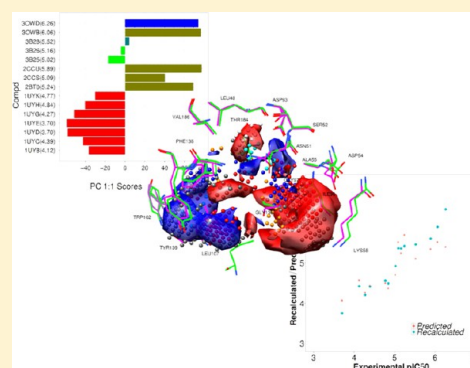


# Hsp90 Inhibitors, Part 1: Definition of 3-D QSAutogrid/R Models as a Tool for Virtual Screening

Flavio Ballante,<sup>†</sup> Antonia Caroli,<sup>‡</sup> Richard B. Wickersham, III,<sup>†,§</sup> and Rino Ragno<sup>\*,†</sup><sup>†</sup>Rome Center for Molecular Design, Dipartimento di Chimica e Tecnologie del Farmaco, Sapienza Università di Roma, P. le A. Moro 5, 00185 Roma, Italy<sup>‡</sup>Department of Physics, Sapienza Università di Roma, P.le Aldo Moro 5, 00185, Roma, Italy<sup>§</sup>Department of Biochemistry and Molecular Biophysics, Washington University in St. Louis School of Medicine, 700 South Euclid Avenue, St. Louis, Missouri 63110, United States

## Supporting Information

**ABSTRACT:** The multichaperone heat shock protein (Hsp) 90 complex mediates the maturation and stability of a variety of oncogenic signaling proteins. For this reason, Hsp90 has emerged as a promising target for anticancer drug development. Herein, we describe a complete computational procedure for building several 3-D QSAR models used as a ligand-based (LB) component of a comprehensive ligand-based (LB) and structure-based (SB) virtual screening (VS) protocol to identify novel molecular scaffolds of Hsp90 inhibitors. By the application of the 3-D QSAutogrid/R method, eight SB PLS 3-D QSAR models were generated, leading to a final multiprobe (MP) 3-D QSAR pharmacophoric model capable of recognizing the most significant chemical features for Hsp90 inhibition. Both the monoprobe and multiprobe models were optimized, cross-validated, and tested against an external test set. The obtained statistical results confirmed the models as robust and predictive to be used in a subsequent VS.



## INTRODUCTION

Molecular chaperones are protein machines, which play a key role in maturation and stabilization of protein substrates within the cell and are essential to restoring misfolded proteins created under stress conditions.<sup>1</sup> Among chaperones, the 90 kDa heat shock protein (Hsp90) is the most extensively studied.<sup>2,3</sup> Hsp90 consists of three functional domains: (1) a highly conserved N-terminal domain that contains an ATP-binding pocket, (2) a middle domain that regulates the interaction with substrate client proteins and also contains residues critical for ATPase activity, and (3) a highly conserved C-terminal region that regulates Hsp90 dimerization.<sup>4</sup> In eukaryotic cells, Hsp90 comprises 1%–2% of cellular proteins in basal conditions, but its expression increases significantly after exogenous injury.<sup>5</sup> Hsp90 is present in cells in two isoforms— $\alpha$  (inducible, major form) and  $\beta$  (constitutive, minor form) located predominantly in the cytosol—and two paralogues—ER-resident Grp94 and mitochondrial tumor necrosis factor receptor-associated protein 1 (Trap 1).<sup>5</sup> Under some cellular stress conditions, such as elevated temperature, abnormal pH, nutrient unavailability, and malignancy, the accumulation of misfolded proteins and the increased synthesis of new proteins commonly occurs. Under these conditions, Hsp90 is overexpressed to refold denatured proteins back into their native conformation. Thus, Hsp90 acts as a conformational buffer, and alteration of this function can

have devastating effects on cell viability.<sup>6</sup> The Hsp90 protein is the chaperone promoter of the folding of several clients, including many oncology targets (p53 mutants, Raf1, Akt, Bcr-Abl, Her2, EGFR) operating in tumor generation and progression.<sup>7</sup> Oncogenic cells depend on their ability to withstand endogenous (anoxia, nutrient deprivation, pH changes, and deranged signaling pathways) and exogenous (chemotherapy and radiation therapy) stressors for survival. Pharmacological inhibition of Hsp90 destabilizes proteins, leading to their degradation through the proteasome.<sup>8</sup> Therefore, one strategy in medicine is to block the ATPase activity of Hsp90 and enhance the degradation of aberrant proteins by the ubiquitin-proteasome pathway, which, in turn, leads to growth arrest or apoptosis in cancer cells.<sup>9</sup> Hsp90 has been confirmed as a particularly exciting cancer target, because of its involvement in the activity of numerous oncogenic proteins. In the past decade, the inhibition of Hsp90 has been one of the most used strategies in cancer therapy and the attention on it has increased impressively: starting from 2000 until November 2013, 302 X-ray structures (with and without inhibitors) were reported in Protein Data Bank (PDB),<sup>10</sup> and, among them, 196 were from *Homo sapiens* and 109 were published since 2011.

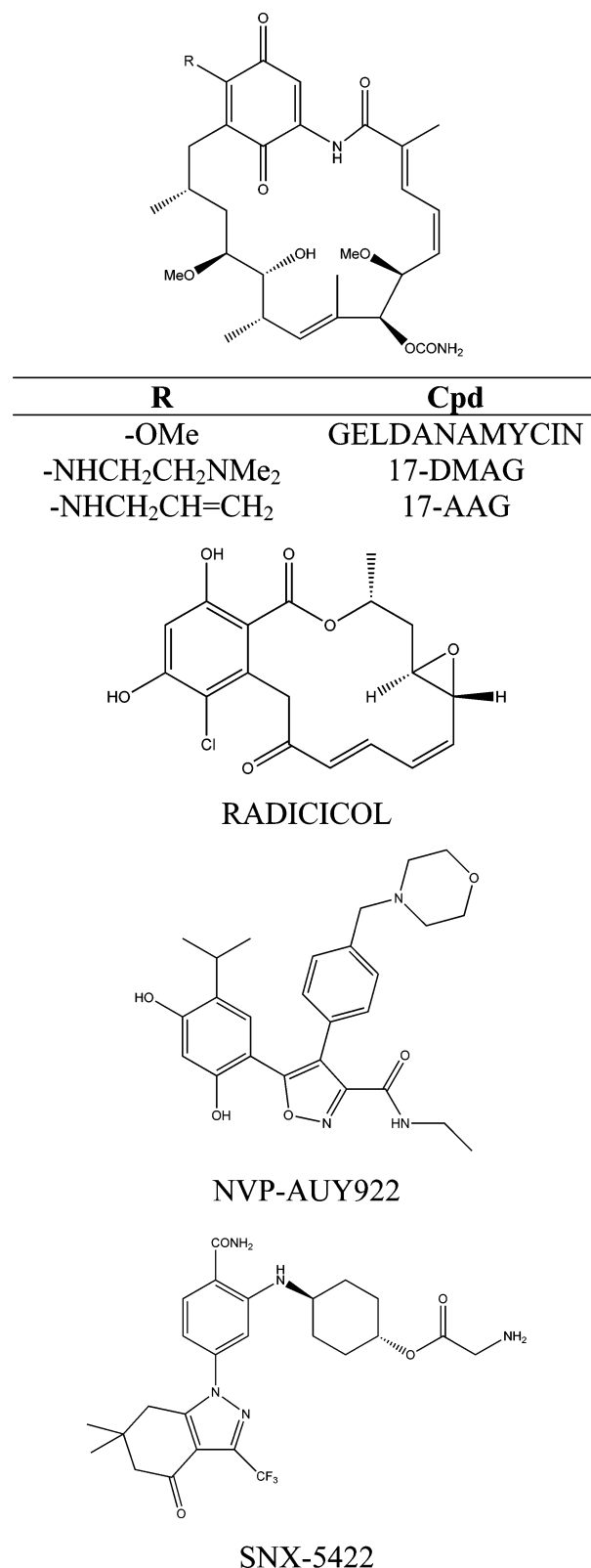
Received: December 22, 2013

Published: February 24, 2014

Moreover, a survey of the SciFinder database (in November 2013) using Hsp90 and inhibitor as key queries, resulted in 6642 results as concept, and 2095 results as entered. Several drugs that inhibit the ATPase activity of Hsp90 were identified in the past decade, including compounds that advanced to clinical trials against different type of cancers.<sup>11</sup> Natural products such as geldanamycin<sup>12</sup> and radicicol<sup>13</sup> were among the first Hsp90 inhibitors discovered (Figure 1). Because of their poor solubility, limited bioavailability and hepatotoxicity, some derivatives were obtained with the aim to improve pharmacological properties (17-AAG<sup>14</sup> and 17-DMAG;<sup>15</sup> see Figure 1). Considering that blocking the ATPase activity is an effective approach, de novo drug design was accomplished and small molecules able to bind the N-terminal Hsp90 binding site were identified. Purine-scaffold inhibitors,<sup>16</sup> dihydroxyphenylpyrazoles,<sup>17</sup> isoxazole derivatives (NVP-AUY922),<sup>18</sup> and carbazol-4-one benzamide derivatives (SNX-5422)<sup>19</sup> were promising candidates for cancer therapy and advanced to clinical studies (Figure 1). Despite this huge amount of data, only a limited number of three-dimensional (3-D) quantitative structure–activity relationship (QSAR) applications were reported (nine papers resulted by querying SciFinder using 3D QSAR and Hsp90 as keys) by different research groups.<sup>4,20–27</sup> Some of the reported 3-D QSARs used pharmacophore-based data interpretation and predictions,<sup>21,24–26</sup> and just five used true 3-D QSAR technology (i.e., CoMFA)<sup>4,20,23,25,27</sup> Regarding the 3-D data availability, only two papers reported the massive use of structural information but confined it in a pharmacophoric/docking protocol.<sup>22,24</sup> In the present work, the recently introduced 3-D QSAutogrid/R procedure<sup>28</sup> was applied to a dataset of 24 Hsp90/inhibitor co-crystals to define 3-D QSAR models. These were used to derive pharmacophoric quantitative models,<sup>29</sup> by the application of Multi Probe Guided Region Selection (MPGRS package),<sup>28</sup> to correlate structural Hsp90 inhibitors features with their biological data. All 3-D QSAR models were externally validated for robustness and predictiveness as a tool in a virtual screening (VS) protocol.<sup>30</sup>

## ■ COMPUTATIONAL METHODS

The 3-D QSAR still represents a useful tool to design new molecules, saving time and financial resources.<sup>31</sup> In any 3-D QSAR model, after careful selection of the training set, several steps must be taken.<sup>32–35</sup> Among them, alignment (ligand-based (LB) and/or structure-based (SB)), statistical analysis (PLS and cross-validation)<sup>36</sup> and external validation are of fundamental importance. Moreover, the application of the recently reported procedure, 3-D QSAutogrid/R,<sup>28</sup> yields quantitative pharmacophore models,<sup>29</sup> summarizing in a single representation all the diverse chemical information, which is a useful tool for the prediction of untested molecules' biological activities. The following section describes how these steps were accomplished in order to develop an adequate procedure to efficiently predict the activity of new Hsp90 inhibitors (see Figure 2). First, all the available ligand–protein complexes (for a total of 39 co-crystals) were acquired from the PDB.<sup>10</sup> From these, 24 were selected to compose the dataset (see the Training and Test Set Selection section). After a single-point minimization, the ligands were extracted from the complexes and split into a training set of 15 compounds and a test set of 9 compounds, to derive, by means of 3-D QSAutogrid/R,<sup>28</sup> nine 3-D QSAR PLS models: eight monoprobe models and a final multiprobe (MP) model. Moreover, all of the training-set



**Figure 1.** Example of Hsp90 inhibitors currently in clinical trials: the ansamycin compound geldanamycin and its derivatives radicicol, isoxazole (NVP-AUY922), and carbazol-4-one benzamide (SNX-5422) are depicted.

inhibitors were used in LB and SB approaches to validate the alignment procedure that was, then, applied on new candidates (with unknown binding mode), as explained in the later VS.<sup>30</sup>

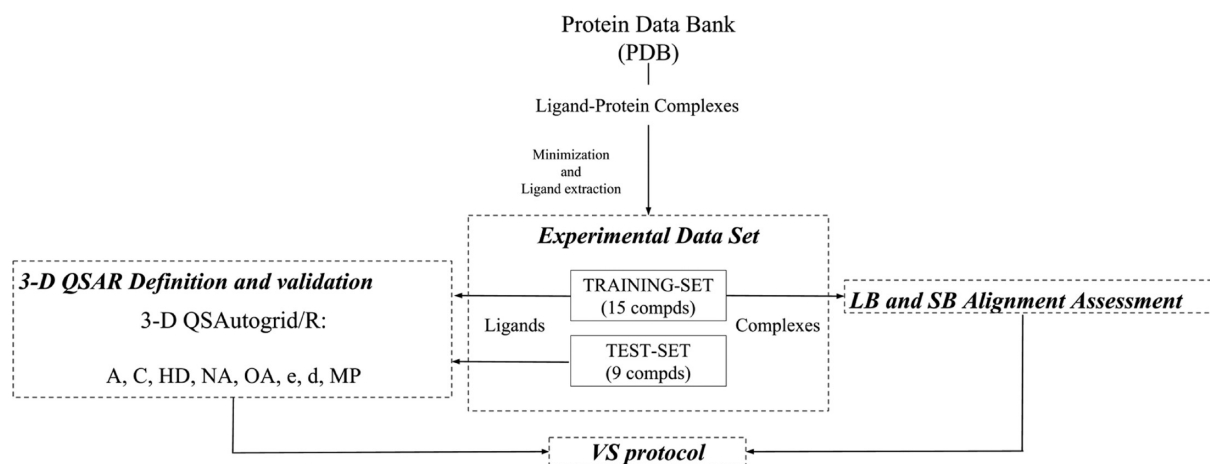


Figure 2. General procedure.

The 3-D QSAR models were used as an external scoring function in order to predict the activity of yet untested ligands in the succeeding virtual screening procedure.<sup>30</sup>

**Training and Test Set Selection.** Co-crystallized Hsp90/inhibitor complexes with known bioactivities showing the inhibitors bound within the N-terminal Hsp90 binding site were selected. At the beginning of the present study, the bioactivities for 39 co-crystallized inhibitors were available (see Table S1 in the Supporting Information).

Because of the different techniques used for the biological measurements, not all compounds could be used for deriving a sound mathematical model. In vitro activity of Hsp90 inhibitors is determined mostly through two conventional biological assays: the measure of the ATPase activity (referenced herein as ATPase assay) and the measure of competitive inhibition using a fluorescent probe (fluorescence polarization). In vivo tests were conducted on several different cancer cell types, giving nonhomogeneous results. Twenty-four (24) compounds in Table S1 in the Supporting Information had biological activities reported as IC<sub>50</sub> values obtained from the ATPase assay, and these were chosen as the dataset (see Table 1). The selected dataset compounds are structurally related to purine,<sup>16</sup> pyrazole,<sup>17,37</sup> 2-aminopyrimidine,<sup>7</sup> triazine,<sup>7</sup> and *N*-aryl-benzimidazolone<sup>38</sup> scaffolds. To find the best training set composition, the dataset was iteratively partitioned, maintaining a similar range of activities in training and test sets, and the corresponding 3-D QSAR models were validated. The final combination showing the best  $q^2$  and SDEP values was then selected: 15 compounds (shown in Table 2, presented later in this work) with affinity values spanning more than 2.5 orders of magnitude, from pIC<sub>50</sub> values of 3.70 to 6.26, were used as training set; whereas the test set comprised 9 compounds (shown in Table 3, also presented later in this work) with affinity values ranging from 3.70 to 6.28. These training set and test set were used to build and validate nine 3-D QSAR models.

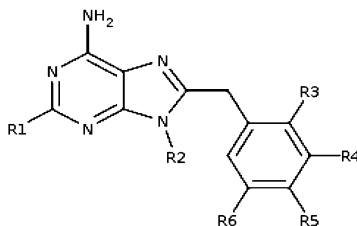
**N-terminal Hsp90 Binding Site-Inhibitor Complex Structure Preparation.** The 24 complexes of the Hsp90 protein downloaded from the PDB<sup>10</sup> were used to build the training and test sets (Tables 2 and 3, respectively). In order to reduce the geometric inaccuracy of the downloaded structures, the complexes were submitted to a molecular modeling protocol, previously tested and reported:<sup>39</sup> (1) using the Chimera software, the Hsp90 complexes were superimposed using the command-line implementation of MatchMaker; (2) all crystal water molecules were deleted, and hydrogen atoms

were added using the leap module of AMBER suite; (3) all ligands were then individually checked for correct protonation states at pH 7.4, and the parameters were calculated by means of the Antechamber module of AMBER suite; (4) the complexes were solvated by water (SOLVATEOCT command) in a box extending 10 Å (TIP3 model), neutralized with Na<sup>+</sup> and Cl<sup>-</sup> ions, and then refined by a single-point minimization using the Sander module of AMBER suite (RMSD values between the original and the minimized PDB complexes, used as training and test sets, are reported in Table S1 in the Supporting Information); and (5) finally, the minimized complexes were realigned with MatchMaker, using the same reference complex (see the Alignment Rules section), and the characterizing minimized ligands were extracted to compose the training set and test set.

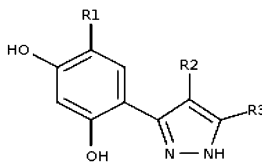
**Molecular Modeling and 3-D QSAR.** All calculations were done on a Linux cluster running a Debian operating system. Several 3-D QSAR studies on the 24 inhibitors extracted from experimentally observed complexes with Hsp90 were carried out using the 3-D QSAutogrid/R procedure.<sup>28</sup>

**Molecular Interaction Fields Calculation.** As implemented in the 3-D QSAutogrid/R procedure,<sup>28</sup> MIFs were generated using the AutoGrid Software (AutoDock Suite,<sup>40</sup> based on the AMBER united-atom force field), considering eight different probes (see Table S6 in the Supporting Information for probe description). Interaction energies between the selected probes and each molecule were computed using a grid spacing of 1 Å; the *xyz* coordinates (in angstroms) of the grid rectangular box used for the computation were Xmin/Xmax = 25.721/47.721, Ymin/Ymax = 29.276/53.276, Zmin/Zmax = 18.289/38.289 to embrace all the SB aligned inhibitors spanning 10 Å in all three dimensions (see Table S8 in the Supporting Information).

**Statistical Analysis.** By application of the D2M package,<sup>28</sup> eight 3-D QSAR PLS models were built. Two cross-validation (CV package)<sup>28</sup> procedures were applied during the model definition to assess both models' quality and robustness: (1) leave-one-out (LOO) and (2) *k*-fold (KF, 5-random groups and 100 iterations). Optimization of the raw models (Table S2 in the Supporting Information) was conducted through the Combinatorial Analysis of Pretreatment Parameters procedure (CAPP package;<sup>28</sup> see Table 4), using *k*-fold cross-validation with 5-random groups and 100 iterations while monitoring the  $q^2$  and SDEP values. Consequently, 4961 combinations were

Table 1. Hsp90 Inhibitors Datasets: (A) Purine, (B) Pyrazole Scaffold Derivatives, and (C) 2-Aminopyrimidine, Triazine, Pyrazole, and *N*-Aryl-benzimidazolone<sup>c</sup>**A) Purine<sup>16</sup> scaffold derivatives**

| entry           | PDB code | R1 | R2         | R3  | R4           | R5  | R6  | IC <sub>50</sub><br>ATPase(μM) |
|-----------------|----------|----|------------|-----|--------------|-----|-----|--------------------------------|
| 1 <sup>a</sup>  | 1UY8     | H  | n-butyl    | H   | OMe          | H   | H   | 75                             |
| 2 <sup>a</sup>  | 1UYC     | H  | n-butyl    | OMe | H            | H   | OMe | 41                             |
| 3 <sup>a</sup>  | 1UYD     | H  | n-butyl    | Cl  | OMe          | OMe | OMe | >200                           |
| 4 <sup>a</sup>  | 1UYE     | H  | 1-pentynyl | Cl  | OMe          | OMe | OMe | >200                           |
| 5 <sup>a</sup>  | 1UYG     | F  | H          | OMe | H            | H   | OMe | 53.5                           |
| 6 <sup>a</sup>  | 1UYH     | F  | n-butyl    | OMe | H            | H   | OMe | 14.3                           |
| 7 <sup>a</sup>  | 1UYK     | F  | n-butyl    | H   | OCH2O bridge |     | H   | 17.1                           |
| 8 <sup>b</sup>  | 1UY6     | H  | n-butyl    | H   | OMe          | OMe | OMe | >200                           |
| 9 <sup>b</sup>  | 1UY7     | H  | n-butyl    | H   | H            | OMe | H   | >200                           |
| 10 <sup>b</sup> | 1UY9     | H  | n-butyl    | H   | OCH2O bridge |     | H   | 15.3                           |
| 11 <sup>b</sup> | 1UYF     |    |            |     |              |     |     |                                |
| 12 <sup>b</sup> | 1UYI     | F  | 1-pentynyl | OMe | H            | H   | OMe | 4.1                            |

**B) Pyrazole<sup>17, 37</sup> scaffold derivatives**

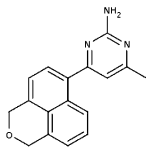
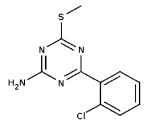
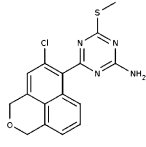
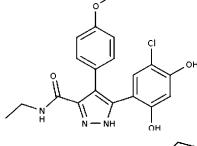
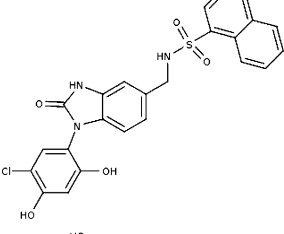
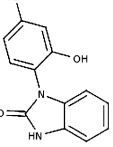
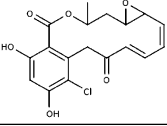
| entry           | PDB code | R1 | R2           | R3     | IC <sub>50</sub><br>ATPase(μM) |
|-----------------|----------|----|--------------|--------|--------------------------------|
| 13 <sup>a</sup> | 2BT0     | Et |              | Me     | 5.7                            |
| 14 <sup>a</sup> | 2CCS     | Cl | 1-piperazine | H      | 8.2                            |
| 15 <sup>b</sup> | 2CCT     | Cl | 1-piperazine | CONHEt | 6.3                            |
| 16 <sup>a</sup> | 2CCU     | Cl |              | H      | 1.3                            |

**C) 2-aminopyrimidine,<sup>7</sup> triazine, pyrazole and *N*-aryl-benzimidazolone<sup>38</sup>**

| entry           | PDB code | IC <sub>50</sub><br>ATPase(μM) |
|-----------------|----------|--------------------------------|
| 17 <sup>a</sup> | 3B25     | 9.6                            |

Table 1. continued

C) 2-aminopyrimidine,<sup>7</sup> triazine, pyrazole and *N*-aryl-benzimidazolone<sup>38</sup>

| entry           | PDB code |                                                                                     | IC <sub>50</sub><br>ATPase(μM) |
|-----------------|----------|-------------------------------------------------------------------------------------|--------------------------------|
| 18 <sup>a</sup> | 3B26     |    | 6.9                            |
| 19 <sup>b</sup> | 3B27     |    | 0.75                           |
| 20 <sup>a</sup> | 3B28     |   | 3                              |
| 21 <sup>a</sup> | 3OWB     |   | 0.88                           |
| 22 <sup>a</sup> | 3OWD     |  | 0.55                           |
| 23 <sup>b</sup> | 3OW6     |  | 85                             |
| 24 <sup>b</sup> | 4EGK     |  | 0.2                            |

<sup>a</sup>Fifteen (15) compounds were used as the training set (see Table 2) and 9 compounds were used as the test set (see Table 3). Data shown represent known values for ATPase activity. <sup>b</sup>Compounds comprising the training set. <sup>c</sup>Compounds used as the external test set.

processed for each 3-D QSAR model, applying 5% sPRESS reduction<sup>28</sup> to select the optimal pretreatment combination and derive the pretreated PLS models; this led to an average  $q^2$  KSFCV value increment equal to 4% (see Tables S3 and S4 in the Supporting Information). In addition, the presence of chance correlations was investigated via the scrambling approach, Y-Scrambling (package YS),<sup>28</sup> using 100 iterations. No further variable selection steps were performed due to the obtained satisfactory overall statistical coefficients, together with the absence of chance correlations (Table 5 and Table S3 in the Supporting Information). Then, a MP 3-D QSAR PLS model was derived (MPGRS package)<sup>28</sup> by selecting the most informative subregions for each of the eight considered probes (Table S6 in the Supporting Information); the same CV and scrambling procedures as those used on the monoprobe models were performed, and the optimal MP 3-D QSAR model, characterized by statistical coefficients similar to those of the

monoprobe models (Table 6, Figure S9 in the Supporting Information), was selected according to the  $q^2_{FL:SL}$  values.<sup>28</sup> No further variable selection procedures were performed. Finally, the most relevant MIFs subregions were selected applying a  $q^2$  threshold value of 0.5.

**Alignment Rules.** The 3-D coordinates of the training-set and test-set compounds were taken from their respective minimized PDB complexes aligned on the 3B28 one, that having the best resolution (1.35 Å) was taken as the template.

## RESULTS AND DISCUSSION

**Structure-Based Design.** Several 3-D QSAR PLS models, characterized by a training set composed of 15 Hsp90 inhibitors comprised of purine,<sup>16</sup> pyrazole,<sup>17,37</sup> 2-aminopyrimidine,<sup>7</sup> triazine,<sup>7</sup> and *N*-aryl-benzimidazolone<sup>38</sup> derivatives (Table 2), were built through the 3-D QSAutogrid/R<sup>28</sup> protocol. Eight monoprobe 3-D QSAR PLS models were generated (see



Table 2. Training Set Activities

| entry | PDB code | IC <sub>50</sub> ATPase (μM) | pIC <sub>50</sub> <sup>a</sup> |
|-------|----------|------------------------------|--------------------------------|
| 1     | 1UY8     | 75                           | 4.12                           |
| 2     | 1UYC     | 41                           | 4.39                           |
| 3     | 1UYD     | 200                          | 3.70                           |
| 4     | 1UYE     | 200                          | 3.70                           |
| 5     | 1UYG     | 53.5                         | 4.27                           |
| 6     | 1UYH     | 14.3                         | 4.84                           |
| 7     | 1UYK     | 17.1                         | 4.77                           |
| 13    | 2BT0     | 5.7                          | 5.24                           |
| 14    | 2CCS     | 8.2                          | 5.09                           |
| 16    | 2CCU     | 1.3                          | 5.89                           |
| 17    | 3B25     | 9.6                          | 5.02                           |
| 18    | 3B26     | 6.9                          | 5.16                           |
| 20    | 3B28     | 3                            | 5.52                           |
| 21    | 3OWB     | 0.88                         | 6.06                           |
| 22    | 3OWD     | 0.55                         | 6.26                           |

$$^a\text{pIC}_{50} = -\log_{10}[\text{IC}_{50}\text{ATPase}(\mu\text{M}) \times 10^{-6}].$$

Table 3. Test Set Activities

| entry | PDB code | IC <sub>50</sub> ATPase (μM) | pIC <sub>50</sub> <sup>a</sup> |
|-------|----------|------------------------------|--------------------------------|
| 8     | 1UY6     | 200                          | 3.70                           |
| 9     | 1UY7     | 200                          | 3.70                           |
| 10    | 1UY9     | 15.3                         | 4.82                           |
| 11    | 1UYF     | 30                           | 4.52                           |
| 12    | 1UYI     | 41                           | 5.39                           |
| 15    | 2CCT     | 6.3                          | 5.20                           |
| 19    | 3B27     | 0.75                         | 6.12                           |
| 23    | 3OW6     | 85                           | 4.07                           |
| 24    | 4EGK     | 0.2                          | 6.28                           |

$$^a\text{pIC}_{50} = -\log_{10}[\text{IC}_{50}\text{ATPase}(\mu\text{M}) \times 10^{-6}].$$

Table 4. CAPP Settings Adopted for the 3-D QSAR Models

| min value | parameter                | max value | step  |
|-----------|--------------------------|-----------|-------|
| 0         | positive cutoff, PCO     | 10        | 1     |
| 0         | zeroing <sup>a</sup>     | 0.1       | 0.01  |
| 0         | minimum SD cutoff, MSDCO | 1         | 0.025 |

<sup>a</sup>Zeroing of very low data points.

Table 5. 3-D QSAutogrid/R PLS Models Statistical Results<sup>a</sup>

| model | P  | PC | r <sup>2</sup> | q <sup>2</sup> <sub>LOO</sub> | q <sup>2</sup> <sub>KSFCV</sub> | r <sup>2</sup> <sub>YS</sub> | q <sup>2</sup> <sub>YS</sub> |
|-------|----|----|----------------|-------------------------------|---------------------------------|------------------------------|------------------------------|
| 1     | A  | 2  | 0.93           | 0.61                          | 0.59                            | 0.69                         | -0.46                        |
| 2     | C  | 2  | 0.93           | 0.61                          | 0.58                            | 0.69                         | -0.50                        |
| 3     | HD | 2  | 0.85           | 0.55                          | 0.54                            | 0.44                         | -0.52                        |
| 4     | NA | 2  | 0.93           | 0.61                          | 0.59                            | 0.71                         | -0.42                        |
| 5     | N  | 2  | 0.93           | 0.62                          | 0.58                            | 0.68                         | -0.52                        |
| 6     | OA | 2  | 0.93           | 0.61                          | 0.59                            | 0.69                         | -0.47                        |
| 7     | e  | 2  | 0.93           | 0.63                          | 0.60                            | 0.73                         | -0.50                        |
| 8     | d  | 1  | 0.72           | 0.61                          | 0.60                            | 0.12                         | -0.22                        |

<sup>a</sup>The CAPP process was applied. Parameters given in this table are as follows: P, Autogrid Probe; PC, optimal number of principal components/latent variables; r<sup>2</sup>, conventional square-correlation coefficient; q<sup>2</sup><sub>LOO</sub>, cross-validation correlation coefficient using the leave-one-out method; q<sup>2</sup><sub>KSFCV</sub>, cross-validation correlation coefficient using the *k*-fold cross-validation with 5 random groups and 100 iterations; r<sup>2</sup><sub>YS</sub>, average square correlation coefficient obtained after Y-scrambling process using 100 iterations; q<sup>2</sup><sub>YS</sub>, average cross-validation correlation coefficient using the leave-one-out (LOO) method obtained after Y-scrambling process, using 100 iterations.

Table 6. Statistical Results Obtained from MPGRS Analysis<sup>a</sup>

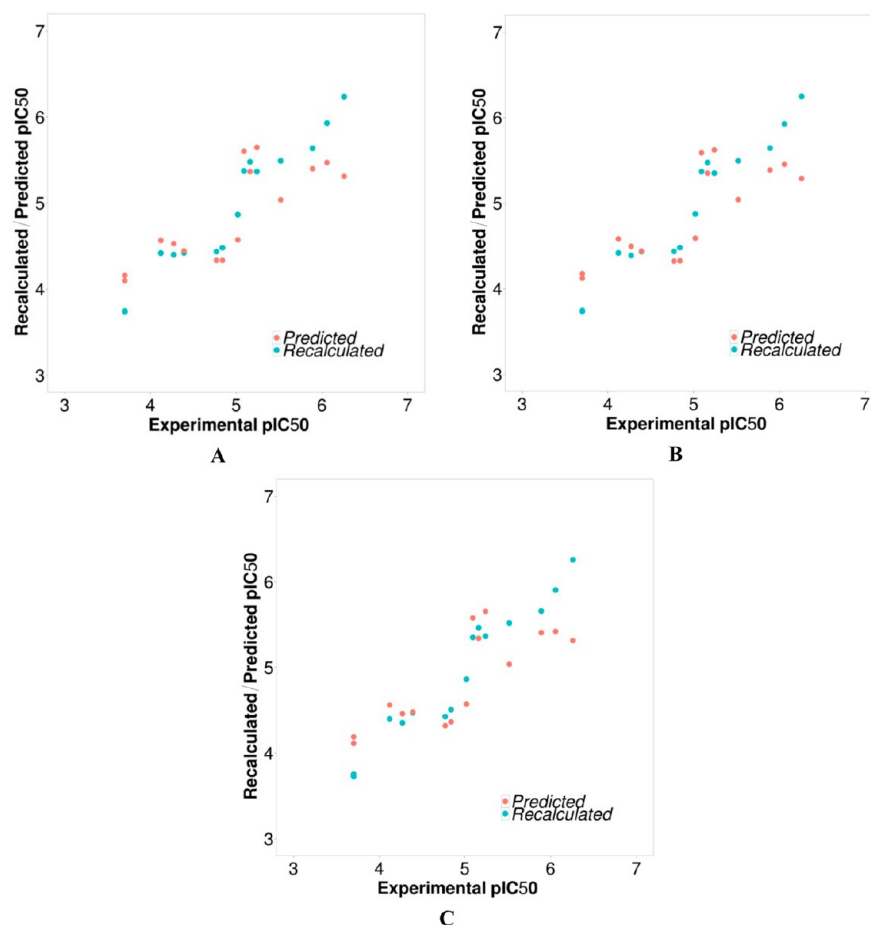
| MPGRS 3-D QSAR                  |       |
|---------------------------------|-------|
| parameter                       | value |
| PC <sub>FL:SL</sub>             | 1:2   |
| r <sup>2</sup>                  | 0.94  |
| q <sup>2</sup> <sub>LOO</sub>   | 0.69  |
| q <sup>2</sup> <sub>KSFCV</sub> | 0.68  |
| SDEP <sub>LOO</sub>             | 0.43  |
| SDEP <sub>KSFCV</sub>           | 0.44  |
| r <sup>2</sup> <sub>YS</sub>    | 0.67  |
| q <sup>2</sup> <sub>YS</sub>    | -0.50 |

<sup>a</sup>Parameters used in this table: PC<sub>FL:SL</sub>, optimal number of principal first level (FL) and second level (SL) components for the MPGRS model; r<sup>2</sup>, conventional square-correlation coefficient; q<sup>2</sup><sub>LOO</sub>, cross-validation correlation coefficient using the leave-one-out method; q<sup>2</sup><sub>KSFCV</sub>, cross-validation correlation coefficient using the *k*-fold cross-validation with 5 random groups and 100 iterations; SDEP<sub>LOO</sub>, standard deviation error of prediction using the leave-one-out method; SDEP<sub>KSFCV</sub>, standard deviation error of prediction using the *k*-fold cross-validation with 5 random groups and 100 iterations; r<sup>2</sup><sub>YS</sub>, average square correlation coefficient obtained after Y-scrambling process using 100 iterations; and q<sup>2</sup><sub>YS</sub>, average cross-validation correlation coefficient using the leave-one-out method obtained after Y-scrambling process using 100 iterations.

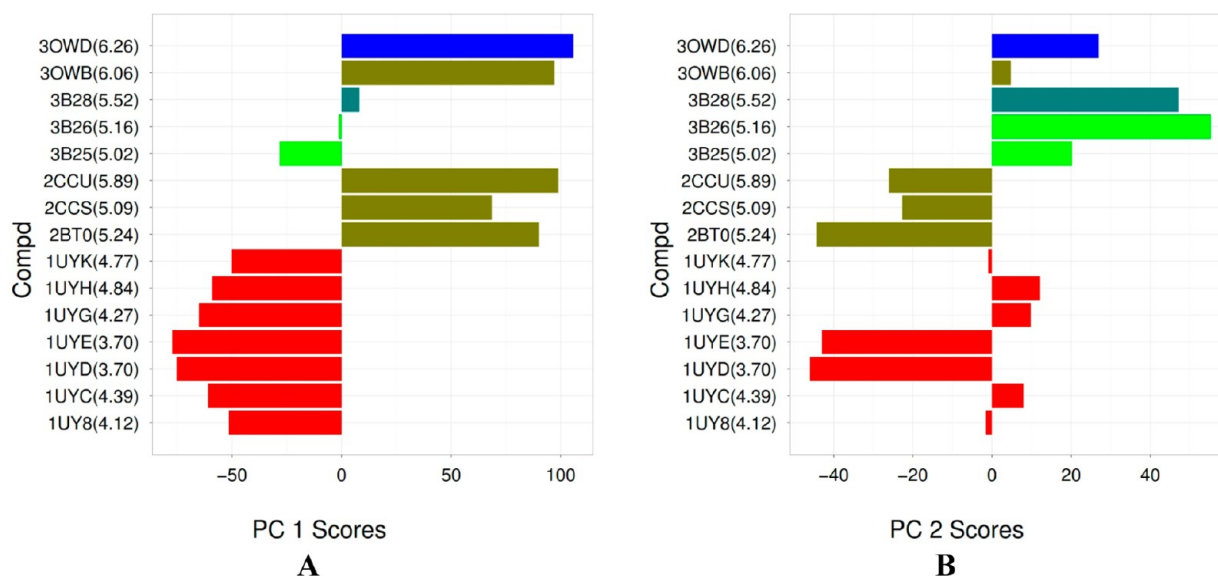
Tables S2 and S6 in the Supporting Information) and optimized via the CAPP<sup>28</sup> procedure (see Tables 4 and 5, as well as Table S3 in the Supporting Information). A final multiprobe (MP) model then was derived via application of the MPGRS<sup>28</sup> (see the Application of Multi-Probe Guided Region-Variable Selection section below) to define the pharmacophoric features required for Hsp90 inhibitor activity and quantitatively correlate them with molecular structures. Activity data, originally reported as IC<sub>50</sub> values from ATPase assays, were transformed to pIC<sub>50</sub> (-log<sub>10</sub>(IC<sub>50</sub>)) values on a molar basis. Satisfactory statistical results were obtained, confirming the predictive capabilities and robustness of the monoprobes models (see Table 5, as well as Table S3 in the Supporting Information and the discussion given earlier in the Statistical Analysis section). Three of them were selected for further investigations: these were obtained from A, N, and OA probes (see Figure 3).

The association of score and PLS-loadings plots was useful to identify the most influencing variables (see Figures S1–S5 in the Supporting Information). Considering the aromatic probe (A), PC1 scores show a clear discrimination between lower active compounds (pIC<sub>50</sub> ≤ 5) and higher active compounds (pIC<sub>50</sub> > 5; see Figure 4A, as well as Figures S1A1 and S1A3 in the Supporting Information): those characterized by the lowest activities are clustered in the negative field, comprising of all the purine-based inhibitors<sup>16</sup> (1UY8, 1UYC, 1UYD, 1UYE, 1UYG, 1UYH, and 1UYK, denoted by red bars in Figure 4A) and 3B25<sup>7</sup> (light green bar), whereas pyrazole<sup>17,37</sup> and N-aryl-benzimidazolone<sup>38</sup> derivatives (2BT0, 2CCS, 2CCU, 3OWB, and 3OWD, denoted by olive drab and blue bars, respectively) are grouped in the positive PC1 score field (these, on average, have the highest activities). Finally, 2-aminopyrimidine and triazine derivatives<sup>7</sup> (3B26 and 3B28, denoted by light green and dark green bars, respectively) were characterized by low absolute PC1 score values.

Even if they are useful in classical QSAR analyses, 2-D PLS-loadings plots (see Figures S2A–S2C in the Supporting Information) are not capable of highlighting the most important variables in a volumetric space, making them



**Figure 3.** Recalculated ( $r^2$ , cyan points) and predicted ( $q^2$   $k$ -5-fold, orange points)  $pIC_{50}$  values: (A) from the A probe model at PC2, (B) from the N probe model at PC2, and (C) from the OA probe model at PC2.



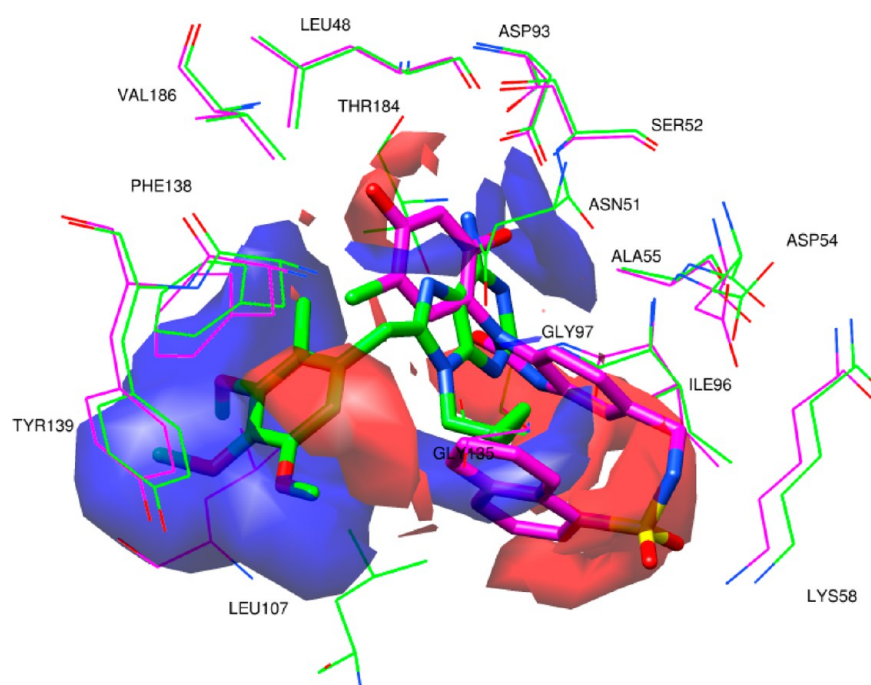
**Figure 4.** Aromatic atom (A) probe scores plots. X-axes are related to score values and Y-axes are related to compounds ( $pIC_{50}$ ). Compounds are color-coded, according to their scaffold: purine-based inhibitors (entries 1, 2, 3, 4, 5, 6, 7), denoted by red bars; 2-aminopyrimidine inhibitors (entries 17 and 18), denoted by light green bars; pyrazole inhibitors (entries 13, 14, 16, and 21), denoted by olive drab bars; *N*-aryl-benzimidazolone inhibitors (entry 22), denoted by blue bars; and triazine inhibitors (entry 20), denoted by dark green bars.

ineffective for 3-D QSAR investigation. In contrast, superimposing the 3-D PLS-loadings plots (LB-extracted data) with the residues active site (SB info) led to identifying the

putatively most involved residues for each PLS score cluster (see Figures S3A1, S3A2, and S4A in the Supporting Information): a first series composed of GLY97, ILE96,

Table 7. LB/SB Convergence Points

| No. | reference probe | reference PC | information                     | LB/SB convergence                                                                                                                                                                       |
|-----|-----------------|--------------|---------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1   | A, N, OA/MP     | 1/1:1        | PLS scores, PLS loadings, X-ray | interactions with GLY97, ILE96, ASP54, ALA55, LEU48, LYS58, VAL 186 are desirable                                                                                                       |
| 2   | A, N, OA/MP     | 2/1:2        | PLS scores, PLS loadings, X-ray | steric hindrance limitation near TYR139, TRP162                                                                                                                                         |
| 3   | A, N, OA/MP     | 2/1:2        | PLS-coefficients, X-ray         | hydrophobic subpockets composed of LYS58, ILE96 and GLY97, and VAL186 and LEU48, respectively, should be fulfilled                                                                      |
| 4   | A, N, OA        | 2            | PLS-coefficients, X-ray         | steric hindrance limitation exerted by TYR139 and PHE138                                                                                                                                |
| 5   | A, N, OA/MP     | 2/1:2        | PLS-coefficients, X-ray         | electrostatic interactions with ALA55, SER52, ASP93, THR184, and ASN51 should be preferable                                                                                             |
| 6   | MP              | 1:2          | PLS-coefficients, X-ray         | electrostatic interactions with the <i>para</i> -hydroxylic group of TYR139 could be advantageous; weak attractive interactions with the indole nitrogen of TRP162 could be established |



**Figure 5.** PLS-coefficients contour maps considering 2 PCs. Only the highest active compound (3OWD in magenta) and the lowest active compound (1UYD in green) are shown. PLS-coefficients contour maps derived from probe A analysis (contour levels: 65%, positive red, negative blue; UCSF Chimera sessions are available as Supporting Information).

ASP54, ALA55, LEU48, LYS58, and VAL186 for positive clustered molecules (2BT0, 2CCS, 2CCU, 3OWB, 3OWD) and a second series composed of LEU107, TYR139, and PHE138 for the negatives (1UY8, 1UYC, 1UYD, 1UYE, 1UYG, 1UYH, 1UYK, and 3B25). This evidence revealed the presence of two “posing” areas as the main discriminating aspect between these two molecular series, suggesting the importance in establishing interactions with the first series residues (LB/SB convergence point 1, Table 7).

From PC2 (see Figure 4B, as well as Figures S1A2, S1A3, S3B1, S3B2, S4B, and S5 in the Supporting Information), a specific area, in which different clustered molecules show a diverse space filling feature, can be highlighted: for example, considering 1UYD, 3B25, and 3B26 (the first belonging to the negative cluster and the latter two belonging to the positive cluster; see Figure 4B, as well as Figures S5A1 and S5A2 in the Supporting Information), a steric group overlapping the

corresponding benzyl *para*-methoxy group of 1UYD (recall Table 1) is detrimental for the activity. In fact, 3B25 and 3B26, which are compounds that fill only regions overlapped with the 1UYD benzyl *meta*- and *ortho*- groups, are endowed with higher activities. Furthermore, from the superimposed crystals, the presence of a conformationally conserved TYR139 side chain in this area suggests a consequent steric limitation for bulkier ligands (LB/SB convergence point 2, Table 7).

A similar clustering profile was observed for N and OA probes from PC1 (Figures S1B1, S1B3, S1C1, and S1C3 in the Supporting Information) and PC2 (Figures S1B2, S1B3, S1C2 and S1C3 in the Supporting Information).

At a higher level of information, analyses of PLS-coefficients (ligand-based information) overlapped to the binding pockets (structure-based information) allow a qualitative/quantitative definition of the different ligand/receptor interactions,



extending the evidence provided by the PLS-loadings/PLS-scores interpretation.

A good agreement was observed between the quantitative derivation (PLS-coefficients) and the experimental information (binding pocket residues, Figure 5): i.e., considering the aromatic atom probe A at PC2, the biggest positive coefficients area was mainly located in the subpocket delimited by LYS58, ILE96, and, GLY97 residues (Figure 5), suggesting that bioactivity improvement could be obtained by filling this hydrophobic area (LB/SB convergence point 3, Table 7), in agreement with previous reports.<sup>7,22</sup> In addition, positive coefficients were found near GLY135, suggesting its potential involvement in hydrophobic interactions. Negative PLS-coefficients were mainly placed in a space surrounded by LEU107, TYR139, and PHE138, confirming, as anticipated above (see PLS-loadings interpretation), a steric hindrance limitation exerted by TYR139 and PHE138 side chains (LB/SB convergence point 4; see Table 7 and Figure 5). As a consequence, molecules such as the ligand in 1UYD, which overlaps this area, are less active.

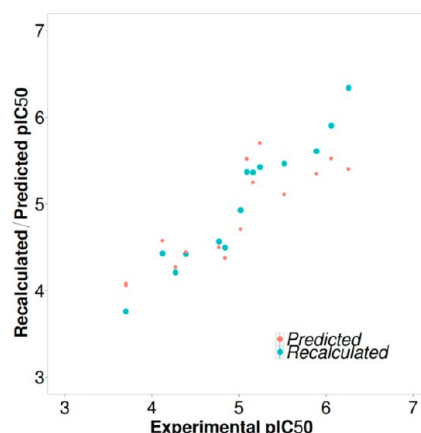
Further smaller negative PLS-coefficients area runs among ALA55, SER52, and ASP93 (small blue area in Figure 5) and, considering also both OA probe PLS-coefficients (also negative) and MIF field plots (negative orange fields, see Figure S6A-A2 in the Supporting Information), these suggest the presence of both steric and electrostatic features in this zone. In fact, these features represent a further LB/SB convergence point (LB/SB convergence point 5, Table 7) by the presence of electrostatic interactions with the three above listed residues (i.e., H bonds), in agreement with previous reports.<sup>41</sup>

Similar information was derived from the N probe (Figure S6B in the Supporting Information), confirming the above-reported assumptions; therefore, the relative discussion is omitted to avoid redundancy.

**Application of Multi-Probe Guided Region-Variable Selection.** A Multi Probe (MP) 3-D QSAR model was also derived by the application of the Multiprobe Guided Region Selection (MPGRS package).<sup>28</sup> As reported, the MP 3-D QSAR model is obtained at the optimal second level PC of the selected first-level PC subregions; in this case,  $PC_{FL,SL} = PC_{1,2}$ , and its associated statistical coefficients were slightly improved, with respect to those of the monoprobe models (see Table 6 and Figure 6).

Briefly, the MP model was derived by applying a  $q^2$  threshold value of 0.5 (different threshold values did not lead to better models), thus selecting only the most relevant MIFs subregions (see Figure 7, as well as Figure S9 and Table S6 in the Supporting Information for color coding reference) to compose the final multiprobe MIF. Analyses of multiprobe scores and loadings led to similar conclusions as the monoprobe models (LB/SB convergence points 1 and 2, Table 7); however, a higher level interpretation allowed identification of both the areas of major interest and the more-profitable associated features (see Figure 8, as well as Figures S10–S12 in the Supporting Information).

At deep analysis, of particular interest are areas simultaneously characterized by negative MP PLS-coefficients associated to atom probes able to derive negative fields, such as OA, NA, and HD. In particular, two distinct areas of negative MP PLS-coefficients derived from the OA probe are among THR184, ASP93, SER52, and ASN51 (Figure 9A) and TRP162, TYR139, and LEU107 (Figure 9B), suggesting that



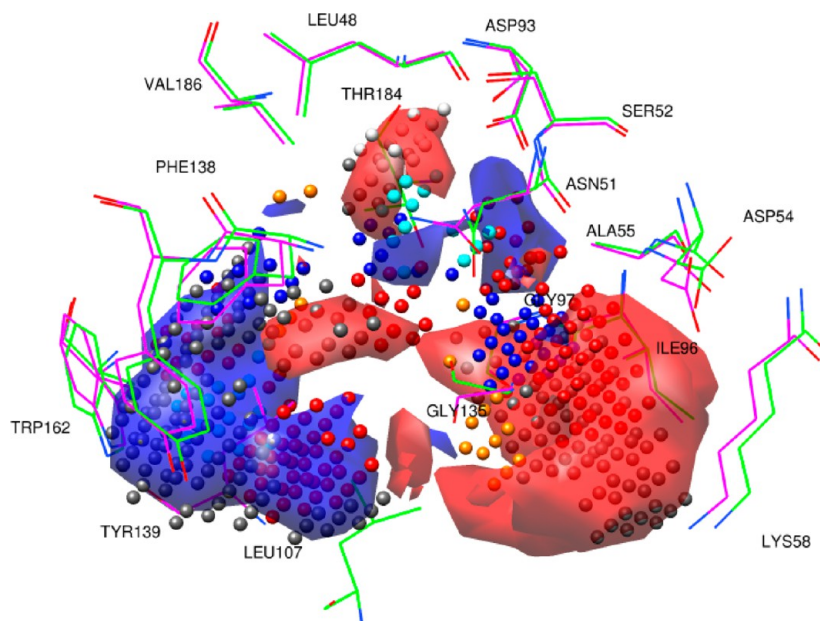
**Figure 6.** MPGRS. Recalculated ( $r^2$ , cyan points) and predicted ( $q^2$  k-5-Fold, orange points)  $pIC_{50}$  values from the multiprobe (MP) model at  $PC_{1,2}$ .

the establishment of attractive interactions (i.e., hydrogen bond) should be advantageous in the proximity of the first residues' series as for the *para*-hydroxylic group of TYR139 (LB/SB convergence points 5 and 6; see Table 7). Furthermore, between TYR139 and TRP162, a negative PLS-coefficient area (Figure 9B), characterized by NA and C probes, was observed, which suggests, besides the steric limitation, that some weak attractive interaction with the indole nitrogen of TRP162 could be established (LB/SB convergence points 2 and 6; see Table 7).

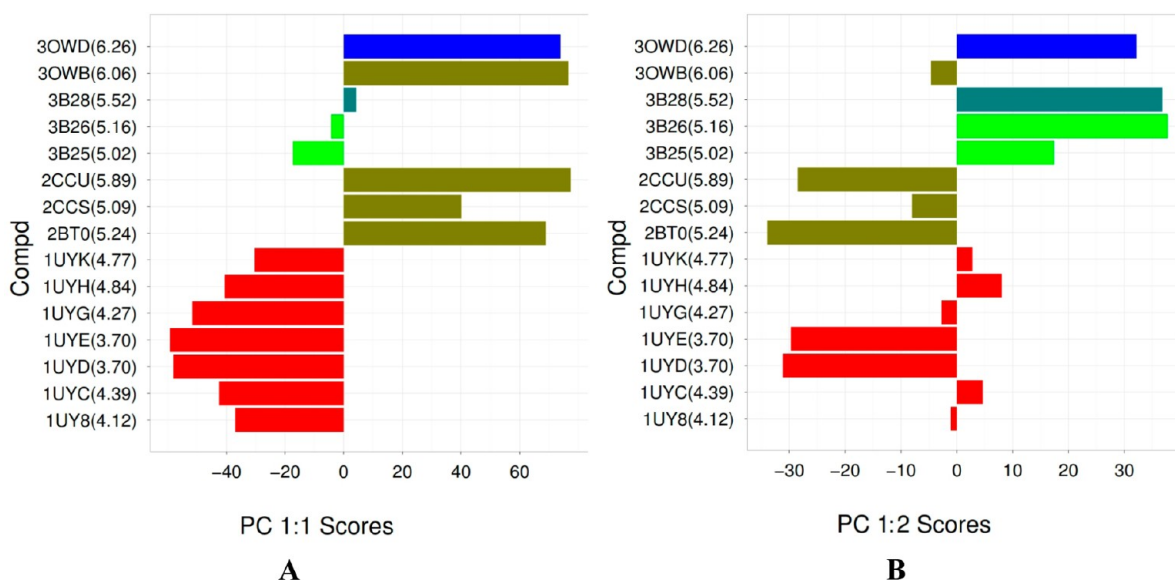
Finally, two positive coefficient polyhedra, characterized both by steric/hydrophobic and electrostatic probes (A, C, OA, N, and NA), were recognized in the proximity of LYS58 and ILE96, and VAL186 and LEU48 (Figures 9C and 9D, respectively), confirming the importance of placing steric groups in these areas (LB/SB convergence point 3, Table 7). Interestingly, a certain parallelism between the type of side chain and the type of selected probe (i.e., NA and OA probes near the indole nitrogen of TRP162, C probes in the proximity of the aromatic rings of TRP162 and TYR139, OA probes near the hydroxylic group of TYR139, and, again, OA probes and C probes near ASP54 and LYS58, respectively, etc.; see Figure S9 in the Supporting Information) was noticed, possibly implying the model's ability to predict the chemical properties of the binding pocket at each grid point.

As a further assessment, the multiprobe information was compared with a previously reported pharmacophoric model,<sup>22,42</sup> showing a high degree of agreement (data not shown).

**External Test Set Prediction Analysis.** The external test set composed of 9 Hsp90 inhibitors (see data split in the Training and Test Set Selection section, Table 3) was applied to each of the eight 3-D QSAutogrid/R monoprobe models to assess their predictive capability. Generally, low errors of prediction were observed for all models (see Table 8), and the correct activity trend was reproduced (see Figure S7 and Table S5 in the Supporting Information). Only one molecule (3OW6) was predicted, with an average absolute error of prediction (AAEP) of  $>1$  (see Table S5 in the Supporting Information): this bias seems to derive from the LB treatment of the available experimental data in any 3-D QSAR approach. In fact, considering the 3-D QSAR models' indications, 3OW6 was correctly predicted more active than experimentally observed, because of the fact that its benzimidazolone moiety overlaps



**Figure 7.** MPGRS. Key points with PLS-coefficients contour maps (contour levels 75%: positive: red; negative: blue). 3OWD in magenta and 1UYD in green. The points are color-coded, according to that reported in Table S6 in the Supporting Information. UCSF Chimera sessions are available as Supporting Information.

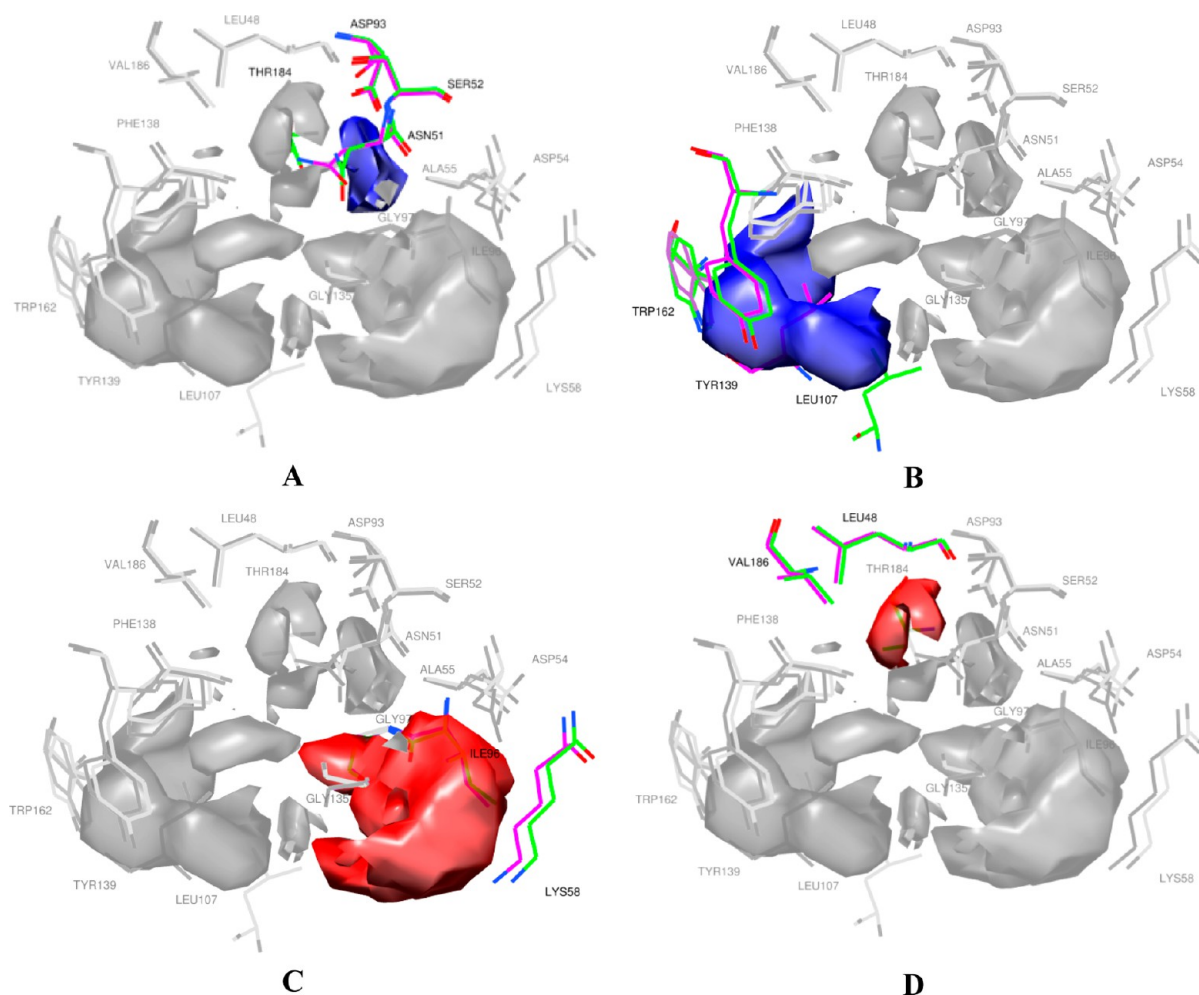


**Figure 8.** Multi Probe (MP) scores plots. X-axes are related to score values and Y-axes are related to compounds ( $\text{pIC}_{50}$ ). Compounds are color-coded according to their scaffold: purine-based inhibitors (entries 1, 2, 3, 4, 5, 6, 7), red bars; 2-aminopyrimidine inhibitors (entries 17 and 18), light green bars; pyrazole inhibitors (entries 13, 14, 16, and 21), olive drab bars; *N*-aryl-benzimidazolone inhibitors (entry 22), blue bars; and triazine inhibitors (entry 20), dark green bars.

the positive PLS-coefficients area discussed in the Structure-Based Design section. A deeper SB analysis revealed that, indeed, in the 3OW6 crystal complex, the benzimidazolone is missing some important interactions with the LYS58 side chain (because of the residue's flexibility; see Figure S8 in the Supporting Information) likely responsible for the decreased activity, reflecting that the LYS58 side chain flexibility is not implemented in the 3-D QSAR model itself. On the other hand, i.e., considering 2CCT, displaying its carboxyaminoethyl moiety overlapped with the positive PLS-coefficients area (Figure S8 in the Supporting Information), the biological activity was slightly overpredicted (AAEP = 0.44); however, in this case, the LYS58

is at hydrogen bonding distance (2.93 Å), thus making favorable ligand/protein interactions.

PLS-coefficients contour maps are useful to graphically represent, for a given 3-D QSAR model, how different types of interactions (according to the used probes and training set molecules) could affect the biological response; differently, activity contribution plots have the capability to show, the recalculated (training set) or predicted (test set) 3-D activity profile of the molecules of each training set or test set, respectively. However, so far, to the best of our knowledge, the latter plots have been applied only on the molecules of the training sets, thereby limiting the analyses of test set



**Figure 9.** MPGRS. (A–D): PLS-coefficients contour maps (contour levels 75%: positive: red; negative: blue), each area is highlighted maintaining its own color.

**Table 8. Test Set Predictions: SDEP Values, Considering the Optimal PCs<sup>a</sup>**

| P  | PC | SDEP <sub>EXT</sub> |
|----|----|---------------------|
| A  | 2  | 0.79                |
| C  | 2  | 0.79                |
| HD | 2  | 0.78                |
| NA | 2  | 0.80                |
| N  | 2  | 0.79                |
| OA | 2  | 0.81                |
| e  | 2  | 0.79                |
| d  | 1  | 0.86                |

<sup>a</sup>Legend of symbols used in this table: P, AutoGrid Probe; PC, optimal number of principal components/latent variables; and SDEP<sub>EXT</sub>, standard deviation error of prediction (or root-mean-squared error of prediction, RMSEP) for the external test set.

compounds to a numerical investigation (predicted biological response and SDEP), or superimposition with the PLS-coefficients maps. To overcome this restriction, in the 3-D QSAutogrid/R procedure, the possibility to derive activity contribution plots also for the external test set molecules is implemented, to perform a comprehensive quantitative analysis, in particular, for untested compounds against the target of interest. This feature is particularly important in the application of the developed 3-D QSAR models to the activity prediction

for molecules dissimilar from those of the training set, as in the case of a virtual screening (VS) protocol.<sup>30</sup>

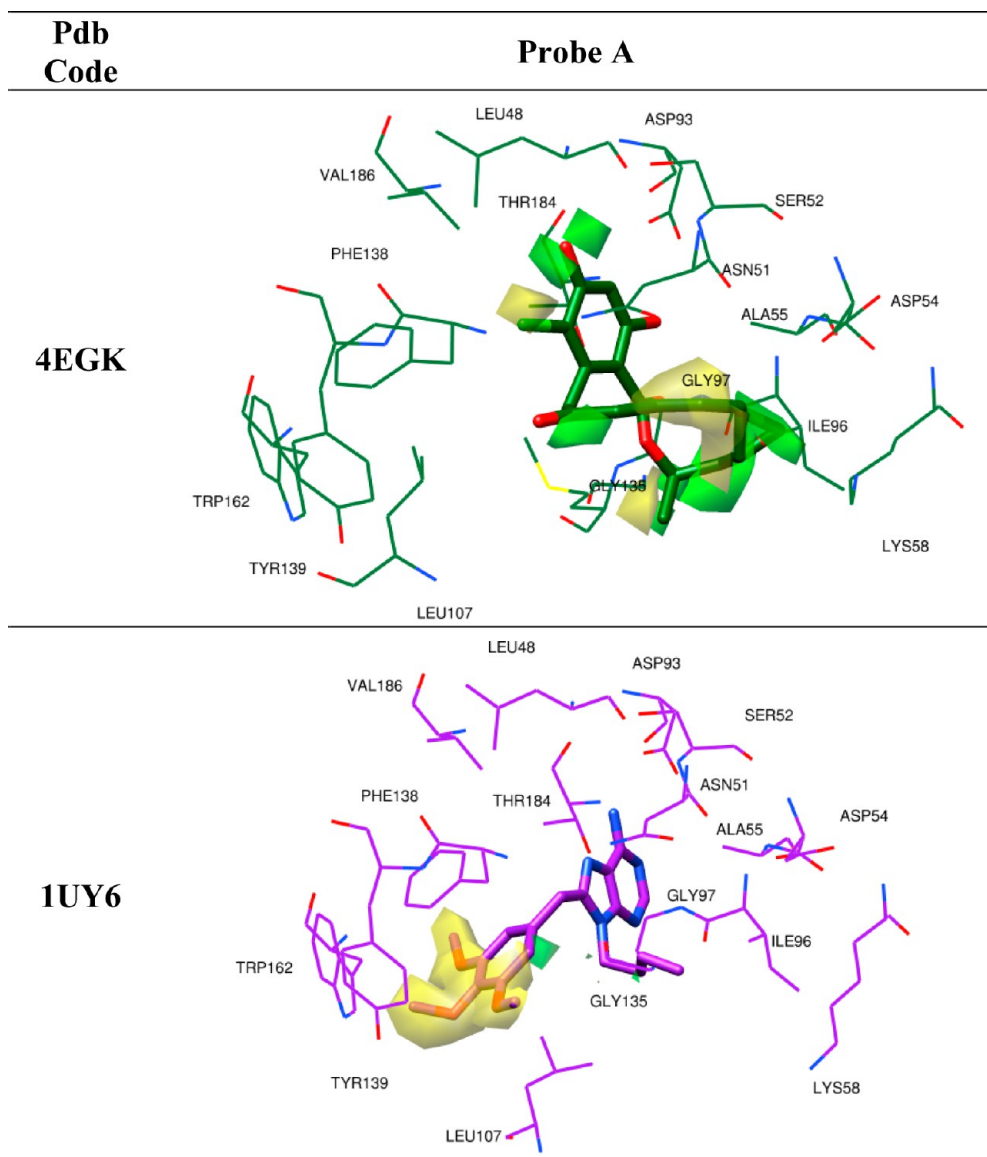
Herein, as an example, the activity contribution plots for the most active and least active molecules of the test set (see Figure 10, PDB codes 4EGK and 1UY6, respectively), are reported for the aromatic probe model (A) at the selected PC (Table 5; see Figure S14 in the Supporting Information for plots derived from the N and OA PLS models).

As for the monoprobe models, the MP is endowed with good predictive capabilities (see Table S7 in the Supporting Information), reproducing, with good approximation, the test set activity trend (Figure S13 in the Supporting Information). As previously reported,<sup>28</sup> the MP approach is characterized by a better explanation of the capacity of the models, while maintaining a similar predictive capability to those of the monoprobe models (MP SDEP<sub>EXT</sub> = 0.81); in this sense, the above-described limitations, with regard to predictions, remain the same: i.e., 3OW6 was predicted, again, with the highest absolute error of prediction (see Table S7 in the Supporting Information).

## CONCLUSION

Structure-based (SB) three-dimensional (3-D) quantitative structure–activity relationship (QSAR) models coupled with an experimental SB alignment method were developed for Hsp90 inhibitors. As discussed, this application allowed us to





**Figure 10.** Activity contribution maps derived from probe A on test set molecules 4EGK and 1UY6. Contour levels: 85%, positive contribution (green), negative contribution (yellow).

highlight several LB/SB convergence points that support how the correct alignment in a 3-D QSAR could enhance the models' robustness, predictiveness, and interpretability. Therefore, the models developed herein can be eligible as a predictive tool in a virtual screening (VS) protocol. In particular the 3-D QSAR models will be used as an ad hoc external scoring function to discover new potential Hsp90 inhibitors. This approach is reported in the second paper in this series.<sup>30</sup>

## ■ ASSOCIATED CONTENT

### Supporting Information

This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [rino.ragno@uniroma1.it](mailto:rino.ragno@uniroma1.it).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

F.B. acknowledges Sapienza Università di Roma ("Progetti per Avvio alla Ricerca" Grant No. C26N12JZCT). R.B.W. acknowledges Washington University in St. Louis (uSTAR Fellowship). R.R. acknowledges Italian Minister of University and Research (MIUR) (PRIN Grant No. 20105YY2HL\_003 and FIRB Grant No. RBFR10ZJQT).

## ■ DEDICATION

Dedicated to Prof. Marino Artico for his 80th birthday.

## ■ REFERENCES

- (1) Buchstaller, H. P.; Eggenweiler, H. M.; Sirrenberg, C.; Gradler, U.; Musil, D.; Hoppe, E.; Zimmermann, A.; Schwartz, H.; Marz, J.; Bomke, J.; Wegener, A.; Wolf, M. Fragment-based discovery of hydroxy-indazole-carboxamides as novel small molecule inhibitors of Hsp90. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 4396–4403.
- (2) Whitesell, L.; Lindquist, S. L. HSP90 and the chaperoning of cancer. *Nat. Rev. Cancer* **2005**, *5*, 761–772.

- (3) Trepel, J.; Mollapour, M.; Giaccone, G.; Neckers, L. Targeting the dynamic HSP90 complex in cancer. *Nat. Rev. Cancer* **2010**, *10*, 537–549.
- (4) Liu, J.; Wang, F.; Ma, Z.; Wang, X.; Wang, Y. Structural Determination of Three Different Series of Compounds as Hsp90 Inhibitors Using 3D-QSAR Modeling, Molecular Docking and Molecular Dynamics Methods. *Int. J. Mol. Sci.* **2011**, *12*, 946–970.
- (5) Chaudhury, S.; Welch, T. R.; Blagg, B. S. Hsp90 as a target for drug development. *ChemMedChem* **2006**, *1*, 1331–1340.
- (6) Nahleh, Z.; Tfayli, A.; Najm, A.; El Sayed, A.; Nahle, Z. Heat shock proteins in cancer: Targeting the “chaperones”. *Future Med. Chem.* **2012**, *4*, 927–935.
- (7) Miura, T.; Fukami, T. A.; Hasegawa, K.; Ono, N.; Suda, A.; Shindo, H.; Yoon, D. O.; Kim, S. J.; Na, Y. J.; Aoki, Y.; Shimma, N.; Tsukuda, T.; Shiratori, Y. Lead generation of heat shock protein 90 inhibitors by a combination of fragment-based approach, virtual screening, and structure-based drug design. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 5778–5783.
- (8) Den, R. B.; Lu, B. Heat shock protein 90 inhibition: rationale and clinical potential. *Ther. Adv. Med. Oncol.* **2012**, *4*, 211–218.
- (9) Ge, J.; Normant, E.; Porter, J. R.; Ali, J. A.; Dembski, M. S.; Gao, Y.; Georges, A. T.; Grenier, L.; Pak, R. H.; Patterson, J.; Sydor, J. R.; Tibbitts, T. T.; Tong, J. K.; Adams, J.; Palombella, V. J. Design, synthesis, and biological evaluation of hydroquinone derivatives of 17-amino-17-demethoxygeldanamycin as potent, water-soluble inhibitors of Hsp90. *J. Med. Chem.* **2006**, *49*, 4606–4615.
- (10) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (11) Biamonte, M. A.; Van de Water, R.; Arndt, J. W.; Scannevin, R. H.; Perret, D.; Lee, W. C. Heat shock protein 90: Inhibitors in clinical trials. *J. Med. Chem.* **2010**, *53*, 3–17.
- (12) Stebbins, C. E.; Russo, A. A.; Schneider, C.; Rosen, N.; Hartl, F. U.; Pavletich, N. P. Crystal structure of an Hsp90-geldanamycin complex: Targeting of a protein chaperone by an antitumor agent. *Cell* **1997**, *89*, 239–250.
- (13) Schulte, T. W.; Akinaga, S.; Soga, S.; Sullivan, W.; Stensgard, B.; Toft, D.; Neckers, L. M. Antibiotic radicicol binds to the N-terminal domain of Hsp90 and shares important biologic activities with geldanamycin. *Cell Stress Chaperones* **1998**, *3*, 100–108.
- (14) Schulte, T. W.; Neckers, L. M. The benzoquinone ansamycin 17-allylamino-17-demethoxygeldanamycin binds to HSP90 and shares important biologic activities with geldanamycin. *Cancer Chemother. Pharmacol.* **1998**, *42*, 273–279.
- (15) Ramanathan, R. K.; Egorin, M. J.; Erlichman, C.; Remick, S. C.; Ramalingam, S. S.; Naret, C.; Holleran, J. L.; TenEyck, C. J.; Ivy, S. P.; Belani, C. P. Phase I pharmacokinetic and pharmacodynamic study of 17-dimethylaminoethylamino-17-demethoxygeldanamycin, an inhibitor of heat-shock protein 90, in patients with advanced solid tumors. *J. Clin. Oncol.* **2010**, *28*, 1520–1526.
- (16) Wright, L.; Barril, X.; Dymock, B.; Sheridan, L.; Surgenor, A.; Beswick, M.; Drysdale, M.; Collier, A.; Massey, A.; Davies, N.; Fink, A.; Fromont, C.; Aherne, W.; Boxall, K.; Sharp, S.; Workman, P.; Hubbard, R. E. Structure-activity relationships in purine-based inhibitor binding to Hsp90 isoforms. *Chem. Biol.* **2004**, *11*, 775–785.
- (17) Dymock, B. W.; Barril, X.; Brough, P. A.; Cansfield, J. E.; Massey, A.; McDonald, E.; Hubbard, R. E.; Surgenor, A.; Roughley, S. D.; Webb, P.; Workman, P.; Wright, L.; Drysdale, M. J. Novel, potent small-molecule inhibitors of the molecular chaperone Hsp90 discovered through structure-based design. *J. Med. Chem.* **2005**, *48*, 4212–4215.
- (18) Brough, P. A.; Aherne, W.; Barril, X.; Borgognoni, J.; Boxall, K.; Cansfield, J. E.; Cheung, K. M.; Collins, I.; Davies, N. G.; Drysdale, M. J.; Dymock, B.; Eccles, S. A.; Finch, H.; Fink, A.; Hayes, A.; Howes, R.; Hubbard, R. E.; James, K.; Jordan, A. M.; Lockie, A.; Martins, V.; Massey, A.; Matthews, T. P.; McDonald, E.; Northfield, C. J.; Pearl, L. H.; Prodromou, C.; Ray, S.; Raynaud, F. I.; Roughley, S. D.; Sharp, S. Y.; Surgenor, A.; Walmsley, D. L.; Webb, P.; Wood, M.; Workman, P.; Wright, L. 4,5-diarylisoazole Hsp90 chaperone inhibitors: Potential therapeutic agents for the treatment of cancer. *J. Med. Chem.* **2008**, *51*, 196–218.
- (19) Rajan, A.; Kelly, R. J.; Trepel, J. B.; Kim, Y. S.; Alarcon, S. V.; Kummar, S.; Gutierrez, M.; Crandon, S.; Zein, W. M.; Jain, L.; Mannargudi, B.; Figg, W. D.; Houk, B. E.; Shnaidman, M.; Brega, N.; Giaccone, G. A phase I study of PF-04929113 (SNX-5422), an orally bioavailable heat shock protein 90 inhibitor, in patients with refractory solid tumor malignancies and lymphomas. *Clin. Cancer Res.* **2011**, *17*, 6831–6839.
- (20) Liu, H. C.; Lyu, P. C.; Leong, M. K.; Tsai, K. C.; Hsiue, G. H. 3D-QSAR studies on PU3 analogues by comparative molecular field analysis. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 731–734.
- (21) Saxena, S.; Chaudhaery, S. S.; Varshney, K.; Saxena, A. K. Pharmacophore-based virtual screening and docking studies on Hsp90 inhibitors. *SAR QSAR Environ. Res.* **2010**, *21*, 445–462.
- (22) Sakkiiah, S.; Thangapandian, S.; John, S.; Kwon, Y. J.; Lee, K. W. 3D QSAR pharmacophore based virtual screening and molecular docking for identification of potential HSP90 inhibitors. *Eur. J. Med. Chem.* **2010**, *45*, 2132–2140.
- (23) Huang, X. Y.; Shan, Z. J.; Zhai, H. L.; Li, L. N.; Zhang, X. Y. Molecular design of anticancer drug leads based on three-dimensional quantitative structure–activity relationship. *J. Chem. Inf. Model.* **2011**, *51*, 1999–2006.
- (24) Roy, K. K.; Singh, S.; Saxena, A. K. Integration-mediated prediction enrichment of quantitative model for Hsp90 inhibitors as anti-cancer agents: 3D-QSAR study. *Mol. Diversity* **2011**, *15*, 477–489.
- (25) Yang, Y.; Liu, H.; Du, J.; Qin, J.; Yao, X. A combined molecular modeling study on a series of pyrazole/isoxazole based human Hsp90alpha inhibitors. *J. Mol. Model.* **2011**, *17*, 3241–3250.
- (26) Jia, J.; Xu, X.; Liu, F.; Guo, X.; Zhang, M.; Lu, M.; Xu, L.; Wei, J.; Zhu, J.; Zhang, S.; Zhang, S.; Sun, H.; You, Q. Identification, Design and Bio-Evaluation of Novel Hsp90 Inhibitors by Ligand-Based Virtual Screening. *PLoS One* **2013**, *8*, No. e59315.
- (27) Zhao, H.; Moroni, E.; Yan, B.; Colombo, G.; Blagg, B. S. J. 3D-QSAR-Assisted Design, Synthesis, and Evaluation of Novobiocin Analogues. *ACS Med. Chem. Lett.* **2013**, *4*, 57–62.
- (28) Ballante, F.; Ragno, R. 3-D QSAutogrid/R: An alternative procedure to build 3-D QSAR models. Methodologies and applications. *J. Chem. Inf. Model.* **2012**, *52*, 1674–1685.
- (29) Friggeri, L.; Ballante, F.; Ragno, R.; Musmuca, I.; De Vita, D.; Manetti, F.; Biava, M.; Scipione, L.; Di Santo, R.; Costi, R.; Feroci, M.; Tortorella, S. Pharmacophore assessment through 3-D QSAR: Evaluation of the predictive ability on new derivatives by the application on a series of antitubercular agents. *J. Chem. Inf. Model.* **2013**, *53*, 1463–1474.
- (30) Caroli, A.; Ballante, F.; Wickersham, R. B., III; Corelli, F.; Ragno, R. Hsp90 Inhibitors, Part 2: Combining Ligand-Based and Structure-Based Approaches for Virtual Screening Application. *J. Chem. Inf. Model.* **2014**, DOI: 10.1021/ci400760a.
- (31) Artese, A.; Cross, S.; Costa, G.; Distinto, S.; Parrotta, L.; Alcaro, S.; Ortuso, F.; Cruciani, G. Molecular interaction fields in drug discovery: recent advances and future perspectives. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, 594–613.
- (32) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
- (33) Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S. Generating Optimal Linear Pls Estimations (Golpe)—An Advanced Chemometric Tool for Handling 3D-QSAR Problems. *Quant. Struct.-Act. Relat.* **1993**, *12*, 9–20.
- (34) Cruciani, G.; Watson, K. A. Comparative Molecular-Field Analysis Using Grid Force-Field and Golpe Variable Selection Methods in a Study of Inhibitors of Glycogen-Phosphorylase-B. *J. Med. Chem.* **1994**, *37*, 2589–2601.
- (35) Clark, M.; Cramer, R. D.; Jones, D. M.; Patterson, D. E.; Simeroth, P. E. Comparative molecular field analysis (CoMFA). 2. Toward its use with 3D-structural databases. *Tetrahedron Comput. Methodol.* **1990**, *3*, 47–59.



(36) Wold, S.; Johansson, E.; Cocchi, M. PLS—Partial least squares projections to latent structures. In *3D QSAR in Drug Design*, Vol. 1; Kubinyi, H., Ed.; Kluwer/ESCOM: Dordrecht, The Netherlands, 1993; pp 523–550.

(37) Barril, X.; Beswick, M. C.; Collier, A.; Drysdale, M. J.; Dymock, B. W.; Fink, A.; Grant, K.; Howes, R.; Jordan, A. M.; Massey, A.; Surgenor, A.; Wayne, J.; Workman, P.; Wright, L. 4-Amino derivatives of the Hsp90 inhibitor CCT018159. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2543–2548.

(38) Bruncko, M.; Tahir, S. K.; Song, X.; Chen, J.; Ding, H.; Huth, J. R.; Jin, S.; Judge, R. A.; Madar, D. J.; Park, C. H.; Park, C. M.; Petros, A. M.; Tse, C.; Rosenberg, S. H.; Elmore, S. W. N-aryl-benzimidazolones as novel small molecule HSP90 inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 7503–7506.

(39) Musmuca, I.; Caroli, A.; Mai, A.; Kaushik-Basu, N.; Arora, P.; Ragno, R. Combining 3-D quantitative structure-activity relationship with ligand based and structure based alignment procedures for in silico screening of new hepatitis C virus NSSB polymerase inhibitors. *J. Chem. Inf. Model.* **2010**, *50*, 662–676.

(40) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

(41) Gopalsamy, A.; Shi, M.; Golas, J.; Vogan, E.; Jacob, J.; Johnson, M.; Lee, F.; Nilakantan, R.; Petersen, R.; Svenson, K.; Chopra, R.; Tam, M. S.; Wen, Y.; Ellingboe, J.; Arndt, K.; Boschelli, F. Discovery of benzisoxazoles as potent inhibitors of chaperone heat shock protein 90. *J. Med. Chem.* **2008**, *51*, 373–375.

(42) Sakkiah, S.; Thangapandian, S.; John, S.; Lee, K. W. Pharmacophore based virtual screening, molecular docking studies to design potent heat shock protein 90 inhibitors. *Eur. J. Med. Chem.* **2011**, *46*, 2937–2947.