. Davis, W. B. Arendall, P. I. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, D. C. Richardson, Structure validation by C-α geometry: phi, psi and C-β deviation, *Proteins* **50**, 437-450 (2003).

119. A. Chapman-Smith, M. L. Whitelaw, Novel DNA Binding by a Basic Helix-Loop-Helix Protein: The role of the dioxin receptor PAS domain, *J. Biol. Chem.* **281**, 12535-12545 (2006).

120. A. Chapman-Smith, J. K. Lutwyche, M. L. Whitelaw, Contribution of the Per/Arnt/Sim (PAS) domains to DNA binding by the basic helix-loop-helix PAS transcriptional regulators, *J. Biol. Chem.* **279**, 5353-5362 (2004).

121. B. H. Jiang, E. Rue, G. L. Wang, R. Roe, G. L. Semenza, Dimerization, DNA binding, and transactivation properties of hypoxia-inducible factor 1, *J. Biol. Chem.* **271**, 17771-17778 (1996).

122. T. Liu, Y. Lin, X. Wen, R. N. Jorissen, M. K. Gilson, BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities, *Nucleic Acids Res.* **35**, D198-201 (2007).

123. G. Michel, E. Minet, I. Ernest, I. Roland, F. Durant, J. Remacle, C. Michiels, A model for the complex between the hypoxia-inducible factor-1 (HIF-1) and its consensus DNA sequence, *J. Biomol. Struct. Dyn.* **18**, 169-179 (2000).

124. Glide, version 5.5, Schrödinger, LLC, New York, NY, 2009.

125. G. Klebe, Virtual ligand screening: strategies, perspectives and limitations, *Drug Discov. Today* **11**, 580-594 (2006).

126. M. H. J. Seifert, J. Kraus, B. Kramer, Virtual high-throughput screening of molecular databases, *Curr. Opin.Drug Discov. Devel.* **10**, 298-307 (2007).

127. B. H. Jiang, J. Z. Zheng, S. W. Leung, R. Roe, G. L. Semenza, Transactivation and inhibitory domains of hypoxia-inducible factor 1α. Modulation of transcriptional activity by oxygen tension, *J. Biol. Chem.* **272**, 19253-19260 (1997).

128. C. W. Pugh, J. F. O'Rourke, M. Nagao, J. M. Gleadle, P. J. Ratcliffe, Activation of hypoxia-inducible factor-1; definition of regulatory domains within the α subunit, *J. Biol. Chem.* **272**, 11205-11214 (1997).

129. Z. Arany, L. E. Huang, R. Eckner, S. Bhattacharya, C. Jiang, M. A. Goldberg, H. F. Bunn, D. M. Livingston, An essential role for p300/CBP in the cellular response to hypoxia, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12969-12973 (1996).

130. M. Ema, K. Hirota, J. Mimura, H. Abe, J. Yodoi, K. Sogawa, L. Poellinger, Y. Fujii-Kuriyama, Molecular mechanisms of transcription activation by HLF and HIF-1α in response to hypoxia: their stabilization and redox signal-induced interaction with CBP/p300, *EMBO J.* **18**, 1905-1914 (1999).

131. P. Carrero, P. Carrero, K. Okamoto, P. Coumailleau, S. O'Brien, H. Tanaka, L. Poellinger, Redox-regulated recruitment of the transcriptional coactivators CREB-binding protein and SRC-1 to hypoxia-inducible factor 1α, *Mol. Cell. Biol.* **20**, 402-415 (2000).

132. D. Lando, D. J. Peet, D. A. Whelan, J. J. Gorman, M. L. Whitelaw, Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch, *Science* **295**, 858-861 (2002).

133. Desmond Molecular Dynamics System, version 2.2, D. E. Shaw Research, New York, NY,2009. Maestro-Desmond interoperability tools, version 2.2, Schrödinger, New York, NY,2009.

134. J. J. Irwin, B. K. Shoichet, ZINC--a free database of commercially available compounds for virtual screening, *J. Chem. Inf. Model.* **45**, 177-182 (2005).

135. D. C. Fry, S. D. Emerson, S. Palme, B. T. Vu, C. M. Liu, F. Podlaski, NMR structure of a complex between Mdm2 and a small molecule inhibitor, *J. Biomol. NMR* **30**, 163-173 (2004).

136. Schrödinger Suite 2009 Induced Fit Docking protocol; Glide version 5.5, Schrödinger, LLC,New York, NY, 2009; Prime version 2.1, Schrödinger, LLC, New York, NY, 2009.

137. H. Zhong, H. A. Carlson, Computational studies and peptidomimetic design for the human p53–Mdm2 complex, *Proteins* **58**, 222-234 (2005).

138. A. Carotti, A. Macchiarulo, N. Giacchè, R. Pellicciari, Targeting the conformational transitions of Mdm2 and MdmX: insights into key residues affecting p53 recognition, *Proteins* **77**, 524-535 (2009).

139. L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi, E. A. Liu, *In vivo* activation of the p53 pathway by small-molecule antagonists of Mdm2, *Science* **303**, 844-848 (2004).

140. F. G. Sajjadi, G. S. Firestein, cDNA cloning and sequence analysis of the human A3 adenosine receptor, *Biochim. Biophys. Acta* **1179**, 105-107 (1993).

141. C. A. Salvatore, M. A. Jacobson, H. E. Taylor, J. Linden, R. G. Johnson, Molecular cloning and characterization of the human A3 adenosine receptor, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10365-10369 (1993).

142. A. Fiser, A. Sali, ModLoop: automated modeling of loops in protein structures, *Bioinformatics* **19**, 2500-2501 (2003).

143. P. Benkert, M. Künzli, T. Schwede, QMEAN server for protein model quality estimation, *Nucleic Acids Res.* **37**, W510-514 (2009).

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

145. H. Frauenfelder, S. G. Sligar, P. G. Wolynes, The energy landscapes and motions of proteins, *Science* **254**, 1598-1603 (1991).

146. V. Colotta, D. Catarzi, F. Varano, F. R. Calabri, O. Lenzi, G. Filacchioni, C. Martini, L. Trincavelli, F. Deflorian, S. Moro, 1,2,4-Triazolo[4,3-a]quinoxalin-1-one moiety as an attractive scaffold to develop new potent and selective human A3 adenosine receptor antagonists: synthesis, pharmacological, and ligand-receptor modeling studies, *J. Med. Chem.* **47**, 3580-3590 (2004).

147. O. Lenzi, V. Colotta, D. Catarzi, F. Varano, G. Filacchioni, C. Martini, L. Trincavelli, O. Ciampi, K. Varani, F. Marighetti, E. Morizzo, S. Moro, 4-Amido-2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-ones as new potent and selective human A3 adenosine receptor antagonists. synthesis, pharmacological evaluation, and ligand-receptor modeling studies, *J. Med. Chem.* **49**, 3916-3925 (2006).

148. V. Colotta, D. Catarzi, F. Varano, O. Lenzi, G. Filacchioni, C. Martini, L. Trincavelli, O. Ciampi, C. Traini, A. M. Pugliese, F. Pedata, E. Morizzo, S. Moro, Synthesis, ligand-receptor modeling studies and pharmacological evaluation of novel 4-modified-2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-one derivatives as potent and selective human A3 adenosine receptor antagonists, *Bioorg. Med. Chem.* **16**, 6086-6102 (2008).

149. V. Colotta, D. Catarzi, F. Varano, F. Capelli, O. Lenzi, G. Filacchioni, C. Martini, L. Trincavelli, O. Ciampi, A. M. Pugliese, F. Pedata, A. Schiesaro, E. Morizzo, S. Moro, New 2-arylpyrazolo[3,4-c]quinoline derivatives as potent and selective human A3 adenosine receptor antagonists. Synthesis, pharmacological evaluation, and ligand-receptor modeling studies, *J. Med. Chem.* **50**, 4061-4074 (2007).

150. E. Morizzo, F. Capelli, O. Lenzi, D. Catarzi, F. Varano, G. Filacchioni, F. Vincenzi, K. Varani, P. A. Borea, V. Colotta, S. Moro, Scouting human A3 adenosine receptor antagonist binding mode using a molecular simplification approach: from triazoloquinoxaline to a pyrimidine skeleton as a key study, *J. Med. Chem.* **50**, 6596-6606 (2007).

151. C. Bolcato, C. Cusan, G. Pastorin, G. Spalluto, B. Cacciari, K. N. Klotz, E. Morizzo, S. Moro, Pyrazolo-triazolo-pyrimidines as adenosine receptor antagonists: Effect of the N-5 bond type on the affinity and selectivity at the four adenosine receptor subtypes, *Purinergic Signal.* **4**, 39-46 (2008).

152. Y. C. Kim, X. D. Ji, K. A. Jacobson, Derivatives of the triazoloquinazoline adenosine antagonist (CGS15943) are selective for the human A3 receptor subtype, *J. Med. Chem.* **39**, 4142-4148 (1996).

153. Phase, version 3.1, Schrödinger, LLC, New York, NY, 2009.

154. J. Zupan, M. Novič, I. Ruisánchez, Kohonen and counterpropagation artificial neural networks in analytical chemistry, *Chemometr. Intell. Lab.* **38**, 1-23 (1997).

155. J. Gasteiger, J. Zupan, Neural networks in chemistry, *Angew. Chem. Int. Edit.* **32**, 503-527 (1993).

156. QikProp, version 3.2, Schrödinger, LLC, New York, NY, 2009.

157. S. Wu, J. Skolnick, Y. Zhang, *Ab initio* modeling of small proteins by iterative TASSER simulations, BMC Biology 5, 17 (2007).

158. Y. Zhang, A. Kolinski, J. Skolnick, TOUCHSTONE II: a new approach to *ab initio* protein structure prediction, *Biophys. J.* **85**, 1145-1164 (2003).

159. Y. Zhang, D. Kihara, J. Skolnick, Local energy landscape flattening: parallel hyperbolic Monte Carlo sampling of protein folding, *Proteins* **48**, 192-201 (2002).

160. Y. Zhang, J. Skolnick, SPICKER: A clustering approach to identify near-native protein folds, *J. Comput. Chem.* **25**, 865-871 (2004).

161. Y. Zhang, J. Skolnick, TM-align: a protein structure alignment algorithm based on the TM-score, *Nucleic Acids Res.* **33**, 2302 -2309.

162. M. Feig, P. Rotkiewicz, A. Kolinski, J. Skolnick, C. L. Brooks III, Accurate reconstruction of all-atom protein representations from side-chain-based low-resolution models, *Proteins* **41**, 86-97 (2000).

163. A. A. Canutescu, A. A. Shelenkov, R. L. Dunbrack Jr, A graph-theory algorithm for rapid protein side-chain prediction, *Protein Sci.* **12**, 2001-2014 (2003).

164. H. J. C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation, *Comput. Phys. Commun.* **91**, 43-56 (1995).

165. E. Lindahl, B. Hess, D. V. D. Spoel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, *J. Mol. Model.* **7**, 306-317 (2001).

166. A. Pedretti, L. Villa, G. Vistoli, VEGA: a versatile program to convert, handle and visualize molecular structure on Windows-based PCs, *J. Mol. Graph. Model.* **21**, 47-49 (2002).

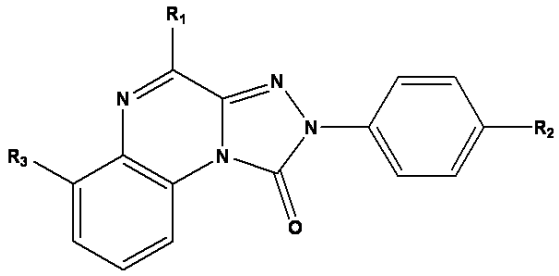
167. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph.* **14**, 33-38, 27-28 (1996).

168. W. Sherman, T. Day, M. P. Jacobson, R. A. Friesner, R. Farid, Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects, *J. Med. Chem.* **49**, 534-553 (2005).

169. A. E. Cho, V. Guallar, B. J. Berne, R. Friesner, Importance of accurate charges in molecular docking: Quantum Mechanical/Molecular Mechanical (QM/MM) approach, *J. Comput. Chem.* **26**, 915-931 (2005).

# APPENDIX

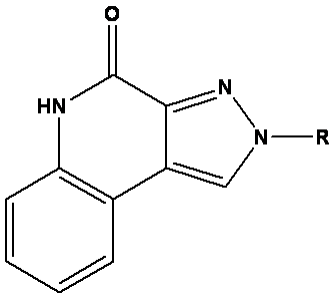
## Group 1: triazolo-quinoxaline derivatives



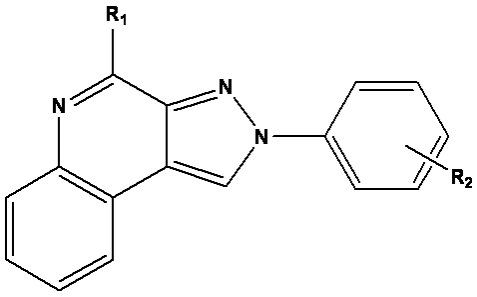

| Ligand | $R_1$ | $R_2$ | $R_3$ | pKi |
|---|---|---|---|---|
| 1 | $NHCOCH_3$ | H | H | 8,70 |
| 2 | $NHCOCH_3$ | OMe | H | 7,45 |
| 3 | $NHCOCH_3$ | H | $NH_2$ | 7,32 |
| 4 | $NHCOCH_3$ | OMe | $NH_2$ | 8,26 |
| 5 | NHCOPh | H | H | 8,83 |
| 6 | NHCOPh | OMe | H | 8,54 |
| 7 | NHCOPh | $NO_2$ | H | 7,00 |
| 8 | NHCOPh | H | $NO_2$ | 7,66 |
| 9 | NHCOPh | OMe | $NO_2$ | 6,66 |
| 10 | NHCOPh | H | $NH_2$ | 7,66 |
| 11 | NHCOPh | OMe | $NH_2$ | 9,00 |
| 12 | $NHCOCHPh_2$ | OMe | H | 7,36 |
| 13 | $NHCOCHPh_2$ | H | H | 9,09 |
| 14 | $NHCOCHPh_2$ | H | $NO_2$ | 7,83 |
| 15 | $NHCOCHPh_2$ | OMe | $NO_2$ | 9,10 |
| 16 | $NHCOCHPh_2$ | H | $NH_2$ | 8,06 |
| 17 | $NHCOCHPh_2$ | OMe | $NH_2$ | 8,59 |
| 18 | $N(COPh)_2$ | H | H | 8,28 |
| 19 | $N(COPh)_2$ | OMe | H | 8,48 |
| 20 | $N(COPh)_2$ | OMe | $NO_2$ | 6,46 |
| 21 | $N(COPh)_2$ | H | $NH_2$ | 5,91 |
| 22 | NHCOPh | H | H | 8,83 |
| 23 | $NHCOC_6H_4$-4-COOMe | OMe | H | 5,86 |
| 24 | NHCO-4-Pyridyl | H | H | 8,21 |
| 25 | NHCO-4-Pyridyl | OMe | H | 7,17 |
| 26 | $NHSO_2Ph$ | H | H | 7,49 |
| 27 | $NHSO_2Ph$ | OMe | H | 8,66 |
| 28 | $NHSO_2Ph$ | H | $NO_2$ | 7,00 |
| 29 | $NHSO_2CH_3$ | H | H | 5,85 |
| 30 | $NHSO_2CH_3$ | OMe | H | 6,31 |

| | | | | |
|---|---|---|---|---|
| 31 | $N(SO_2CH_3)_2$ | H | H | 8,26 |
| 32 | $N(SO_2CH_3)_2$ | OMe | H | 6,41 |
| 33 | $NHCONHCH_2Ph$ | H | H | 7,08 |
| 34 | $NHCONHCH_2Ph$ | OMe | H | 7,19 |
| 35 | $NHCONHCH_2Ph$ | H | $NO_2$ | 7,20 |
| 36 | NHCONHCOPh | H | H | 5,89 |
| 37 | $NHCONH\text{-}C_6H_4\text{-}3I$ | H | H | 6,02 |
| 38 | $OCH_2Ph$ | H | H | 7,68 |
| 39 | $OCH_2Ph$ | OMe | H | 8,19 |
| 40 | OH | $NO_2$ | H | 9,22 |

**Group 2: pyrazolo-quinoline derivatives**

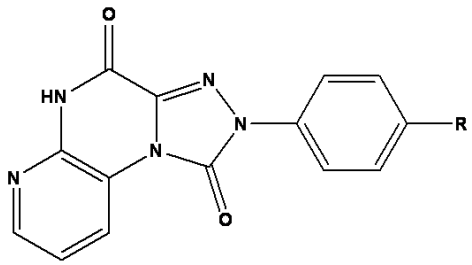



| Ligand | R | pKi |
|---|---|---|
| 41 | Ph | 7,51 |
| 42 | $C_6H_4$-3-Me | 8,30 |
| 43 | $C_6H_4$-4-Me | 8,49 |
| 44 | $C_6H_4$-4-OMe | 8,49 |
| 45 | $C_6H_4$-3-OMe | 8,89 |
| 46 | $C_6H_4\text{-}4\text{-}NO_2$ | 7,07 |
| 47 | $CH_2Ph$ | 7,13 |

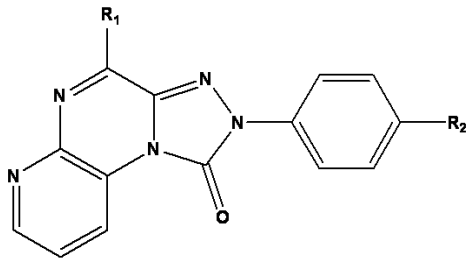



| Ligand | $R_1$ | $R_2$ | pKi |
|---|---|---|---|
| 48 | $NH_2$ | H | 6,26 |
| 49 | $NH_2$ | 3-Me | 7,00 |
| 50 | $NH_2$ | 4-Me | 6,73 |
| 51 | $NH_2$ | 4-OMe | 7,04 |
| 52 | $NH_2$ | 3-OMe | 6,64 |
| 53 | NHCOMe | H | 7,32 |
| 54 | NHCOMe | 3-Me | 7,51 |
| 55 | NHCOMe | 4-Me | 6,91 |
| 56 | NHCOMe | 4-OMe | 6,99 |
| 57 | NHCOPh | H | 8,68 |
| 58 | NHCOPh | 3-Me | 8,37 |
| 59 | NHCOPh | 4-Me | 8,36 |
| 60 | NHCOPh | 4-OMe | 8,47 |
| 61 | $NHCOCH_2Ph$ | H | 8,00 |
| 62 | $NHCOCH_2Ph$ | 3-Me | 8,41 |
| 63 | $NHCOCH_2Ph$ | 4-Me | 8,25 |
| 64 | $NHCOCH_2Ph$ | 4-OMe | 8,35 |
| 65 | $NHCOCHPh_2$ | H | 8,00 |
| 66 | $NHCOCHPh_2$ | 3-Me | 8,41 |
| 67 | $NHCOCHPh_2$ | 4-Me | 8,25 |
| 68 | $NHCOCHPh_2$ | 4-OMe | 8,35 |
| 69 | $NHCONHCH_2Ph$ | H | 8,08 |
| 70 | $NHCONHCH_2Ph$ | 3-Me | 8,47 |
| 71 | $NHCONHCH_2Ph$ | 4-Me | 6,59 |
| 72 | $N(COPh)_2$ | H | 8,21 |
| 73 | $N(COPh)_2$ | 3-Me | 7,63 |
| 74 | $N(COPh)_2$ | 4-Me | 7,52 |
| 75 | $N(COPh)_2$ | 4-OMe | 7,76 |

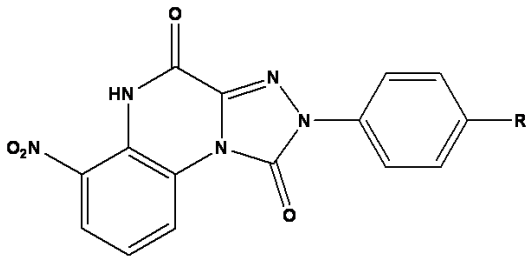
**Group 3: annelated triazolo-pyrazine derivatives**



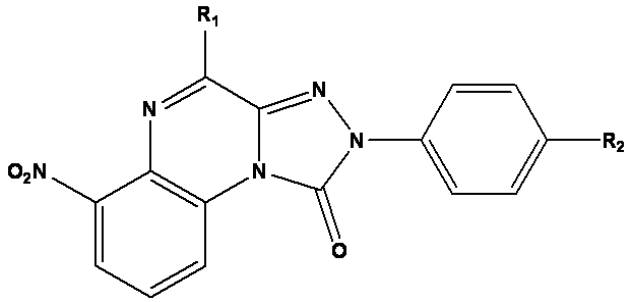
| Ligand | R | pKi |
|---|---|---|
| 76 | H | 6,60 |
| 77 | OMe | 8,48 |
| 78 | F | 6,23 |



| Ligand | $R_1$ | $R_2$ | pKi |
|---|---|---|---|
| 79 | $NH_2$ | H | 8,51 |
| 80 | $NH_2$ | OMe | 6,80 |
| 81 | $NH_2$ | OH | 5,87 |
| 82 | $NH_2$ | F | 6,31 |
| 83 | $NHC_6H_{11}$ | H | 7,81 |
| 84 | $NHC_5H_9$ | H | 8,08 |
| 85 | NHCOMe | H | 6,86 |
| 86 | NHCOPh | H | 7,15 |
| 87 | $NHCOCH_2Ph$ | H | 7,93 |
| 88 | NHCOMe | OMe | 7,39 |
| 89 | NHCOPh | OMe | 8,34 |
| 90 | $N(COPh)_2$ | H | 6,47 |
| 91 | $N(COPh)_2$ | OMe | 8,11 |
| 92 | H | H | 6,18 |
| 93 | H | OMe | 7,24 |

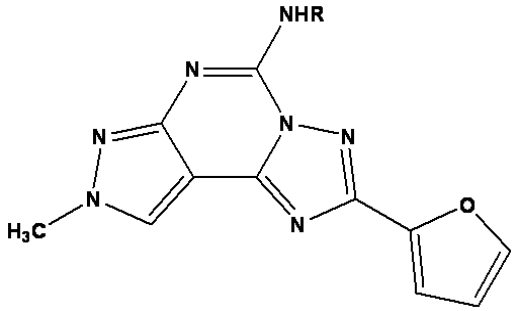



| Ligand | R | pKi |
|---|---|---|
| 94 | H | 6,55 |
| 95 | OMe | 8,33 |

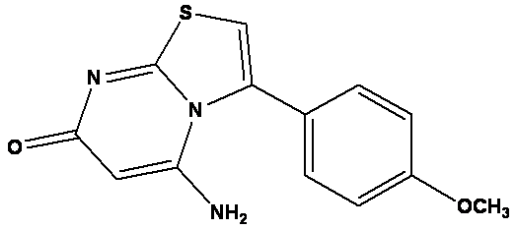



| Ligand | $R_1$ | $R_2$ | pKi |
|---|---|---|---|
| 96 | H | H | 8,32 |
| 97 | H | OMe | 7,33 |
| 98 | $NHC_6H_{11}$ | H | 6,55 |
| 99 | $NHC_5H_9$ | H | 6,94 |
| 100 | NHCOPh | H | 7,66 |
| 101 | NHCOPh | OMe | 6,66 |
| 102 | $N(COPh)_2$ | OMe | 6,46 |

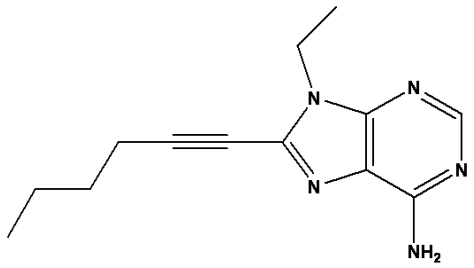
**Group 4: heterocyclic derivatives**



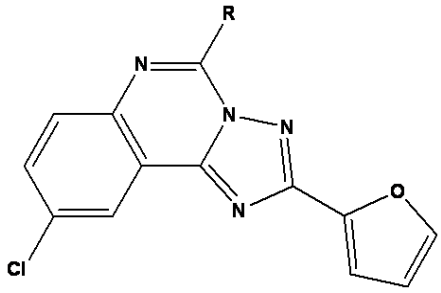

| Ligand | R | pKi |
|---|---|---|
| 103 | CONHPh | 8,74 |
| 104 | COCH-$Ph_2$ | 9,04 |
| 105 | COPh | 7,80 |
| 106 | $SO_2$-Ph | 6,13 |



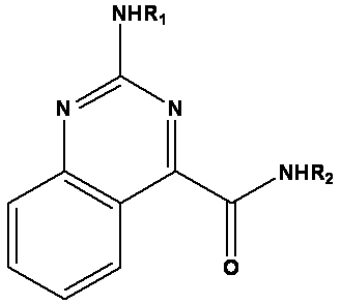

| Ligand | pKi |
|---|---|
| 107 | 7,74 |



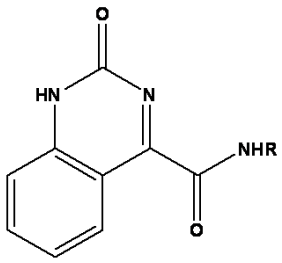

| Ligand | pKi |
|---|---|
| 108 | 6,20 |



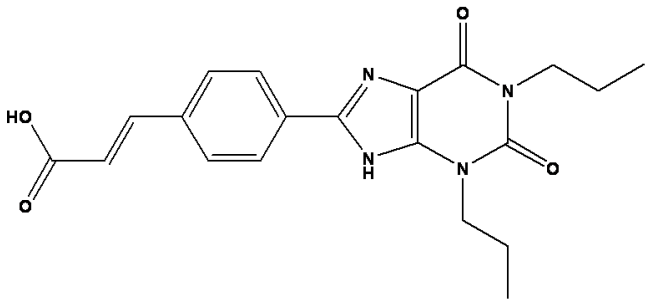

| Ligand | R | pKi |
|---|---|---|
| 109 | $NH_2$ | 7,85 |
| 110 | $NHCOCH_2Ph$ | 9,19 |




| Ligand | $R_1$ | $R_2$ | pKi |
|---|---|---|---|
| 111 | H | $C_6H_4$-4-OMe | 7,06 |
| 112 | H | $C_6H_5$ | 6,46 |
| 113 | H | $C_6H_4$-4-Me | 7,01 |
| 114 | H | $C_6H_4$-4-Br | 6,26 |
| 115 | $COCH_3$ | Ph | 7,60 |
| 116 | COPh | Ph | 6,74 |




| Ligand | R | pKi |
|---|---|---|
| 117 | $C_6H_4$-4-OMe | 7,71 |
| 118 | $C_6H_5$ | 7,30 |
| 119 | $C_6H_4$-4-Me | 7,57 |
| 120 | $C_6H_4$-4-Br | 7,57 |




| Ligand | pKi |
|---|---|
| 121 | 7,27 |