



**HAL**  
open science

## Ligand selection from the analysis of protein conformational substates: new leads targeting the N-terminal domain of Hsp90

Alessandro Genoni, Pennati Marzia, Morra Giulia, Zaffaroni Nadia, Colombo Giorgio

### ► To cite this version:

Alessandro Genoni, Pennati Marzia, Morra Giulia, Zaffaroni Nadia, Colombo Giorgio. Ligand selection from the analysis of protein conformational substates: new leads targeting the N-terminal domain of Hsp90. RSC Advances, 2012, 2, pp.4268-4282. 10.1039/C2RA00911K . hal-02196507

**HAL Id: hal-02196507**

**<https://hal.univ-lorraine.fr/hal-02196507>**

Submitted on 27 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Ligand Selection from the Analysis of Protein Conformational Substates:

### New Leads Targeting the N-Terminal Domain of Hsp90

Alessandro Genoni <sup>(1,2)</sup>, Marzia Pennati <sup>(3)</sup>, Giulia Morra <sup>(1)</sup>, Nadia Zaffaroni <sup>(3)</sup> and Giorgio Colombo <sup>(1)\*</sup>

(1) Istituto di Chimica del Riconoscimento Molecolare, CNR, Via Mario Bianco 9, 20131 Milano, Italy.

(2) Equipe de Chimie et Biochimie Théoriques, UMR 7565 SRSMC – CNRS, **Université de Lorraine**, Boulevard des Aiguillettes BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex, France.

(3) Department of Experimental Oncology and Molecular Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Via Amadeo 42, 20133 Milano, Italy.

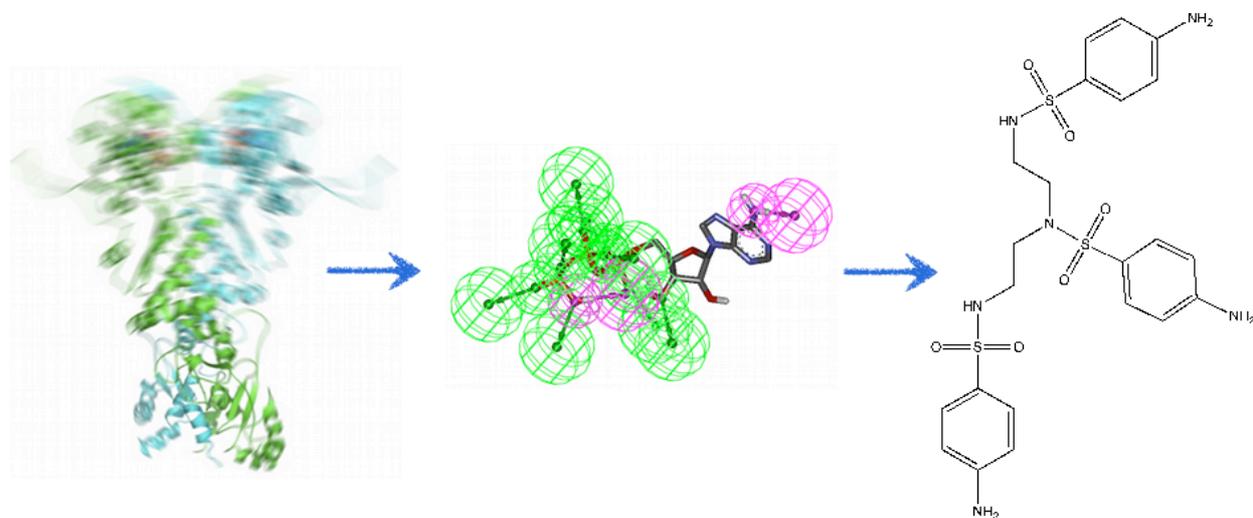
---

\* Correspondence to:

Giorgio Colombo, Istituto di Chimica del Riconoscimento Molecolare, CNR, Via Mario Bianco 9, 20131 Milano, Italy; E-mail: [giorgio.colombo@icrm.cnr.it](mailto:giorgio.colombo@icrm.cnr.it); Phone: +39-02-28500031; Fax: +39-02-28901239.

## Table of contents entry

The conformational dynamics of the complex between the molecular chaperone Hsp90 and its natural ligand ATP is fully considered to identify new possible N-terminal inhibitors.



## **Abstract**

The dynamic properties of proteins underlie every aspect of their functions in the cell. The atomistic description of protein motions and their inclusion in ligand selection processes may provide new opportunities for hit identification and drug discovery. Herein, we present a novel rational strategy that allowed us to computationally select new N-terminal targeted inhibitors of the molecular chaperone Hsp90 starting from the atomistic analysis of the conformational dynamics of the complex between the protein and its natural ligand ATP. First of all, we have identified the relevant representatives of distinct conformational substates of the Hsp90-ATP complex through the application of a novel structural clustering strategy and, for each of them, we have afterwards characterized the nucleotide-protein interactions to build a pharmacophore model recapitulating the binding hotspots conserved in different ensembles of protein conformations. The resulting pharmacophore has been finally used to screen a database of small molecules and allowed us to identify novel drug-like molecules with interesting activities against Hsp90 functions in experimental models of cancer cells. The results and the experimental validation of the selected molecules provide support to the feasibility of including protein flexibility in drug selection strategies through the characterization of relevant substates.

## Introduction

Proteins are inherently flexible systems in constant motion among different conformational substates and their dynamic properties are frequently associated with molecular recognition, function and catalysis. However, internal motions and the concept of conformational substates have often been ignored in classical structure-based drug discovery despite recent successes in the design of small-molecule inhibitors<sup>1-7</sup>.

The role of functional dynamics has often been conceptualized in terms of a model in which the protein pre-exists in an ensemble of energetically similar conformations. In this context, conformational selection is defined as a process in which a ligand selectively binds to one (or a subset) of these conformers, stabilizing it and increasing its relative population with respect to the total protein population, eventually resulting in the final observed complex. The concepts of ligand-induced shifts in protein conformational dynamics can be usefully exploited in the context of drug design and discovery. In fact, in some cases ligand binding induces specific conformational states that might not be anticipated from the simple analysis of the protein apo structure and even result in major structural transitions.

One notable instance of ligand-regulated protein functional dynamics is represented by Heat shock protein 90 (Hsp90). Hsp90 is an ATPase-directed molecular chaperone, operating in a complex cycle driven by ATP binding, hydrolysis, and ATP/ADP exchange at the N-terminal nucleotide-binding site<sup>8</sup>. The protein is a homodimer in which each protomer is characterized by a modular architecture with three domains: an N-terminal regulatory Domain (NTD), responsible for ATP binding, a Middle Domain (M-domain), which completes the ATPase site necessary for ATP hydrolysis and also binds client proteins, and a C-terminal Domain (CTD)<sup>9</sup> required for dimerization. Crystal structures, solution and biochemical studies suggest a role for the nucleotide in selecting and stabilizing different conformations of Hsp90, related to specific functions in the chaperone cycle<sup>10, 11</sup>. Overall, the mechanism of conformational coupling to the ATPase cycle appears to involve a “tense” and closed conformational state of Hsp90 upon ATP binding, whereas

subsequent hydrolysis to ADP leads to a more “relaxed” and structurally flexible state of Hsp90. Finally, in the nucleotide-free form, the dimer moves to an “open” state<sup>9</sup>. Given its importance in cancer development and maintenance, Hsp90 inhibition has been intensely pursued in recent years. Thirteen inhibitors, all targeting the N-terminal ATP-binding pocket, have entered clinical trials<sup>12</sup>. In this paper, we present a novel rational strategy for the computational discovery of N-terminal targeted inhibitors of Hsp90, which is based on the explicit inclusion of the flexibility and conformational dynamics of the active complex between the chaperone and its natural ligand ATP. Our approach starts from the identification of the relevant representatives of distinct conformational substates visited during a long timescale molecular dynamics (MD) simulation of the complex between full length Hsp90 and ATP. The analysis of the interactions between the nucleotide and the binding site and their fluctuations within different substates is used to build a pharmacophore model in which the relative importance of each pharmacophoric feature is weighted according to the conservation of the interaction it mimics in the different substates. The resulting pharmacophore takes into account in a simple and compact manner the effects of protein mobility on ligand binding and the reciprocal effect of the nucleotide on the internal dynamics of the protein. Afterwards, the pharmacophore is used to screen a database of small molecules with the aim of identifying novel antagonists of the Hsp90 chaperone function with possible anticancer activity. **It is important to underline that the pharmacophore model proposed here relies mainly on mimicking hydrogen bond interactions within the active site. Since this approximation might lead to the discovery of ligands that may have non-optimal drug-like properties, we have alleviated this problem by filtering the first selection of compounds through the application of the set of rules introduced by Lipinski<sup>13</sup>.** The results and the experimental validation of the activities of the selected molecules provide support to the feasibility of including flexibility in drug selection strategies and give a vivid illustration of how the knowledge of the internal dynamics of a system can be exploited to expand the molecular diversity space of possible ligands.

## Material and Methods

*Molecular Dynamics studies for the selection of relevant ATP-induced substates.* The MD simulation trajectories used in this work were obtained as already described in a previously published paper<sup>14</sup>. The details and the full description of the MD set up and runs can thus be found in that published paper, which deals with the characterization of the ligand-modulation of the Hsp90 internal dynamics<sup>14</sup>.

Briefly, the crystal structure<sup>9</sup> (pdb entry 2CG9) containing Yeast Hsp90 dimer bound to ATP was employed as a starting point for the simulations. The system was solvated in a tetrahedral solvation box containing around 57000 particles. The simulations and the analysis of the trajectories were performed using the GROMACS software package<sup>15</sup> using the GROMOS96 force field<sup>16</sup> and the SPC water model<sup>17</sup>. In order to remove possible bad contacts in the initial structure, the ATP-bound Hsp90 dimer system was initially optimized by means of a 2000-step steepest descent energy minimization and afterwards relaxed by a further 2000-step conjugate gradient energy optimization. The system was then equilibrated by a 50 ps MD run with position restraints on the protein and ligand to allow the relaxation of the solvent molecules, equilibration that was followed by another 50 ps MD run without position restraints on the solute. It is worthwhile to note that all the trajectory equilibrations were checked by monitoring quantities such as the RMSD with respect to the initial structure, the internal protein energy and the fluctuations calculated on different time-intervals. Furthermore, in all the performed MD simulations we took into account the long-range Coulomb interactions exploiting the particle mesh Ewald summation method<sup>18, 19</sup>, we used the LINCS algorithm<sup>20</sup> to constrain all the bond lengths and the SETTLE program to deal with the water molecules<sup>21</sup>. A dielectric permittivity,  $\epsilon = 1$ , and a time step of 2 fs were also used. All atoms were given an initial velocity obtained from a Maxwellian distribution at the desired initial temperature of 300K. The density of the system was adjusted performing the first equilibration runs at NPT condition by weak coupling to a bath of constant pressure ( $P_0 = 1$  bar, coupling time  $\tau_P = 0.5$  ps)<sup>22</sup>. In all the simulations the temperature was maintained close to the intended values by weak coupling to

an external temperature bath<sup>22</sup> with a coupling constant of 0.1 ps. The proteins and the rest of the system were coupled separately to the temperature bath.

The production MD simulation was performed for 170 ns, but the first 5 ns were not considered in order to minimize convergence artefacts. The resulting trajectory was afterwards analyzed and partitioned according to the clustering procedure that has been recently introduced by Micheletti and coworkers<sup>23</sup> and that will be briefly described in the following subsection.

***Clustering of the trajectory for the definition of functional substates.*** The clustering procedure adopted in our work to identify the relevant representatives of distinct conformational substates from an MD simulation is based on the  $K$ -medoids clustering scheme<sup>24</sup>. This technique considers as input only the pairwise RMSD distances between all the pairs of recorded structures and provides a pre-assigned number  $K$  of non-empty and homogeneous clusters, each of them characterized by a representative conformation. The method identifies the clusters iteratively minimizing the “dissimilarity score” obtained by summing the RMSD of each structure from its cluster representative and, in particular, consists in the following three steps:

- 1) random assignment of the members to the  $K$  clusters;
- 2) identification of the representative for each cluster as the element with the smallest total distance from the other cluster members;
- 3) redefinition of the clusters assigning each structure to the closest representative.

Steps 2 and 3 are iterated until the “dissimilarity score” converges to a minimum, but it is worthwhile to note that, in order to avoid trapping in local minima, the whole procedure is consistently repeated considering several starting random clusters.

It is obvious that the standard  $K$ -medoids technique<sup>24</sup> relies only on geometrical considerations (namely, the RMSD distance of any pairs of structures), but it does not take into account any time-information, such as the time-succession of the recorded conformations in the examined trajectory. Time-information is fundamental to partition the trajectory in exam in a series of progressively

visited substates and, therefore, in order to overcome this drawback Micheletti and coworkers have recently devised a modified  $K$ -medoids scheme to ensure that all the structures assigned to a cluster span a continuous time-interval<sup>23</sup>. Of course, the time-continuity constraint simplifies the definition of the cluster members because they are simply specified by  $K-1$  time-partitions of the trajectory. In particular, the new modified technique starts with an equally spaced  $K-1$  subdivision of the simulation time-length and the representative of each cluster is afterwards determined as in step 2 of the standard technique. After calculating the associated “dissimilarity score”, a new clustering is proposed by randomly reassigning one or more of the  $K-1$  subdivisions with the only constraint that there are not identical subdivisions that would result in empty clusters. At this point the new cluster representatives can be identified and a new “dissimilarity score” can be computed. Following a Monte Carlo philosophy, the proposed clustering is accepted if it is associated with a lower “dissimilarity score”, otherwise the previous one is retained and another new partition is randomly proposed. The procedure is usually repeated until convergence of the “dissimilarity score”.

***Dynamic Pharmacophore Design.*** In order to design a dynamic pharmacophore, first of all, we have performed the analysis of the conserved hydrogen bonds in all the relevant states of the Hsp90-ATP complexes, selected according to the clustering procedure described above. This analysis has been performed exploiting the `g_hbond` utility of the GROMACS simulation package<sup>15</sup>. We assumed the following geometric criteria to identify the hydrogen bonds: 1) distance between the hydrogen atom and the acceptor atom lower than 3.5 Å; 2) angle formed by the donor atom, the hydrogen atom and the acceptor atom greater than 120°; 3) angle formed by the hydrogen atom, the acceptor atom and the acceptor-antecedent atom (namely, an atom one covalent bond away from the acceptor) greater than 90°. As we will explain in the “Results” section, the hydrogen bonds analysis allowed us to construct the pharmacophore ATP-PHARM-1 model (see Figure 2) using the module to create pharmacophore queries in the DISCOVERY STUDIO 2.5 package<sup>25</sup>. In particular we created a ten-point model by assembling eight hydrogen bond acceptors mapped over

the OB, O5, O8, OA, O9, OC, O3 and O6 atoms of the ATP molecule (see Figure 1) and two hydrogen bond donors mapped over the ATP groups OC-H5 and N4-H1 (see Figure 1). Furthermore, we assigned a weight to all these pharmacophore features according to the overall occurrence of the corresponding hydrogen bond descriptors in the selected Hsp90-ATP complexes (see Tables 1 and 2).

***Small-Molecule database screening and hit identification.*** The obtained pharmacophore model has been considered as query for a search in the NCI\_3D database. This search has been performed by means of the DISCOVERY STUDIO 2.5 package setting the “SEARCH” method equal to “BEST” and the “LIMIT” option equal to “BEST N”. This returned 294 ligands that have been further processed accepting only those compounds fulfilling three of the four Lipinski rules<sup>13</sup>. In fact, for the sake of completeness, it is important to remind that the Lipinski rule of five<sup>13</sup> states that a molecule, to potentially be an orally active drug, must fulfill three of the following four criteria: a) no more than five hydrogen bond donors; b) no more than 10 hydrogen bond acceptors; c) a molecular mass lower than 500 daltons; d) an octanol-water partition coefficient LogP not greater than 5. Actually, it is worthwhile to observe that we have used modified Lipinski rules since the desired molecules have been required to have a mass not greater than 550 daltons instead of 500 daltons. The screening using the modified Lipinski rules has been performed using the InstantJChem software<sup>26</sup> and returned 25 compounds.

***Small Molecules.*** Compounds for this study were kindly provided by the National Cancer Institute Repository (Bethesda, MD, USA). The 17-(Allylamino)-17-demethoxygeldanamycin (17-AAG) was purchased from Sigma Aldrich (St. Louis, MO, USA). To obtain a 100 mM stock solution of the molecules, a variable amount of a solution of 1% dimethyl sulfoxide (DMSO) made in sterile water was added to each powdered stock and the compound was completely dissolved, stored at -

20°C, and diluted in complete culture medium, at the appropriate working solution, immediately before use.

**Cell Culture Conditions.** Mycoplasma-free human breast cancer cell lines (BT-474, HCC1937, MCF-7, MDA-MB-231, T-47D and ZR75.1) were obtained from American Type Culture Collection (Rockville, MD, USA). The HCC1937, ZR75.1, MCF-7, T-47D and BT-474 cells were grown in RPMI-1640 medium (Lonza Group Ltd, Basel, Switzerland) supplemented with 10% heat-inactivated fetal bovine serum; the MDA-MB-231 cells were cultured using D-MEM/F12 medium (Lonza Group Ltd) supplemented with 5% heat-inactivated fetal bovine serum. Cells were maintained in the logarithmic growth phase as a monolayer at 37°C with a supply of 5% CO<sub>2</sub>/95% air atmosphere.

**Evaluation of the Antiproliferative Potential of Small Molecules.**  $4 \times 10^3$  cells were plated in 96-well flat-bottomed microtiter plates for 24 h and then treated with increasing concentrations of different compounds (0.01-50  $\mu$ M) for 72 h. Control cells received vehicle alone (1% DMSO). All studies were performed in eight replicates and repeated at least three times independently. At the end of drug exposure, the antiproliferative potential was determined with the sulforhodamine B (SRB) protein staining method as previously reported<sup>27</sup>. Briefly, cells were fixed with cold 50% trichloroacetic acid solution (Sigma Aldrich) and stained with 0.4% SRB (Sigma Aldrich) dissolved in 1% acetic acid (Sigma Aldrich) for 30 min. The plates were then washed four times with 1% acetic acid to remove unbound stain and air-dried. The dye was then dissolved by 10 mM Tris base (pH 10.5). Optical density was read at 550 nm on a microplate reader and the results were expressed as a percentage, relative to DMSO-treated cells. Dose-response curves were created and IC<sub>50</sub> and IC<sub>80</sub> values (i.e., concentrations able to inhibit cell growth by 50% and 80%, respectively) were determined graphically from the curve for each compound.

***Analysis of Hsp90 Chaperone Function.*** To monitor changes in Hsp90 client proteins, at the end of each treatment, cells were scraped, pelleted, washed twice with PBS, and then solubilized in lysis buffer (0.01% NP40, 10 mM Tris [pH 7.5], 50 mM KCl, 5 mM MgCl<sub>2</sub>, 2 mM DTT, 20% glycerol plus protease inhibitors). Samples containing 40 µg of protein were resolved on precast 12% NuPAGE™ bis-tris gels (Invitrogen Life Technology, Camarillo, CA, USA), and were transferred onto Hybond ECL nitrocellulose membranes (GE Healthcare, Chalfont St. Giles, UK) using the NuPAGE transfer buffer (Invitrogen Life Technology). Nitrocellulose membranes were blocked in PBS-Tween 20 with 5% skim milk, first incubated with the primary antibodies specific for survivin (ab469) (Abcam Inc., Cambridge, UK), Hsp90 (sc7947), Hsp70 (sc24), CDK4 (sc136241), CDK6 (sc271364), (Santa Cruz Biotechnology, Santa Cruz, CA), and AKT (9272) (Cell Signaling Technology, Danvers, MA) and then with the secondary peroxidase-linked whole antibodies (GE Healthcare). Bound antibodies were detected using the SuperSignals West PICO chemiluminescent substrate (Pierce Biotechnology Inc., Rockford, IL, USA). β-Actin monoclonal antibody (Abcam Inc.) were used to confirm equal protein loading on the gel.

***Cell Cycle Phase Distribution and Apoptosis Analysis.*** After treatment with small molecules, both adherent and floating cells were fixed in 70% ethanol. Fixed cells were stained for 30 min with propidium iodide (50 µg/ml) and RNase (50 mg/ml) in PBS buffer. Samples were analyzed with a FACScan flow cytometer (Becton Dickinson, Sunnyvale, CA, USA) with an excitation wavelength of 488 nm and an emission wavelength of 585 nm. At least 30,000 events were read and histograms were analyzed using the CellQuest software according to the Modfit model (Becton Dickinson). For apoptosis analysis, an aliquot of propidium-stained cells were observed for nuclear morphology of apoptosis (chromatin condensation and DNA fragmentation) by fluorescence microscopy. The percentage of cells with an apoptotic nuclear morphology was determined by scoring at least 500 cells in each sample.

**Statistical Analysis.** Two-tailed Student's t test was used for statistical analysis. *P* values <0.05 were considered statistically significant.

**Analysis of the molecular interactions established by the most promising hits with Hsp90.** In order to gain an insight into the observed Hsp90 inhibitory/cellular activity of some of the most promising hits that resulted from the performed experimental tests, we have analyzed the molecular interactions that they establish with Hsp90. To accomplish this task we have docked each considered compound into the ATP binding site of all the Hsp90 conformations previously taken into account for the pharmacophore construction (namely, fourteen structures). In fact, in the light of what already stressed above, considering only one conformer would be rather limiting and we believe that examining several conformations would help us to unravel the way in which an inhibitor adapts to / influences the receptor dynamics. For the sake of completeness, it is worth to mention that all the docking calculations were performed using the package AutoDock 4.2<sup>28</sup>. In particular, grid maps were generated by the program AutoGrid with a 0.375 Å spacing, with a grid center corresponding to the position of the ATP-C7 atom (see Figure 1) in the Hsp90-ATP complex of the dynamic substate in exam (namely, the grid was centered over the ATP binding site) and with 55 points in each directions (namely, 166,375 points totally). Polar hydrogen atoms and Gasteiger charges were properly taken into account for the receptor and, furthermore, the Lennard-Jones parameters 12-10 and 12-6 were used to model hydrogen bonds and van der Waals interactions, respectively. Considering the real docking calculations, random orientations and torsions were used for each ligand and the number of active torsions was set equal to the maximum number of rotatable bonds. The Lamarckian genetic algorithm was used with the default docking parameters. Moreover, for each run, the population size, the maximum number of energy evaluations and the maximum number of generations were set equal to 150, 50 millions and 27,000, respectively. A total of 250 runs were performed. At the end of the docking runs, the poses of the ligand were listed in an increasing energy order and the ligand conformation with the lowest binding energy was afterwards

considered as reference. All the poses with an rmsd value lower than 2.0 Å from the reference structure were assigned to the first cluster and it is important to note that after a conformation is assigned to a cluster it was not used again for other energetically less-favorable clusters. The procedure was repeated for all the remaining unclassified structures until all of them were put in a cluster.

After each docking calculation, we have extracted the complex corresponding to the lowest binding energy pose and we have analyzed the associated intermolecular interactions. Repeating this protocol for all the considered receptor conformations we were able to perform a statistical analysis of the Hsp90-inhibitor contacts and, finally, we detected those interactions that conserve and stabilize most the complex also considering the protein fluctuations (see the “Results” section for more details).

## **Results**

***Identification of relevant conformational substates from the dynamics of the Hsp90-ATP complex.*** The conformational dynamics of Hsp90 in complex with ATP, representing the activated state of the chaperone, was characterized by a long time scale (170 ns) all-atom Molecular Dynamics simulation as already described above and by Morra *et al.*<sup>14</sup>. The trajectory was afterwards analyzed with a novel structural clustering scheme developed by Micheletti and coworkers<sup>23</sup> to identify the relevant visited conformational substates. The details of the clustering procedure and of the identification of the substates are reported in the “Materials and Methods” section. The aim of this step was to select a minimal number of key configurations that capture most of the structural fluctuations encountered in the multi-nanosecond trajectory, recapitulating the flexibility of the complex and providing a compact representation of how internal dynamics reverberates on the types and persistence of intermolecular interactions between Hsp90 and ATP in the N-terminal domain.

It is important to underline that the structural clustering strategy used herein accounts for the time-order of the structures, which is essential for partitioning the trajectory in a succession of progressively-visited and coherent substates. The time-continuity of the members of the different clusters is instrumental to simplify the definition of the different homogeneous structural sets of the Hsp90-ATP complex that have been afterwards used as templates for the pharmacophore design.

The clustering was performed by varying  $K$  from 2 to 15 and enforcing the time continuity of the clustered structures, according to the scheme described in the “Materials and Methods” section. The algorithm returned a subdivision of the trajectory into a fixed number  $K$  of non-overlapping intervals (clusters) with maximal internal structural homogeneity. An optimal subdivision was obtained using seven (namely,  $K = 7$ ) non-overlapping intervals and it was consequently used to identify the most prominent conformational substates explored by the trajectory.

During the free dynamic evolution, the chaperone-nucleotide complex populates seven distinct conformational substates with residence times from 21.2 to 30.2 ns. The average RMSD between pairs of structures within each identified substate resulted 2 Å. In contrast, the ensemble of different conformers is structurally heterogeneous (with intersubstates differences up to 5 Å), reflecting the well-known flexibility that characterizes the members of the Hsp90 family.

***Dynamic Pharmacophore Design.*** As briefly mentioned in the introduction, in order to design a dynamic pharmacophore model for the discovery of new possible Hsp90 inhibitors, we have identified the most recurring fundamental interactions between Hsp90 and the ATP molecule, assigning them a weight according to the frequency with which they occur.

In order to accomplish this task, first of all, we have assumed that the intermolecular interactions needed to stabilize the Hsp90-ATP complexes are mainly hydrogen bond-like interactions, which are directional interactions typically found also in drug-like molecules, and, therefore, we have selected all the possible hydrogen bond acceptors and donors (indicated also as potential hydrogen bond descriptors) in ATP (see Tables 1-2 and Figure 1). We have afterwards considered the seven

key configurations that we have previously selected by means of the clustering procedure described above, each corresponding to two protein-nucleotide complexes. For all of these complexes we have determined which are the potential hydrogen bond descriptors able to establish significant interactions in the ATP binding site, namely, those potential ATP hydrogen bond acceptors and donors that form effective hydrogen bonds with Hsp90 (see the “Materials and Methods” section for the details). We have thus determined the number of times each ATP hydrogen bond descriptor plays a fundamental role in the substrate binding (see Tables 1 and 2) and we have afterwards divided these numbers by the number of analyzed Hsp90-ATP complexes (fourteen) in order to obtain the relative weights of all the hydrogen bond acceptors and donors in the pharmacophore definition (see Tables 1 and 2). Nevertheless, the ATP hydrogen bond descriptors with a weight lower than 0.4 have been discarded and this has led to the construction of the pharmacophore model labeled as ATP-PHARM-1 and represented in Figure 2.

It is worth noting that a pharmacophore model such as the one proposed here, namely a model mainly based on reproducing hydrogen bond interactions, may select ligands with charged functionalities, a feature that potentially provides leads that are poor drug candidates. On the other hand, this rationally designed model may provide useful information for the selection of potentially active molecules with new structural and stereoelectronic properties. Moreover, given the directionality of the hydrogen bond interactions used in the pharmacophore, selected hits may display an optimal 3D organization of the functional groups needed to productively bind the ATP site of Hsp90. To evaluate the potentials and pitfalls of this model, an analysis of how known inhibitors fit to or diverge from our pharmacophore constraints was carried out. Therefore, we considered the crystal structures of the Hsp90 N-terminal domain with some known inhibitors: geldanamycin<sup>29</sup>, which is included in the NCI database, radicicol<sup>30</sup>, purine scaffold-based inhibitors<sup>31</sup> (PU-H71, PU-H64 and PU24FCI) and a benzamide tetrahydro-4H-carbazol-4-one analog that represents a Serenex-type compound<sup>32</sup>. From this analysis it results that whereas radicicol, PU-H71, PU-H64 and the Serenex-like inhibitor mainly establish only one important direct hydrogen

bond with the residue Asp 79 (from now on we will use the residue numbering of the 2CG9 pdb file), the network of direct hydrogen bonds is much more extended for geldanamycin and PU24FCI. In particular, geldanamycin forms a hydrogen bond with Asp 79 through its carbamate amino group, which mimics the exocyclic  $\text{-NH}_2$  group of the ATP/ADP adenine, two crucial hydrogen bonds with Lys 44 by means of the oxygen atom of its hydroxyl group and the oxygen atom of the “C29” methoxy group<sup>29</sup>, two hydrogen bonds with Gly 123 and Phe 124 through the ansa-ring carbonyl oxygen, which is the analog of the ATP/ADP  $\alpha$ -phosphate group, and, finally, a hydrogen bond with Lys 98 by means of the benzoquinone oxygen atom, which mimics the ATP/ADP  $\beta$ -phosphate group. For the sake of completeness, we observe that PU24FCI forms two hydrogen bonds with Tyr 125 in addition to the fundamental hydrogen bond with Asp 79. From these last two examples, and especially from the geldanamycin case, we can conclude that our hydrogen bond based pharmacophore model partially reproduces the binding determinants of these two major inhibitors of Hsp90.

At this point, it is worth noting that one of the goals of our work was to select molecular entities that are different from known inhibitors (see the “Discussion” section). Moreover, the pharmacophore was designed using the concept of dynamics-based weighting the Hsp90-ATP interactions. As a consequence, it is not surprising that, compared to already known inhibitors, the model shows differences in the fine details of the molecular recognition determinants that it mimics.

***Small-Molecule database screening and hit identification.*** The new dynamic pharmacophore model was used to perform a screening search of the small molecule database NCI\_3D that approximately contains 160,000 compounds. The filtering with the pharmacophore returned 294 hits that were subject to a further screening **constraint consisting** in the fulfillment of three of the four Lipinski rules<sup>13</sup> (see the “Materials and Methods” section for the details). This resulted in 25 molecules (see Figures 3-5) that approximately correspond to the 0.016% of the reference database. **Despite the presence of charged functionalities, the overall properties of most of the selected**

compounds are consistent with drug-like features, since the second screening imposed the fulfillment of the Lipinski rules<sup>13</sup>. To better investigate this aspect we have also explicitly calculated the LogP values of the 25 hits (see Table 3). It is possible to observe that all the identified compounds fulfill the rule for the octanol-water partition coefficient (LogP lower than or equal to 5) even if most of them show negative values. However, a stricter extension of the Lipinski rules, which was proposed by Ghose and coworkers to better evaluate the druglikeness of molecules<sup>33</sup>, defines the LogP values as acceptable if they are between -0.4 and 5.6. As shown in Table 3, eight of our 25 hits fulfill the Ghose rule for the partition coefficient. More importantly, three members of this subset correspond to compounds that were shown to be active in experimental tests (see the next subsection).

For the sake of completeness, it is important to note that the screening with our pharmacophore model was not able to select geldanamycin, which is part of the NCI database. This is due to the fact that our pharmacophore captures only some of the binding determinants that characterize the macrocyclic inhibitor. Nevertheless, we want to stress again that one of the aims of our work was to develop and test a new pharmacophore design strategy allowing the discovery of new active molecules against Hsp90, possibly different from already identified active compounds.

***Experimental tests on the newly discovered Hsp90 inhibitors.*** Eight of the selected compounds (**2**, **4**, **5**, **10**, **11**, **19**, **20**, **23**) resulting from the virtual screening were obtained from the National Cancer Institute (NCI) and tested in a panel of human breast carcinoma cell lines. Cells were exposed for 72 h to increasing concentrations (from 0.5 to 50  $\mu$ M) of the small molecules, and the effect on cell proliferation was determined by SRB assay. Assessment of IC<sub>50</sub> values showed that only compounds **2**, **4**, **19** and **20** inhibited cell growth in a concentration-dependent manner. However, while compounds **2** and **19** exhibited a comparable potency in all tumor cell lines, compounds **4** and **20** induced a variable growth inhibitory effect in the different cell models (see Table 4). At tested drug concentrations, the IC<sub>50</sub> was not reached for the other compounds (**5**, **11**, **10** and **23**) (see Table

4). For comparative purposes, the antiproliferative effect of 17-AAG was also evaluated and IC<sub>50</sub> values ranging from 0.08±0.01 μM to 3.48±0.22 μM were observed for the different cell lines (see Table 4).

The cellular effects of Hsp90 are mediated in large part by its physical association with several client proteins, which contribute to tumor cell survival, including CDK4, CDK6, AKT, c-Raf, EGFR, HER2/neu, c-Kit<sup>34</sup>. In addition, the cytoprotective role of Hsp90 has been linked to its ability to modulate the conformation and activity of the antiapoptotic protein survivin<sup>35</sup>. In this context, we examined whether the cytotoxic effect exerted by compounds **2**, **4**, **19** and **20** was due to the breakdown of multiple cell survival pathways as a consequence of destabilization of Hsp90 client proteins. Specifically, the exposure of MCF-7 and ZR75.1 cells to IC<sub>50</sub> and IC<sub>80</sub> of compounds **2**, **4** and **19** resulted in a decrease in the expression levels of AKT, CDK4, CDK6, and survivin (see Figure 6). The treatment did not modify the Hsp90 expression levels, whereas an increase of Hsp70 protein abundance was appreciable (see Figure 6). By contrast, the cytotoxic effect exerted by compound **20** does not seem to rely on the interference with the activity of Hsp90. In fact, no alteration in the expression of Hsp90 client proteins was observed after treatment in both cell lines (see Figure 6).

On the basis of the above data, since some of the Hsp90 clients are key cell cycle regulators, the effect of Hsp90 inhibition on cell cycle phase distribution was assessed. Consistent with the inhibition of both CDK4 and CDK6, the exposure of asynchronously growing MCF-7 and ZR75.1 cells to compounds **2**, **4** and **19** induced a marked change in the distribution of the cells in the different phases of the cell cycle (Figure 7A). Specifically, flow cytometric analysis of propidium iodide-stained cells showed that a 72-h treatment with subcytotoxic, IC<sub>50</sub>, and IC<sub>80</sub> concentrations of compounds **2**, **4** and **19** resulted in a concentration-dependent accumulation of cells in G0/G1 phase (see Figure 7A). In addition, chromatin condensation and DNA fragmentation—which are common features of apoptosis— were detected by fluorescence microscopy after staining with propidium iodide of breast cancer cells treated with small molecules. The percentage of cells with

an apoptotic nuclear morphology, as determined on the overall cell population, increased in a concentration-dependent fashion after treatment with all compounds (see Figure 7B).

Notably, the effects on cell cycle phase distribution and induction of apoptosis observed after treatment with compounds **2**, **4** and **19** are very similar to those obtained after exposure of MCF-7 and ZR75.1 cells to 17-AAG (see Figure 7).

*Analysis of the molecular interactions established by the most promising hits with Hsp90.* The interactions of active hits with Hsp90 were analyzed in details by generating structural models that take into account different conformations of the receptor. In particular, we focused on compounds **2** and **19** since compound **4** shows interesting similarities with the ATP and ADP molecules (see Figure 3) and, therefore, its modes of interaction with the molecular chaperone can be easily anticipated. In fact, it is possible to observe that in compound **4** we have a cytosine unit, which can be considered a mimic for the adenine in ATP/ADP, and, above all, a ribose ring and a phosphate group, which have exact correspondences in the two nucleotides.

As described in details in the “Material and Methods” section, the other two promising hits (compounds **2** and **19**) were docked into the ATP binding site of all the Hsp90 conformations considered for the construction of our pharmacophore model. After each docking calculation we extracted the Hsp90-inhibitor complex associated with the lowest binding energy pose and we have afterwards analyzed the corresponding intermolecular interactions. As we will show below, this protocol has allowed us to indentify the most stabilizing interactions for the compounds in exam (frequency greater than or equal to five), taking into account also protein fluctuations. However, before analyzing the obtained results, it is worthwhile to point that compounds **2** and **19** are constituted by “quasi-equivalent” units (see Figures 8 and 9), which have led to degeneracy in the docking solutions. Therefore, in order to better rationalize the results of our calculations, we have grouped together the interactions established by “quasi-equivalent” functional groups. To accomplish this task we have previously defined the following groups for compound **2** (see Figure

8): the aromatic amino groups  $-N(1)H_2$ ,  $-N(27)H_2$  and  $-N(37)H_2$  (from now on indicated as *ARAM*), the aromatic rings (*AR*), the sulfonyl groups  $-S(8)O_2$ ,  $-S(19)O_2$  and  $-S(29)O_2$  (*SU*) and the aliphatic amino groups  $-N(11)H$  and  $-N(17)H$  (*ALAM*). In analogous way, for compound **19** we have (see Figure 9): the hydroxyl groups of the sugar-like units  $-O(1)H$ ,  $-O(30)H$ ,  $-O(36)H$  and  $-O(28)H$  (from now on indicated as *OG*), the oxygen heteroatoms of the sugar-like units  $O(7)$  and  $O(24)$  (*OE*), the sugar-like units (*SLU*) and the purine units (*PU*).

Let us consider the interactions established by compound **2**. From the analysis of Tables 5 and 6 it is easy to see that the sulfonyl groups of the molecule act as hydrogen bond acceptors and form recurring interactions with Asn 37, Asn 92, Gly 121, Val 122, Gly 123, Phe 124 and Tyr 125, while the aromatic amino groups are hydrogen bonds donors and mainly interact with Asp 79 and Asn 92. It is interesting to note that, as discussed above, the interaction with Asp 79 is a recurring and fundamental hydrogen bond in many complexes between Hsp90 and known inhibitors<sup>29-32, 36</sup>, whereas the interactions (interaction) with Gly 123 and Phe 124 (Tyr 125) have (has) been already observed for geldanamycin (PU24FC1)<sup>29, 31</sup>. Taking into account the hydrophobic contacts (see Table 7), the aromatic rings establish significant interactions with Ala 41, Ile 82, Met 84, Leu 93, Met 116, Phe 124, Val 136 and Leu 173, residues that are involved in hydrophobic contacts also with the already known inhibitors mentioned above<sup>29-32</sup>. For instance, the aromatic ring face of radicicol forms extensive hydrophobic interactions with the side chains of Met 84, Leu 93, Phe 124, Thr171 and Leu 173<sup>30</sup> while PU-H71 establishes hydrophobic contacts with Ala 41, Met 84 Phe 124 and Trp 148<sup>31</sup>.

In analogous way, for compound **19** we have observed that the oxygen atoms of the hydroxyl groups act as hydrogen bond acceptors with Asn 37, Gly 83, Phe 120 and Gly 123, whereas the oxygen heteroatoms of the sugar-like units accept hydrogen bonds from the side chain of Gln 119 (see Table 8). Furthermore, also in this case the residue Asp 79 is crucial since it forms recurring hydrogen bond interactions with the hydroxyl groups of the molecule (see Table 9). Finally, considering the hydrophobic contacts we have a similar situation to the one observed for compound

2. In fact, the purine units of compound **19** mainly establish hydrophobic interactions with Ala 41, Met 84, Leu 93, Met 116 and Phe 124 (see Table 10).

## Discussion

The main goal of this work consists in the proposal of a novel simple, rational *in silico* approach for the search of new possible small-molecule inhibitors taking into account the fundamental concepts of protein flexibility and conformational dynamics. In particular, the new strategy is based on the consideration that a ligand selectively binds one (or a subset) of the energetically similar protein conformers, stabilizing it (them) and inducing some specific protein conformations that cannot be predicted through a simple analysis of the corresponding unbound crystal structure. From the applicative point of view, considering the active protein-ligand complex, the proposed approach consists in a proper cluster analysis of a long-time scale all-atom Molecular Dynamics simulation to extract a minimal number of key configurations that take into account most of the structural fluctuations and, afterwards, in identifying the most recurring fundamental interactions between the protein and the ligand, interactions that allow to design a dynamic pharmacophore for the screening of small-molecule databases.

In this manuscript the examined protein is the molecular chaperone Hsp90 that is generally viewed as a “druggable” target for rational cancer therapy due to (a) its role at the crossroads of multiple signalling pathways associated with cell proliferation and cell viability and (b) to its elevated ATPase activity in tumors as opposed to normal tissues<sup>37</sup>. Exploiting our novel drug-discovery rational strategy, we have been able to identify new alternative N-terminal inhibitors explicitly including the flexibility and the conformational dynamics of the activated complex between Hsp90 and its natural ligand ATP. Importantly, the newly discovered Hsp90 inhibitors show interesting activities in blocking Hsp90 dependent cell pathways and interactions, as proved by the selective degradation of Hsp90 clients and the consequent cytotoxicity.

In this case, after selecting the seven most relevant conformational substates visited during the MD simulation of the complex, we have analyzed the occurrence of the intermolecular interactions between the ligand and the protein. In particular, we have observed that the intermolecular interactions that stabilize most the Hsp90-ATP complex are hydrogen bond-like interactions and, therefore, after identifying the hydrogen bond donors and acceptors in the ATP molecule, we have assigned them a weight proportional to the frequency with which they are involved in significant interactions in the extracted key configurations. This allowed us to design a dynamic pharmacophore model that implicitly takes into account the protein flexibility and that has been used to screen the small-molecule database NCI\_3D.

A possible important limitation of our pharmacophore model is that it is mainly based on mimicking hydrogen bond interactions. This may reverberate in the selection of molecules bearing charged groups, which is an undesirable feature in drug-like molecules. To alleviate this problem, we have included an additional level of control based on the application of the Lipinski rules<sup>13</sup>. The virtual screening with this further filter returned 25 compounds, eight of which have been obtained from the National Cancer Institute and tested in a panel of human breast carcinoma cell lines. The LogP values of all the 25 hits fulfill the Lipinski rule for the water-octanol partition coefficient<sup>13</sup> and, above all, eight compounds respect the stricter Ghose rule<sup>33</sup> as well. Three of the molecules passing the more stringent test are active against Hsp90, underlying the importance of including drug-like features in the second-stage analysis of the screening.

Furthermore, despite its limitations, the proposed pharmacophore model interestingly captures some of the binding determinants of the well-known inhibitors geldanamycin and PU24FCI. Indeed, both of these compounds are able to establish an extended network of direct hydrogen bonds with Hsp90. In contrast, our pharmacophore performs poorly with respect to other known inhibitors. However, it has to be noted that such compounds were designed and synthesized basing on the careful analysis of X-ray structures, namely, using a totally different strategy compared to the one proposed here.

In fact, one of the aims of our effort was to consider explicitly the possible dynamic states of the receptor during the stage of the pharmacophore design. Including receptor flexibility may allow to take into account the formation/disappearance/change of pockets due to the conformational response of the protein to the ligand (ATP in our case). In turn, this enables to identify the stereochemical properties of functional groups that should be present in a putative inhibitor to facilitate the optimal fitting to variations in the shape of the binding site. In this framework, the goal of our pharmacophore design was to select molecular entities that are different from known active compounds.

Four of the tested molecules (**2**, **4**, **19** and **20**) have shown an interesting cytotoxic effect and they have been subject to further experimental tests. In particular, exploiting the fact that the chaperone Hsp90 oversees the correct conformational development of a diverse ensemble of client proteins, we have investigated whether the cytotoxic effect exerted by the compounds relies on the interference with the Hsp90 chaperone activity. In particular, we have observed that three compounds (**2**, **4** and **19**) induce a decrease in the expression levels of Hsp90 client proteins for the cell lines MCF-7 and ZR75.1 and, furthermore, other tests have shown that the inhibition of some Hsp90 client proteins also results in a significant change in the distribution of the cells in the different phases of the cell cycle. In particular, the percentage of cells in apoptosis increases as the concentration of administrated compounds rises.

Finally, in order to gain an insight into possible structural determinants of the inhibitory/cellular activity of the most promising hits, we have studied their modes of interaction with Hsp90. To accomplish this task we have docked the active compounds into the ATP binding site of all the Hsp90 conformers used for the pharmacophore construction. Afterwards, from the statistical analysis of the intermolecular interactions in all the obtained Hsp90-inhibitor complexes, we have been able to identify the most recurring and stabilizing contacts and we have shown that they generally correspond to the ones already observed for a wide range of different known inhibitors. The fact that our active compounds reproduces the molecular recognition determinants of highly

optimized drug candidates (e.g., hydrogen bonding similarities to geldanamycin and hydrophobic interactions similarities to the purine-based inhibitors and related derivatives) suggests that their function may be determined by similar mechanisms. In particular, we propose that the identified hits engage the ATP binding site, disrupting the Hsp90 ATPase activity that is necessary for the chaperone cycle.

At this point it is very important to stress again that the main goal of our investigations was not to dock Hsp90 inhibitors into the active site and to analyze the structures of the resulting complexes. The aim of our study was to identify the possible molecular determinants, and their dynamic evolution, required for binding to a certain pocket using the natural ligand as a template and to include these findings in the pharmacophore design stage. Afterwards, this explicit information on the protein conformational motions was used to search for novel structural entities that may interfere with the biological functions of the target.

We believe that the proposed strategy can be extended also to other targets and exploited to identify active compounds using information obtained directly from the atomic-resolution analysis of the dynamics of complex biomolecules. However, the “drug-like properties” of the obtained hits must be carefully examined before starting experimental tests through the application of proper filters, such as the Lipinski<sup>13</sup> or Ghose<sup>33</sup> rules. In fact, a pharmacophore such as the one proposed here may actually be biased towards selecting molecules with a number of charged groups which would make them fail as drug candidates. In this respect, improvements in the pharmacophore design process are clearly desirable.

In the application reported in this paper the activities of molecules are certainly still distant from the ones required for efficient pharmacological applications. Therefore, from this point onwards, we believe that classical drug design and medicinal chemistry **efforts can be started** to improve the activities and pharmacokinetic properties of the compounds.

In conclusion, considering the obtained results, it is worth pointing out again that the inclusion of protein flexibility is an important factor to be considered in the framework of **rational drug**

**discovery.** This task can be accomplished simply considering a minimal number of key configurations that capture most of the structural fluctuations encountered in multi-nanosecond trajectories. The rate-limiting step in the whole procedure might be represented by the time required to run MD simulations. However, the fast pace in methodological developments and the continuous increase in computer power provide unprecedented chances to include atomic level descriptions of molecular motions in the drug discovery process<sup>38-40</sup>. Therefore, we believe that our new technique can open the possibility to rationally expand the chemical space of the inhibitors for any other protein target involved in cancer development and it can help in discovering new possible uses for already known drugs.

### **Acknowledgments**

A.G. and G.C. thank A.I.R.C. (Project IG. 11775), the PRIN project 2008K37RHP\_002 and the project “Ministero della Salute Progetto Giovani Ricercatori” for financial support. We acknowledge the Drug Synthesis and Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute (NCI/DTP, [dtp.nci.nih.gov](http://dtp.nci.nih.gov)) for providing the compounds used in our assays.

## References

1. H. A. Carlson, K. M. Masukawa, J. A. McCammon, *J. Phys. Chem. A* 1999, 103, 10213.
2. H. A. Carlson, K. M. Masukawa, K. Rubins, F. D. Bushman, W. L. Jorgensen, R. D. Lins, J. M. Briggs, J. A. McCammon, *J. Med. Chem.* 2000, 43, 2100.
3. H. A. Carlson, *Curr. Opin. Chem. Biol.* 2002, 6, 447.
4. K. L. Meagher, H. A. Carlson, *J. Am. Chem. Soc.* 2004, 126, 13276.
5. K. L. Damn, H. A. Carlson, *J. Am. Chem. Soc.* 2007, 129, 8225.
6. A. L. Bowman, Z. Nikolovska-Coleska, H. Zhong, S. Wang, H. A. Carlson, *J. Am. Chem. Soc.* 2007, 129, 12809.
7. M. G. Lerner, A. L. Bowman, H. A. Carlson, *J. Chem. Info. Model* 2007, 47, 2358.
8. P. Workman, *Curr. Cancer Drug Targets* 2003, 3, 297.
9. M. M. U. Ali, S. M. Roe, C. K. Vaughan, P. Meyer, B. Panaretou, P. W. Piper, C. Prodromou, L. H. Pearl, *Nature* 2006, 440, 1013.
10. S. Frey, A. Leskovar, J. Reinstein, K. Buchner, *J. Biol. Chem.* 2007, 282, 35612.
11. M. Mickler, M. Hessling, C. Ratzke, J. Buchner, T. Hugel, *Nat. Struct. Mol. Biol.* 2009, 16, 281.
12. J. Trepel, M. Mollapour, G. Giaccone, L. Neckers, *Nat. Rev. Cancer* 2010, 10, 537.
13. C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Deliv. Rev.* 2001, 46, 3.
14. G. Morra, M. A. C. Neves, C. J. Plescia, S. Tsustsumi, L. Neckers, G. Verkhivker, D. C. Altieri, G. Colombo, *J. Chem. Theory Comput.* 2010, 6, 2978.
15. D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* 2005, 26, 1701.
16. W. R. P. Scott, P. H. Hunenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kruger, W. F. van Gunsteren, *J. Phys. Chem. A* 1999, 103, 3596.
17. H. J. C. Berendsen, J. R. Grigera, P. R. Straatsma, *J. Phys. Chem.* 1987, 91, 6269.

18. T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* 1993, 98, 10089.
19. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* 1995, 103, 8577.
20. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* 1997, 18, 1463.
21. S. Miyamoto, P. A. Kollman, *J. Comput. Chem.* 1992, 13, 952.
22. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, J. R. Haak, *J. Chem. Phys.* 1984, 81, 3684.
23. F. Pontiggia, A. Zen, C. Micheletti, *Biophys. J.* 2008, 95, 5901.
24. L. Kaufman, P. J. Rousseuw, in *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York, 2005.
25. Discovery Studio 2.5, Accelrys, USA, 2009.
26. Instant JChem, 2.0.0 , 2006, ChemAxon (<http://chemaxon.com>).
27. P. Skehan, R. Storeng, D. Scudiero, A. Monks, J. McMahon, D. Vistica, J. T. Warren, H. Bokesch, S. Kenney, M. R. Boyd, *J. Natl. Cancer Inst.* 1990, 82, 1107.
28. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* 2009, 30, 2785.
29. C. E. Stebbins, A. A. Russo, C. Schneider, N. Rosen, F. U. Hartl, N. P. Pavletich, *Cell* 1997, 89,239.
30. S. M. Roe, C. Prodromou, R. O'Brien, J. E. Ladbury, P. W. Piper, L. H. Pearl, *J. Med. Chem.* 1999, 42, 260.
31. R. M. Immormino, Y. Kang, G. Chiosis, D. T. Gewirth, *J. Med. Chem.* 2006, 49, 4953.
32. T. E. Barta, J. M. Veal, J. W. Rice, J. M. Partridge, R. P. Fadden, W. Ma, M. Jenks, L. Geng, G. J. Hanson, K. H. Huang, A. F. Barabasz, B. E. Foley, J. Otto, S. E. Hall, *Bioorg. Med. Chem. Lett.* 2008, 18, 3517.
33. A. K. Ghose, V. N. Viswanadhan, J. J. Wendoloski, *J. Comb. Chem.* 1999, 1, 55.

34. M. Taipale, D. F. Jarosz, S. Lindquist, *Nat. Rev. Mol. Cell. Biol.* 2010, 11, 515.
35. P. Fortugno, E. Beltrami, J. Plescia, J. Fontana, D. Pradhan, P. C. Marchisio, W. C. Sessa, D. C. Altieri, *Proc. Natl. Acad. Sci. U.S.A.* 2003, 100, 13791.
36. M. Meli, M. Pennati, M. Curto, M. G. Daidone, J. Plescia, S. Toba, D. C. Altieri, N. Zaffaroni, G. Colombo, *J. Med. Chem.* 2006, 49, 7721.
37. A. Kamal, L. Thao, J. Sensintaffar, L. Zhang, M. F. Boehm, L. C. Fritz, F. J. Burrows, *Nature* 2003, 425, 407.
38. D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. B. Shan, W. Wiggers, *Science* 2010, 330, 341.
39. Y. B. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, D. E. Shaw, *J. Am. Chem. Soc.* 2011, 133, 9181.
40. R. O. Dror, A. C. Pan, D. H. Arlow, D.W. Borhami, P. Maragakis, Y. B. Shan, H. Xu, D. E. Shaw, *Proc. Natl. Acad. Sci. USA* 2011, 108, 13118.

## Figure captions

Figure 1. ATP molecule with labels for atoms considered in the hydrogen-bond analyses for the construction of the pharmacophore model.

Figure 2. The ten-feature pharmacophore model ATP-PHARM-1 plotted on the ATP molecule. Green spheres and arrows represent hydrogen bond acceptors, while purple spheres and arrows represent hydrogen bond donors. This is the pharmacophore used for screening the small molecule NCI\_3D database.

Figure 3. Molecular structures of the compounds **1-8** selected from the virtual screening of the NCI\_3D database using the pharmacophore model ATP-PHARM-1.

Figure 4. Molecular structures of the compounds **9-17** selected from the virtual screening of the NCI\_3D database using the pharmacophore model ATP-PHARM-1.

Figure 5. Molecular structures of the compounds **18-25** selected from the virtual screening of the NCI\_3D database using the pharmacophore model ATP-PHARM-1.

Figure 6. Effects of small molecules on Hsp90 chaperone function. Western blotting experiments showing the loss of Hsp90 client proteins in breast cancer cells treated with 1% DMSO (v/v) (control cells; ctr), or with compounds **2**, **4**, **19** or **20** (IC<sub>50</sub> and IC<sub>80</sub>) for 72 h.

Figure 7. Effects of small molecules **and 17-AAG** on cell cycle phase distribution and induction of apoptosis. Cells were exposed for 72 h to 1% DMSO (v/v) (control cells; ctr) or to compounds **2**, **4**, **19** or **17-AAG** (IC<sub>20</sub>, IC<sub>50</sub> and IC<sub>80</sub>). **(A)** Cell-cycle phase distribution was assessed by flow cytometric analyses. The percentage of cells in G0/G1 (black), S (white) and G2/M (grey) phases are shown. Columns represent the mean of three independent experiments; SDs always within  $\pm 5\%$ . **(B)** Percentage of cells with an apoptotic morphology as assessed by fluorescence microscopy. Data represent the mean values  $\pm$ SD of three independent experiments. **\*\*P** <0.001 and **\*P** <0.01, student's t test.

Figure 8. Compound **2** with labels to identify its “quasi-equivalent” units.

Figure 9. Compound **19** with labels to identify its “quasi-equivalent” units.

Table 1- List of the possible ATP hydrogen bond acceptors, their occurrence in the analyzed Hsp90-ATP complexes, their relative weight and presence in the constructed pharmacophore model ATP\_PHARM-1.

HB Acceptor	Occurrence	Relative Weight	Presence in the pharmacophore model
OB	28	2.00	Yes
O5	19	1.35	Yes
O8	19	1.35	Yes
OA	19	1.35	Yes
O9	11	0.79	Yes
OC	10	0.71	Yes
O3	8	0.57	Yes
O6	8	0.57	Yes
O0	5	0.36	No
O2	3	0.21	No
O7	3	0.21	No
O1	2	0.14	No
O4	0	0.00	No
N2	0	0.00	No
N3	0	0.00	No
N5	0	0.00	No

Table 2- List of the possible ATP hydrogen bond donors, their occurrence in the analyzed Hsp90-ATP complexes, their relative weight and presence in the constructed pharmacophore model ATP-PHARM-1.

HB Acceptor	Occurrence	Relative Weight	Presence in the pharmacophore model
OC-H5	11	0.79	Yes
N4-H1	6	0.43	Yes
O3-H4	3	0.21	No
O2-H3	2	0.14	No

Table 3 – LogP values for the compounds identified using the pharmacophore model ATP-PHARM-1. The fulfillment of the Lipinski and Ghose rules for the octanol-water partition coefficient and the experimental activity of the different hits are also shown. The acronyms TBD stands for “To Be Determined” and indicates that the compound has not been obtained from the National Cancer Institute to be tested.

Compound	LogP	Lipinski	Ghose	Activity
1	-2.58	Yes	No	TBD
2	-0.20	Yes	Yes	Yes
3	-1.64	Yes	No	TBD
4	-1.57	Yes	No	Yes
5	-3.00	Yes	No	No
6	-2.94	Yes	No	TBD
7	-0.44	Yes	No	TBD
8	-0.88	Yes	No	TBD
9	+0.76	Yes	Yes	TBD
10	+0.06	Yes	Yes	No
11	-0.82	Yes	No	No
12	-6.36	Yes	No	TBD
13	-1.63	Yes	No	TBD
14	-4.22	Yes	No	TBD
15	-3.55	Yes	No	TBD
16	-3.30	Yes	No	TBD
17	-0.67	Yes	No	TBD
18	-3.75	Yes	No	TBD
19	-0.10	Yes	Yes	Yes
20	+2.16	Yes	Yes	Yes
21	+0.16	Yes	Yes	TBD
22	-1.80	Yes	No	TBD
23	-6.89	Yes	No	No
24	+0.98	Yes	Yes	TBD
25	+0.90	Yes	Yes	TBD

Table 4- Cytotoxic activity of small molecules and 17-AAG in breast tumor cell lines. Data are reported as IC<sub>50</sub> values (concentration of drug required to inhibit growth by 50%) determined by the SRB assay after 72 h of continuous exposure to each compound. Data represent mean values ± SD of at least three independent experiments.

Cell Line	IC <sub>50</sub> values (μM) for the different compounds								
	2	4	5	10	11	19	20	23	17-AAG
BT-474	27.3±4.8	27.1±3.6	>50	>50	>50	24.1±4.8	21.5±3.4	>50	0.23±0.02
HCC1937	37.3±4.4	27.8±4.2	>50	>50	>50	22.2±3.4	37.2±5.6	>50	0.08±0.01
MCF-7	21.3±2.8	4.2±1.1	>50	>50	>50	15.1±2.1	43.1±2.8	>50	3.48±0.22
MDA-MB-231	25.1±3.2	6.3±1.2	>50	>50	>50	19.3±2.8	16.8±2.9	>50	1.12±0.04
T-47D	25.3±1.7	2.6±0.8	>50	>50	>50	16.4±2.8	28.2±3.9	>50	1.57±0.11
ZR75.1	31.4±3.2	28.6±4.4	>50	>50	>50	18.2±2.2	29.6±3.3	>50	0.14±0.02

Table 5- Hydrogen-bond interactions in which compound **2** acts as acceptor. The frequency of the interactions is shown and the significant interactions (frequency greater than or equal to five) are highlighted in italic. For the sake of clarity interactions with frequency equal to one are omitted. The acronyms are described in the text.

Hydrogen Bond Donor	Hydrogen Bond Acceptor	Frequency
<i>Asn 37 (side-chain)</i>	<i>SU</i>	8
<i>Asn 92 (side-chain)</i>	<i>SU</i>	7
<i>Gly 123 (backbone)</i>	<i>SU</i>	6
<i>Phe 124 (backbone)</i>	<i>SU</i>	6
<i>Tyr 125 (side-chain)</i>	<i>SU</i>	6
<i>Gly 121 (backbone)</i>	<i>SU</i>	5
<i>Val 122 (backbone)</i>	<i>SU</i>	5
Phe 120 (backbone)	SU	4
Arg 345 (side-chain)	SU	4
Lys 102 (backbone)	SU	3
Gly 100 (backbone)	SU	2
Gln 119 (backbone)	SU	2

Table 6- Hydrogen-bond interactions in which compound **2** acts as donor. The frequency of the interactions is shown and the significant interactions (frequency greater than or equal to five) are highlighted in italic. For the sake of clarity interactions with frequency equal to one are omitted. The acronyms are described in the text.

Hydrogen Bond Acceptor	Hydrogen Bond Donor	Frequency
<i>Asp 79 (side-chain)</i>	<i>ARAM</i>	<i>6</i>
<i>Asn 92 (backbone)</i>	<i>ARAM</i>	<i>5</i>
Thr 171 (side-chain)	ARAM	4
Asn 37 (side-chain)	ALAM	3
Leu 34 (backbone)	ARAM	2
Asp 40 (side-chain)	ALAM	2
Met 84 (side-chain)	ALAM	2
Gly 94 (backbone)	ARAM	2
Ser 99 (backbone)	ARAM	2
Ser 99 (side-chain)	ARAM	2
Met 116 (backbone)	ALAM	2
Ile 117 (backbone)	ARAM	2
Gln 119 (backbone)	ARAM	2

Table 7- Hydrophobic interactions in which compound **2** is involved. The frequency of the interactions is shown and the significant interactions (frequency greater than or equal to five) are highlighted in *italic*. For the sake of clarity interactions with frequency equal to one are omitted. The acronyms are described in the text.

Hydrophobic Residue	Inhibitor Interacting Unit	Frequency
<i>Met 84</i>	<i>AR</i>	<i>12</i>
<i>Leu 173</i>	<i>AR</i>	<i>11</i>
<i>Ile 82</i>	<i>AR</i>	<i>9</i>
<i>Phe 124</i>	<i>AR</i>	<i>9</i>
<i>Ala 41</i>	<i>AR</i>	<i>8</i>
<i>Leu 93</i>	<i>AR</i>	<i>7</i>
<i>Met 116</i>	<i>AR</i>	<i>7</i>
<i>Val 136</i>	<i>AR</i>	<i>6</i>
Ala 38	AR	4
Leu 89	AR	4
Ile 117	AR	3
Leu 34	AR	2
Ile 77	AR	2
Ala 97	AR	2
Ala 103	AR	2
Met 116	AR	2
Val 122	AR	2
Leu 127	AR	2

Table 8- Hydrogen-bond interactions in which compound **19** acts as acceptor. The frequency of the interactions is shown and the significant interactions (frequency greater than or equal to five) are highlighted in *italic*. For the sake of clarity interactions with frequency equal to one are omitted. The acronyms are described in the text.

Hydrogen Bond Donor	Hydrogen Bond Acceptor	Frequency
<i>Gly 83 (backbone)</i>	<i>OG</i>	<i>6</i>
<i>Asn 37 (side-chain)</i>	<i>OG</i>	<i>5</i>
<i>Gln 119 (side-chain)</i>	<i>OE</i>	<i>5</i>
<i>Phe 120 (backbone)</i>	<i>OG</i>	<i>5</i>
<i>Gly 123 (backbone)</i>	<i>OG</i>	<i>5</i>
Gly 100 (backbone)	OG	4
Gly 121 (backbone)	OG	4
Val 122 (backbone)	OG	4
Thr 171 (side-chain)	OG	4
Arg 345 (side-chain)	OG	4
Arg 356 (side-chains)	OE	4
Ile 82 (backbone)	OG	3
Asn 37 (side-chain)	OE	3
Gln 119 (backbone)	OG	3
Gly 100 (backbone)	OE	2
Lys 102 (backbone)	OG	2
Lys 102 (backbone)	OE	2
Ala 103 (backbone)	OG	2
Phe 124 (backbone)	OG	2

Table 9 - Hydrogen-bond interactions in which compound **19** acts as donor. The frequency of the interactions is shown and the significant interactions (frequency greater than or equal to five) are highlighted in *italic*. For the sake of clarity interactions with frequency equal to one are omitted. The acronyms are described in the text.

Hydrogen Bond Acceptor	Hydrogen Bond Donor	Frequency
<i>Asp 79 (side-chain)</i>	<i>OG</i>	<i>10</i>
Asn 37 (side-chain)	OG	4
Lys 98 (backbone)	OG	4
Ser 99 (backbone)	OG	3
Ser 99 (backbone)	OG	3
Glu 33 (side-chain)	OG	2
Glu 88 (side-chain)	OG	2
Met 116 (backbone)	OG	2
Tyr 125 (side-chain)	OG	2
Thr 171 (side-chain)	OG	2

Table 10- Hydrophobic interactions in which compound **19** is involved. The frequency of the interactions is shown and the significant interactions (frequency greater than or equal to five) are highlighted in *italic*. For the sake of clarity interactions with frequency equal to one are omitted. The acronyms are described in the text.

Hydrophobic Residue	Inhibitor Interacting Unit	Frequency
<i>Met 84</i>	<i>PU</i>	<i>9</i>
<i>Leu 93</i>	<i>PU</i>	<i>6</i>
<i>Phe 124</i>	<i>PU</i>	<i>6</i>
<i>Ala 41</i>	<i>SLU</i>	<i>5</i>
<i>Met 116</i>	<i>PU</i>	<i>5</i>
Ile 82	SLU	3
Ala 38	SLU	2
Ala 41	PU	2
Ile 82	PU	2
Met 116	SLU	2
Leu 127	PU	2
Leu 173	PU	2

Figure 1-

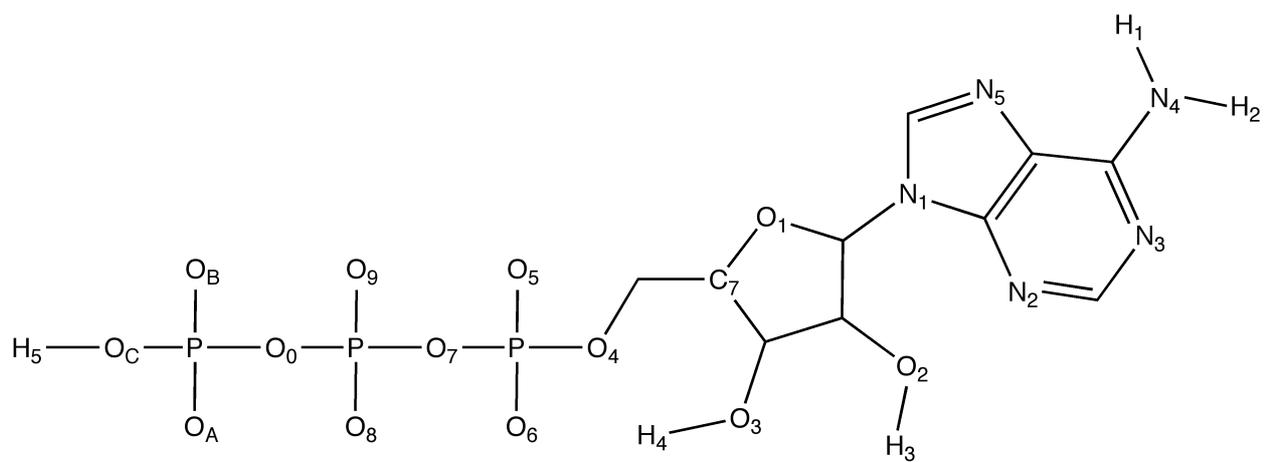


Figure 2 –

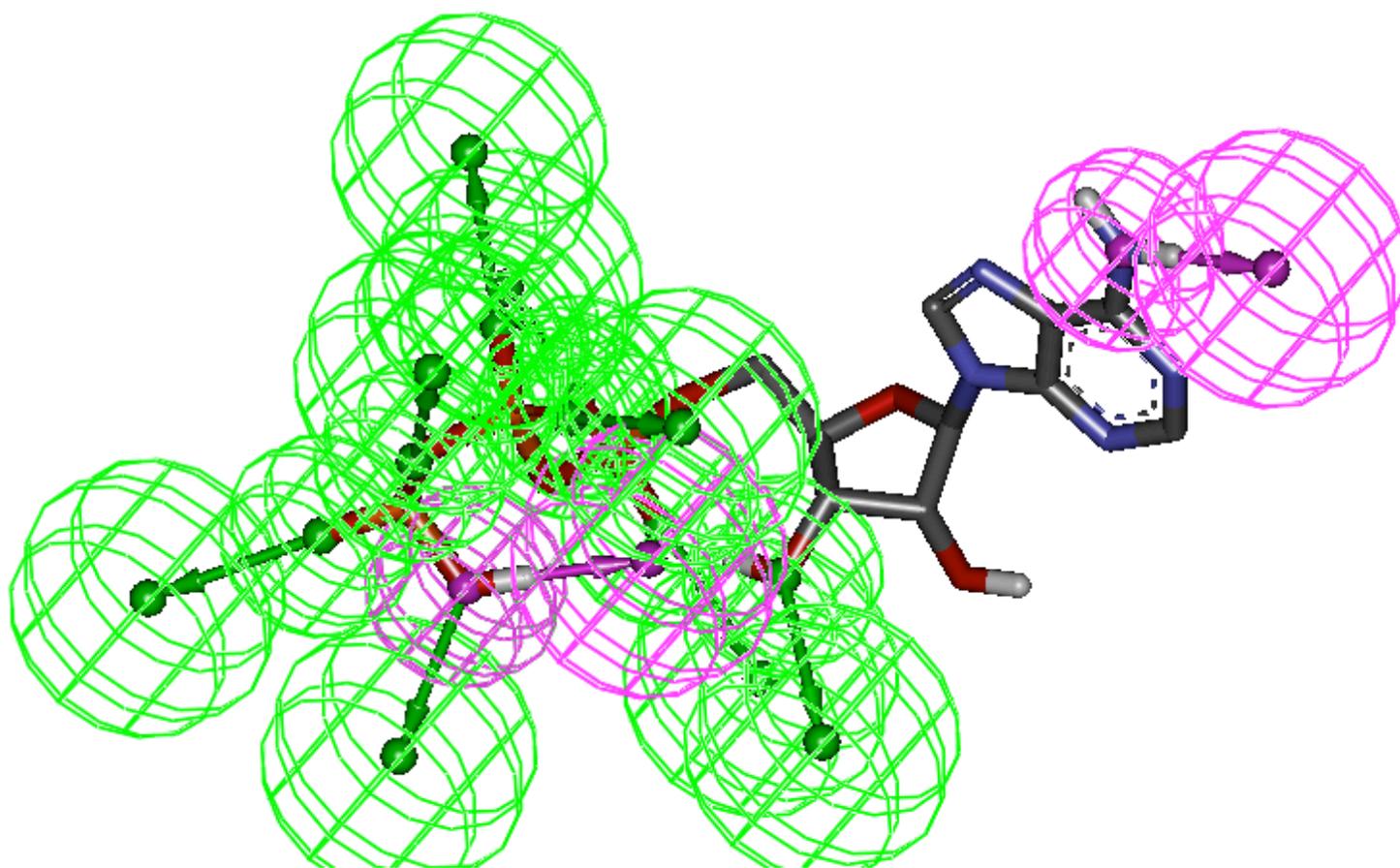
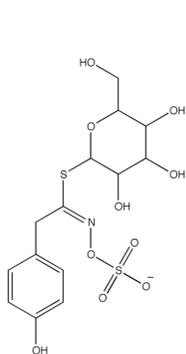
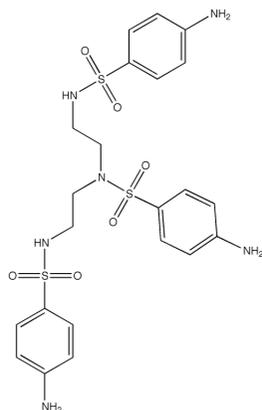


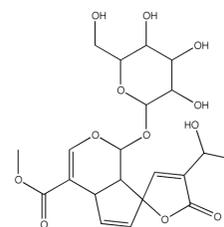
Figure 3 –



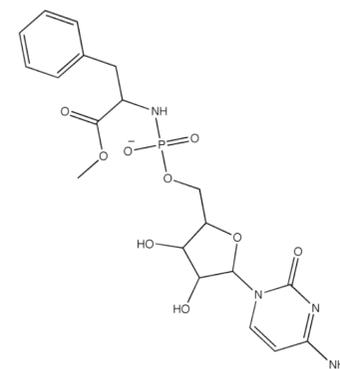
Compound 1  
NSC Number: 70863



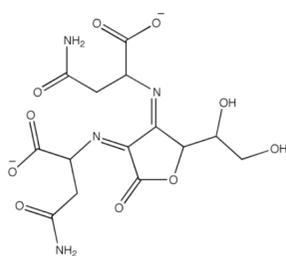
Compound 2  
NSC Number: 120244



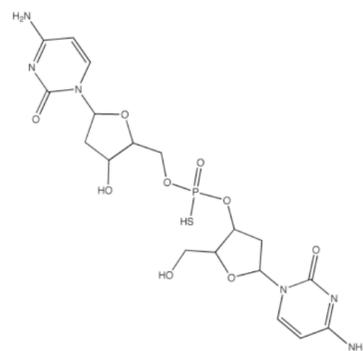
Compound 3  
NSC Number: 609065



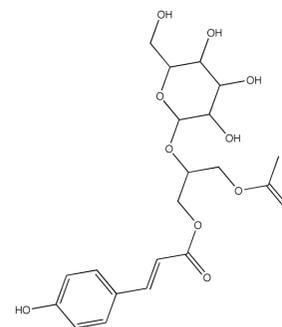
Compound 4  
NSC Number: 674186



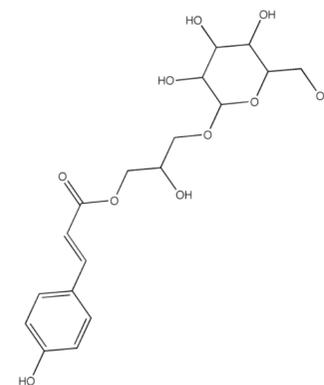
Compound 5  
NSC Number: 668472



Compound 6  
NSC Number: 695649

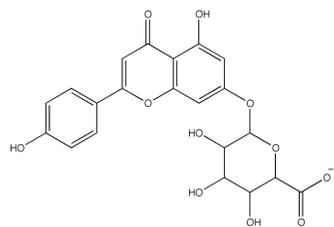


Compound 7  
NSC Number: 627041

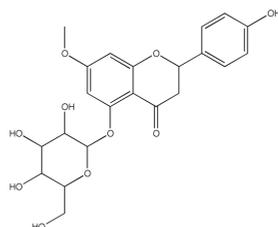


Compound 8  
NSC Number: 626437

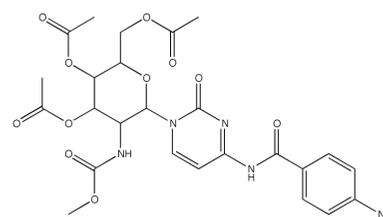
Figure 4 –



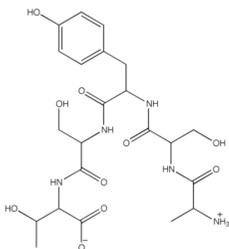
**Compound 9**  
NSC Number: 622811



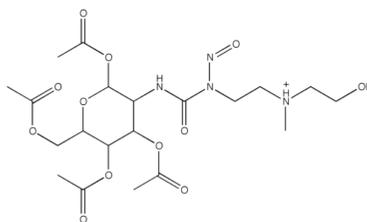
**Compound 10**  
NSC Number: 407308



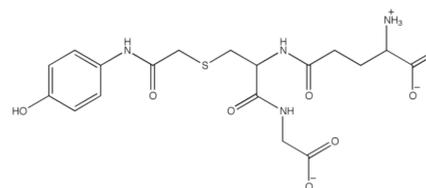
**Compound 11**  
NSC Number: 407130



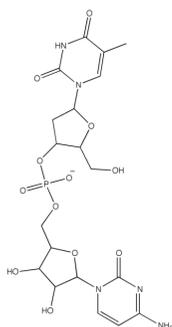
**Compound 12**  
NSC Number: 634822



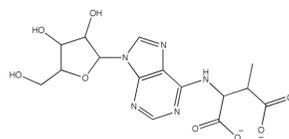
**Compound 13**  
NSC Number: 332568



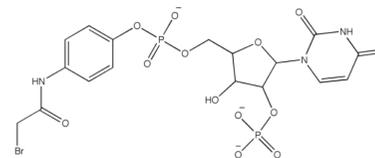
**Compound 14**  
NSC Number: 381988



**Compound 15**  
NSC Number: 365429

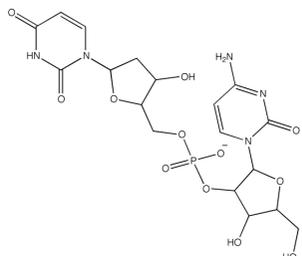


**Compound 16**  
NSC Number: 294196

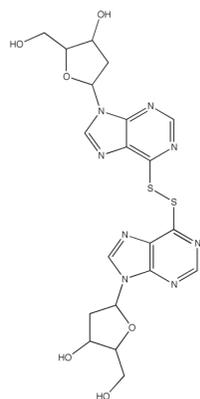


**Compound 17**  
NSC Number: 246114

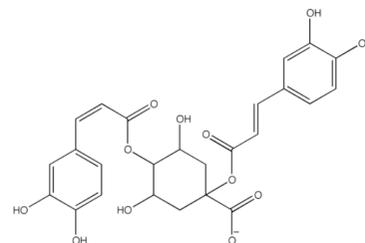
Figure 5 –



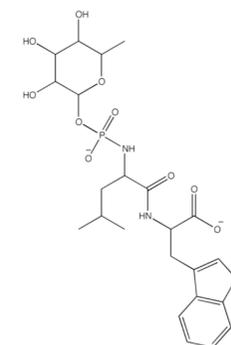
Compound 18  
NSC Number: 128278



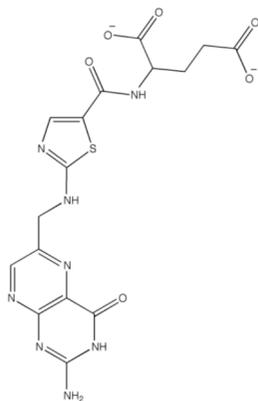
Compound 19  
NSC Number 118612



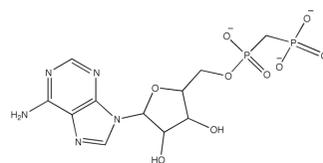
Compound 20  
NSC Number: 91529



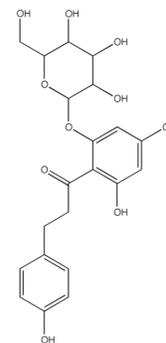
Compound 21  
NSC Number: 694280



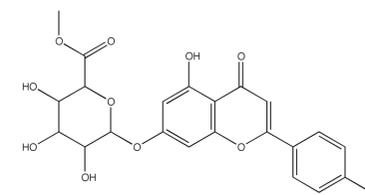
Compound 22  
NSC Number: 143665



Compound 23  
NSC Number: 614641



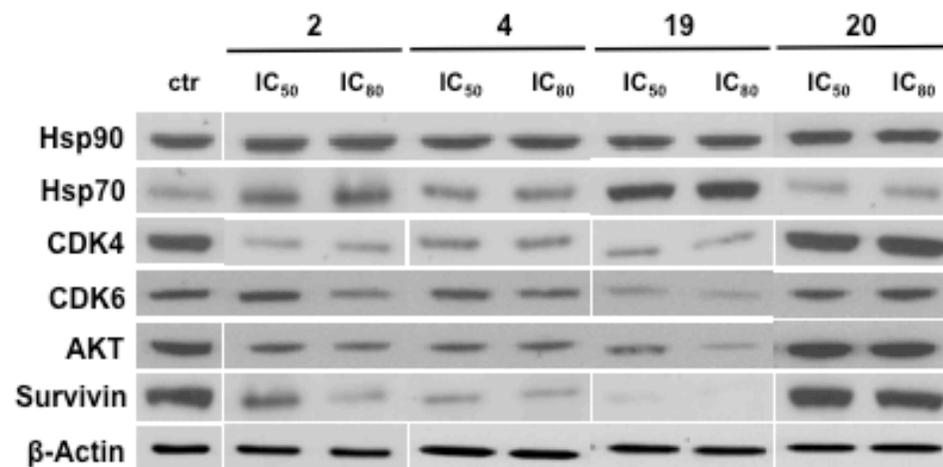
Compound 24  
NSC Number: 2833



Compound 25  
NSC Number: 622810

Figure 6 –

### MCF-7



### ZR75.1

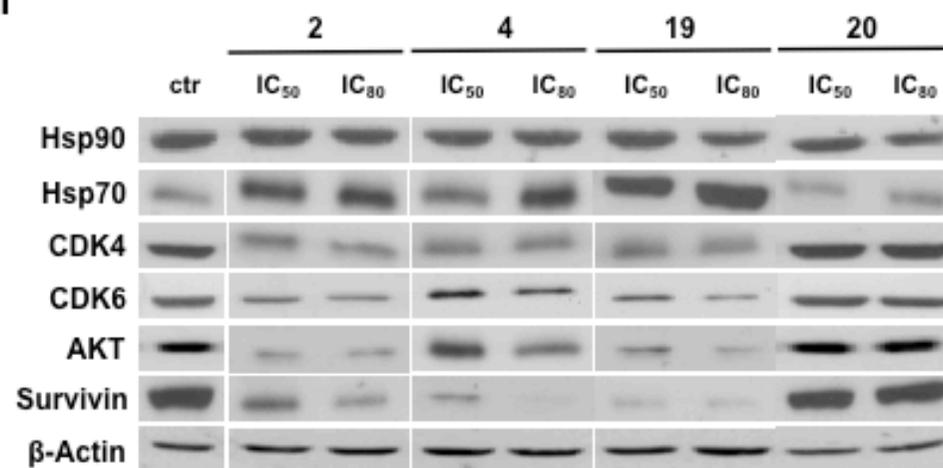


Figure 7 -

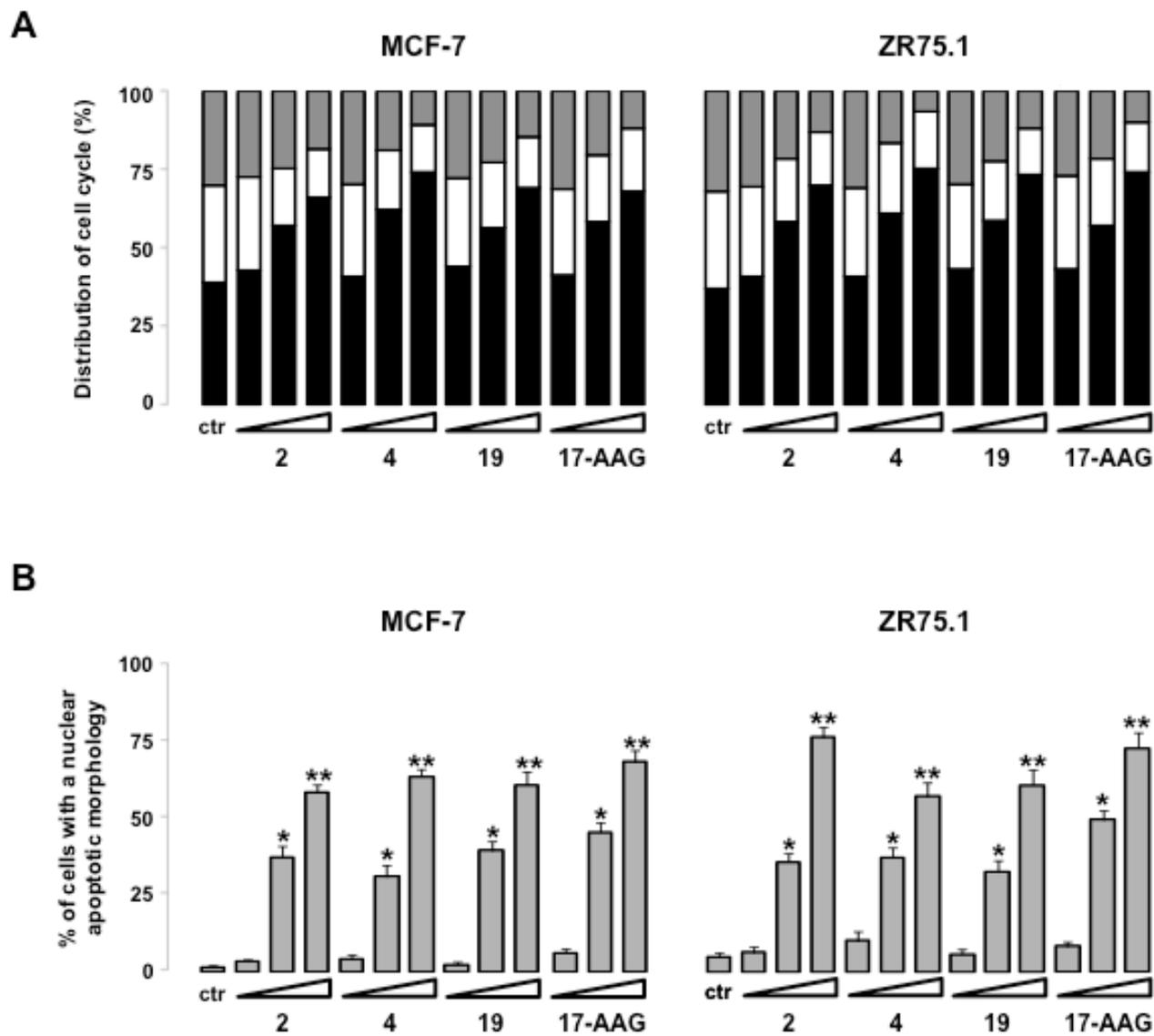




Figure 9 –

