Elaborate Ligand-Based Modeling Reveals New Nanomolar Heat Shock Protein 90α Inhibitors

Mahmoud A. Al-Sha'er and Mutasem O. Taha*

Department of Pharmaceutical Sciences, Faculty of Pharmacy, University of Jordan, Amman, Jordan

Received June 8, 2010

Heat shock protein (Hsp90α) has been recently implicated in cancer prompting several attempts to discover and optimize new Hsp90α inhibitors. Toward this end, we explored the pharmacophoric space of 83 Hsp90α inhibitors using six diverse sets of inhibitors to identify high-quality pharmacophores. Subsequently, genetic algorithm and multiple linear regression analysis were employed to select an optimal combination of pharmacophoric models and 2D physicochemical descriptors capable of accessing a self-consistent quantitative structure−activity relationship (QSAR) of optimal predictive potential ($r_{67}^2 = 0.811$, $F = 42.8$, $r_{LOO}^2 = 0.748$, r_{PRESS}^2 (against 16 external test inhibitors) $= 0.619$). Three orthogonal pharmacophores emerged in the QSAR equation suggesting the existence of at least three binding modes accessible to ligands within the Hsp90α binding pocket. Receiver operating characteristic (ROC) curves analysis established the validity of QSAR-selected pharmacophores. We employed the pharmacophoric models and associated QSAR equation to screen the national cancer institute (NCI) list of compounds and our in-house-built drugs and agrochemicals database (DAC). Twenty-five nanomolar and low micromolar Hsp90α inhibitors were identified. The most potent were formoterol, amodaquine, primaquine, and midodrine with IC_{50} values of 3, 5, 6, and 20 nM, respectively.

1. INTRODUCTION: HEAT SHOCK PROTEIN

Hsp90α belongs to a family of molecular chaperones that play a pivotal role in the conformational maturation, stability and function of protein substrates within the cell.[1] The interaction of ATP with its binding domain in Hsp90α leads to autophosphorylation of certain tyrosine residues, thus activating this chaperon, and provides the necessary energy for refolding of denatured "client" proteins.[1]

Among the client proteins of Hsp90α are many oncogenes essential for the survival, proliferation, invasion, metastasis and angiogenesis of tumors. In fact, 48 oncogenic proteins have been shown to be dependent upon Hsp90α for conformational activation, including telomerase, Her2 (erbB2), Raf-1, focal adhesion kinase, and the steroid hormone receptors.[2]

The validity of Hsp90α as anticancer target for drug discovery was established by emerging clinical trials employing potent Hsp90α inhibitor 17-allylaminogeldanamycin and the natural Hsp90α inhibitors geldanamycin and radicicol.[3−7]

However, despite the high cellular activity and clinical progression of 17-allylaminogeldanamycin,[8] it has several limitations, for example, poor solubility, hepatotoxicity, and extensive metabolism. These issues have led to significant efforts to identify novel rationally designed small molecular inhibitors of Hsp90α.[9]

The main focus of recent efforts toward the development of new Hsp90α inhibitors concentrate on structure-based ligand design[10−13] and high-throughput screening,[14] with a few ligand-based examples.[15,16] To date, several human Hsp90 X-ray complexes are documented in the Protein Data Bank (e.g., PDB codes 1YET, 2CCS, 2FWY, 2CDD, 2CCU, 2CCT, 1YER, 1YES, 1YC3, 1UYM, 1UYC, 1UY8, 1OSF, 1BYQ, and 2H55 with resolution range 1.50−2.7 Å). However, although considered the most reliable structural information that can be used for drug design, crystallographic structures are limited by inadequate resolution[17] and crystallization-related artifacts of the ligand−protein complex.[18−20] Moreover, crystallographic structures generally ignore structural heterogeneity related to protein anisotropic motion and discrete conformational substrates.[21]

The continued interest in designing new Hsp90α inhibitors and lack of adequate ligand-based computer-aided drug discovery efforts combined with the drawbacks of structure-based design and the significant induced fit flexibility observed for Hsp90α[22] prompted us to explore the possibility of developing ligand-based three-dimensional (3D) pharmacophore(s) integrated within a self-consistent quantitative structure−activity relationship (QSAR) model. This approach avoids the pitfalls of structure-based techniques; furthermore, the pharmacophore model(s) can be used as 3D search queries to discover new Hsp90α inhibitory scaffolds. We previously reported the use of this innovative approach toward the discovery of new inhibitory leads against glycogen synthase kinase-3β,[23] bacterial MurF,[24] protein tyrosine phosphatase,[25] DPP IV,[26] hormone-sensitive lipase,[27] β-secretase,[28] influenza neuraminidase,[29] and cholesteryl ester transfer protein.[30]

We employed the HYPOGEN module from the CATALYST software package[31] to construct plausible binding hypotheses for a diverse list of Hsp90α inhibitors.[32−34] Subsequently, genetic function algorithm (GFA) and multiple linear regression (MLR) analyses were employed to search for an optimal QSAR that combines high-quality binding

* Corresponding author. Telephone: 00962 6 5355000, ext. 23305. Fax: 00962 6 5339649. E-mail: mutasem@ju.edu.jo.

10.1021/ci100222k
Published on Web 09/10/2010

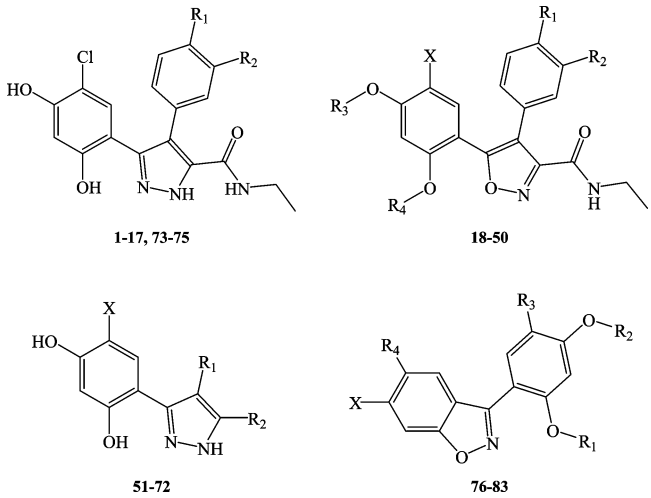


Figure 1. The chemical scaffolds of training compounds. The detailed structures are in Table A, Supporting Information.

pharmacophores with other molecular descriptors and is capable of explaining bioactivity variation across a collection of diverse Hsp90α inhibitors. The optimal pharmacophores were subsequently used as 3D search queries to screen the national cancer institute (NCI) list of compounds and our in-house-built structural database of established drugs and agrochemicals (DAC) for new Hsp90α inhibitory leads.

CATALYST models drug−receptor interaction using information derived only from the ligand structure. HYPOGEN identifies a 3D array of a maximum of five chemical features common to active training molecules, which provides a relative alignment for each input molecule consistent with their binding to a proposed common receptor site. The chemical features considered can be hydrogen bond donors and acceptors (HBDs and HBAs), aliphatic and aromatic hydrophobes (Hbic), positive and negative ionizable (PosIon and NegIon) groups, and aromatic planes (RingArom). The conformational flexibility of training ligands is modeled by creating multiple conformers, judiciously prepared to emphasize representative coverage over a specified energy range. CATALYST pharmacophores have been used as 3D queries for database searching and in 3D-QSAR studies.[23−30,35,36]

2. RESULTS AND DISCUSSION

CATALYST enables automatic pharmacophore construction by using a collection of molecules with activities ranging over a number of orders of magnitude. CATALYST pharmacophores (hypotheses) explain the variability of bioactivity with respect to the geometric localization of the chemical features present in the molecules used to build them. The pharmacophore model consists of a collection of features necessary for the biological activity of the ligands arranged in 3D space (hydrogen bond acceptors and donors, hydrophobic regions, etc.).

Different hypotheses were generated for a series of Hsp90α inhibitors. A total of 83 compounds were used in this study (Figure 1 and Table A in Supporting Information).[32−34] Six training subsets were selected from the collection (Table C in Supporting Information). Each subset consisted of inhibitors of wide structural diversity. The biological activity in the training subsets spanned from 3.5 to 4.0 orders of magnitude.

2.1. Data Mining and Conformational Coverage. The literature was surveyed to collect many reported structurally diverse Hsp90α inhibitors (**1**−**83**, see Table A in Supporting Information and Figure 1).[32−34] The 2D structures of the inhibitors were imported into CATALYST and converted automatically into 3D single conformer representations. The structures were used as starting points for conformational analyses and in the determination of various molecular descriptors for QSAR modeling.

The conformational space of each inhibitor was extensively sampled utilizing the poling algorithm employed within the CONFIRM module of CATALYST.[31] Conformational coverage was performed employing the "Best" module to ensure extensive sampling of conformational space. Pharmacophore generation and pharmacophore-based search procedures are known for their sensitivity to inadequate conformational sampling within the training compounds prompting our extensive conformational sampling.[37]

2.2. Exploration of Hsp90α Pharmacophoric Space. The fact that we have an informative list of 83 Hsp90α inhibitors of evenly spread bioactivities over more than 3.5 orders of magnitude prompted us to employ CATALYST-HYPOGEN to identify possible pharmacophoric binding modes assumed by Hsp90α inhibitors.[32−34]

HYPOGEN implements an optimization algorithm that evaluates a large number of potential binding models for a particular target through fine perturbations to hypotheses that survived the constructive and subtractive phases of the modeling algorithm (see section 4.1.4 and Supporting Information).[38] The extent of the evaluated pharmacophoric space is reflected by the configuration (Config.) cost calculated for each modeling run. It is generally recommended that the Config. cost of any HYPOGEN run not exceed 17 (corresponding to 2^{17} hypotheses to be assessed by CATALYST) to guarantee thorough analysis of all models.[39] The size of the investigated pharmacophoric space is a function of training compounds, selected input chemical features, and other CATALYST control parameters.[39]

Restricting the extent of explored pharmacophoric space should improve the efficiency of optimization via allowing effective evaluation of a limited number of pharmacophoric models. On the other hand, rigorous restrictions imposed on the pharmacophoric space might reduce the possibility of discovering optimal pharmacophoric hypotheses, as they might occur outside the "boundaries" of the pharmacophoric space.

Therefore, we decided to explore the pharmacophoric space of Hsp90α inhibitors under reasonably imposed "boundaries" through 24 HYPOGEN automatic runs and employing six carefully selected training subsets: subsets **I**−**VI** in Table B of the Supporting Information. The training compounds in these subsets were selected in such a way to guarantee maximal 3D diversity and continuous bioactivity spread over more than 3.5 logarithmic cycles. We gave special emphasis to the 3D diversity of the most active compounds in each training subset (Table B, Supporting Information) because of their significant influence on the extent of the evaluated pharmacophoric space during the constructive phase of HYPOGEN algorithm (see section 4.1.4 and Supporting Information).[38]

Guided by our rationally restricted pharmacophoric exploration concept, we restricted the software to explore

Table 1. Success Criteria of Representative Pharmacophoric Hypotheses (Cluster Centers)[a]

RUN[b]	hypotheses[c]	features	Config. cost	total cost	null hypothesis	residual cost[d]	R^e	global R^f	F-stat[g]	Cat. Scramble (%)
1	**7**[h]	**2x HBD, Hbic, RingArom**	**18.88**	**112.3**	**163.0**	**50.7**	**0.950**	**0.690**	**73.61**	**95**
	9	HBA, HBD, 2xHbic	18.88	112.5	163.0	50.5	0.945	0.653	60.12	95
3	8	HBD, 2xHbic, PosIon, RingArom	16.74	81.91	113.12	31.20	0.967	0.674	67.59	95
4	10	HBD, 2xHbic, PosIon, RingArom	15.66	82.38	113.12	30.74	0.957	0.734	94.74	95
5	8	3xHBD, Hbic	19.46	98.39	130.5	32.11	0.917	0.506	27.85	90
	10	2xHBD, Hbic, PosIon	19.46	98.57	130.5	31.93	0.918	0.750	104.1	90
7	9	HBA, HBD, 2xHbic, PosIon	17.44	97.62	130.5	32.88	0.905	0.748	103.2	90
8	**8**	**HBD, 2xHbic, PosIon, RingArom**	**16.35**	**96.38**	**130.5**	**34.12**	**0.904**	**0.756**	**107.7**	**95**
9	**1**	**2xHBD, Hbic, RingArom, 2xExVol**	**17.26**	**87.05**	**103.5**	**16.45**	**0.938**	**0.403**	**15.68**	**90**
	7	HBA, HBD, Hbic, RingArom, 7xExVol	17.26	89.05	103.5	14.45	0.911	0.430	18.40	85
12	1	HBA, HBD, 2xHbic, PosIon, 2xExVol	10.57	87.58	103.5	15.92	0.84	0.670	66.14	95
	2	2xHBD, 2xHbic, PosIon	16.63	87.68	103.5	15.82	0.85	0.589	42.97	95
13	4	HBA, HBD, Hbic, PosIon	21.30	80.32	128.9	48.58	0.974	0.576	40.20	95
	5	HBA, 2xHBD, Hbic	21.30	80.41	128.9	48.49	0.974	0.682	70.45	95
14	4	HBD, PosIon, 2xRingArom	18.30	81.81	111.0	29.19	0.952	0.575	40.05	95
	8	HBD, PosIon, 2xRingArom	18.30	83.00	111.0	28.00	0.939	0.419	17.29	95
16	1	HBA, 3xHbic, RingArom	17.79	79.46	111.0	31.54	0.954	0.737	96.57	95
17	7	HBA, HBD, 2xHbic, PosIon	19.29	85.41	109.8	24.39	0.933	0.498	26.75	95
18	7	HBD, 2xHbic, PosIon, RingArom	18.34	84.17	109.8	25.63	0.927	0.468	22.75	90
20	7	HBA, HBD, 2xHbic, PosIon	16.32	82.84	109.8	26.96	0.921	0.6897	73.48	95
	10	HBA, 2xHbic, PosIon, RingArom	16.31	83.01	109.8	26.79	0.919	0.666	64.79	95
21	9	HBA, HBD, Hbic, RingArom	18.58	123.9	162.3	38.4	0.902	0.725	89.69	90
23	1	2xHBD, 3xHbic	16.61	117.1	147.2	30.1	0.896	0.722	88.07	90
	3	HBA, HBD, 2xHbic, PosIon	16.61	123.4	147.2	23.8	0.848	0.742	99.65	90
	6	HBD, 2xHbic, PosIon, RingArom	16.61	123.6	147.2	23.6	0.846	0.707	81.02	90

[a] Bolded pharmacophores appeared in the best QSAR equations. [b] Correspond to runs in Table C of Supporting Information. [c] High ranking representative hypotheses (in their corresponding clusters, see section 4.1.6). [d] Difference between total cost and the cost of the corresponding null hypotheses. [e] Correlation coefficients between pharmacophore-based bioactivity estimates and bioactivities of corresponding training compound (subsets in Table B of Supporting Information). [f] Correlation coefficients between pharmacophore-based bioactivity estimates and bioactivities of all collected compounds. [g] Fisher statistic calculated based on the linear regression between the fit values of all collected inhibitors (**1**−**83**, Table A of Supporting Information) against pharmacophore hypothesis (employing the "best fit" option and eq C, Supporting Information) and their respective anti-HSP90 bioactivities ($\log(1/IC_{50})$ values). [h] Ranking of hypotheses is as generated by CATALYST in each automatic run.

pharmacophoric models incorporating from zero to one PosIon feature and from zero to three HBA, Hbic, and RingArom features instead of the default range of zero to five, as shown in Table C of the Supporting Information. Furthermore, we instructed HYPOGEN to explore only four- and five-featured pharmacophores, that is, ignore models of lesser number of features, in order to further narrow the investigated pharmacophoric space and to represent the feature-rich nature of known Hsp90α ligands (as shown in Table C of the Supporting Information).

In each run, the resulting binding hypotheses were automatically ranked according to their corresponding "total cost" value, which is defined as the sum of error cost, weight cost, and configuration cost (see section 4.1.5).[31,36,38−41] Error cost provides the highest contribution to total cost and it is directly related to the capacity of the particular pharmacophore as 3D-QSAR model, that is, in correlating the molecular structures to the corresponding biological responses.[31,38−41] HYPOGEN also calculates the cost of the null hypothesis, which presumes that there is no relationship in the data and that experimental activities are normally distributed about their mean. Accordingly, the greater the difference from the null hypothesis cost (residual cost, Table 1) the more likely that the hypothesis does not reflect a chance correlation. An additional validation technique based on Fischer's randomization test[42] was recently introduced into CATALYST: Cat. Scramble.[31] In this test, the biological data and the corresponding structures are scrambled several times, and the software is challenged to generate pharmacophoric models from the randomized data. The confidence in the parent hypotheses (i.e., generated from unscrambled data) is lowered proportional to the number of times the software succeeds in generating binding hypotheses from scrambled data of apparently better cost criteria than the parent hypotheses (see section 4.1.5).[31,36,38−41]

Eventually, 240 pharmacophore models emerged from 24 automatic HYPOGEN runs, out of which only 221 models illustrated Cat. Scramble confidence levels ≥85%. These successful models were clustered and the best representatives (25 models, Table C of the Supporting Information, see section 4.1.6) were used in subsequent QSAR modeling. Clearly from Table 1, the representative models shared comparable features and acceptable statistical success criteria. Emergence of several statistically comparable pharmacophore models suggests the ability of Hsp90α ligands to assume multiple pharmacophoric binding modes within the binding pocket. Therefore, it is quite challenging to select any particular pharmacophore hypothesis as a sole representative of the binding process.

2.3. QSAR Modeling. Despite pharmacophoric hypotheses providing excellent insights into ligand−macromolecule recognition and being useful to mine for new biologically interesting scaffolds, their predictive value as 3D-QSAR models is limited by steric shielding and bioactivity-enhancing or -reducing auxiliary groups.[23−30] This point combined with the fact that pharmacophore modeling of Hsp90α inhibitors furnished several binding hypotheses of comparable success criteria prompted us to employ classical QSAR analysis to search for the best combination of pharmacophore(s) and other 2D descriptors capable of explaining bioactivity variation across the whole list of collected inhibitors (**1**−**83**, Table A, Supporting Information,

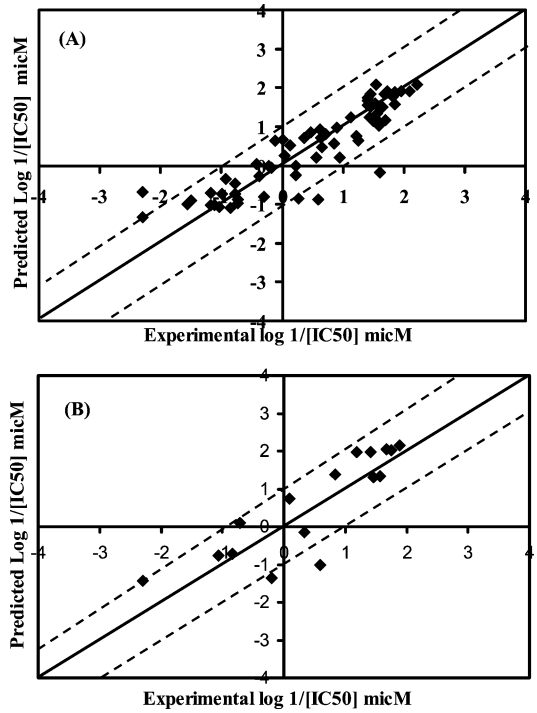


Figure 2. Experimental versus (A) fitted (67 compounds, $r_{LOO}^2 = 0.748$) and (B) predicted (16 compounds, $r_{PRESS}^2 = 0.619$) bioactivities calculated from the best QSAR model, eq 1. The solid lines are the regression lines for the fitted and predicted bioactivities of training and test compounds, respectively, whereas the dotted lines indicate 1.0 log point error margins.

and Figure 1). We employed genetic function approximation and multiple linear regression QSAR (GFA−MLR-QSAR) analysis to search for an optimal QSAR equation(s).

The fit values obtained by mapping representative hypotheses (25 models) against collected Hsp90α inhibitors (**1**−**83**, Table A, Supporting Information, and Figure 1) were enrolled, together with around 100 other physicochemical descriptors, as independent variables (genes) in GFA−MLR-QSAR analysis (see section 4.1.7).[23−30,43,44] However, since it is essential to access the predictive power of the resulting QSAR models on an external set of inhibitors, we randomly selected 16 molecules (marked with double asterisks in Table A, Supporting Information, see section 4.1.7) and employed them as external test molecules for validating the QSAR models (r_{PRESS}^2). Moreover, all QSAR models were cross-validated automatically using the leave-one-out cross-validation in CERIUS2.[43,44]

Equation 1 shows the details of the optimal QSAR model. Figure 2 shows the corresponding scatter plots of experimental versus estimated bioactivities for the training and testing inhibitors.

$$\log(I/IC_{50}) = -1.47 + 2.24 \times 10^{-2}(\text{Hypo1/7})^2 + 1.215 \times 10^{-2}(\text{Hypo8/8})^2 + 2.7 \times 10^{-2}(\text{Hypo9/1})^2 - 9.0 \times 10^{-4}(\text{SsOH})^2 - 0.5446(\text{AtypeN73})^2 + 8.3 \times 10^{-2}(\text{Jurs-RPCS})^2 \quad r_{67}^2 = 0.811, \quad \text{F-statistic} = 42.8 \quad r_{BS}^2 = 0.811, \quad r_{LOO}^2 = 0.748, \quad r_{PRESS(16)}^2 = 0.619 \quad (1)$$

where r_{67}^2 is the correlation coefficient against 67 training compounds, r_{LOO}^2 is the leave-one-out correlation coefficient,

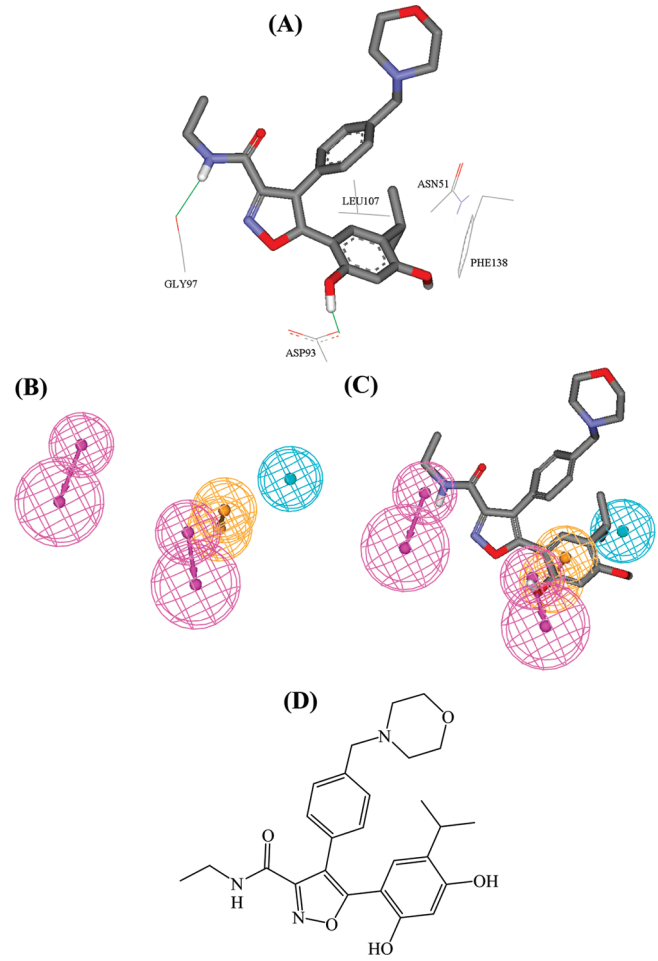


Figure 3. (A) Co-crystal structure of training compound **36** (IC_{50} = 35.0 nM, Table 1) complexed within Hsp90α (PDB code 2VCI, resolution 2.00 Å´). (B) Pharmacophoric features of Hypo1/7: HBD as pink vectored spheres, Hbic as blue spheres, RingArom as vectored orange spheres. (C) Hypo1/7 fitted against **36** employing rigid mapping (i.e., without allowing conformational flexibility). (D) Chemical structure of **36**.

r_{BS}^2 is the bootstrapping regression coefficient and r_{PRESS}^2 is the predictive r^2 determined for the 16 test compounds.[43,44] Hypo1/7, Hypo8/8, and Hypo9/1 represent the fit values of the training compounds against these three pharmacophores (bolded models in Table 1 and Figures 3−5) as calculated from eq C, Supporting Information. Jurs-RPCS describes the relative positive charge surface area (calculated by dividing the solvent-accessible surface area of the most positive atom over the relative positive charge of that atom),[43] AtypeN73 is an atom-type-based AlogP descriptor that encodes for the presence of aromatic NH (e.g., in pyrazole rings) and SsOH is an electrotopological state descriptor for hydroxyl groups[43] (See Table D in Supporting Information).

The emergence of three orthogonal pharmacophoric models, that is, Hypo1/7, Hypo8/8, and Hypo9/1, of cross-correlation $r^2 \leq 0.41$ (Table 2) in eq 1 suggests they represent three complementary binding modes accessible to ligands within the binding pocket of Hsp90α, which means that one of the pharmacophores can optimally explain the bioactivities of some training inhibitors, while the others explain the remaining inhibitors. Similar conclusions were reached about the binding pockets of other targets based on QSAR analysis.[23−30]

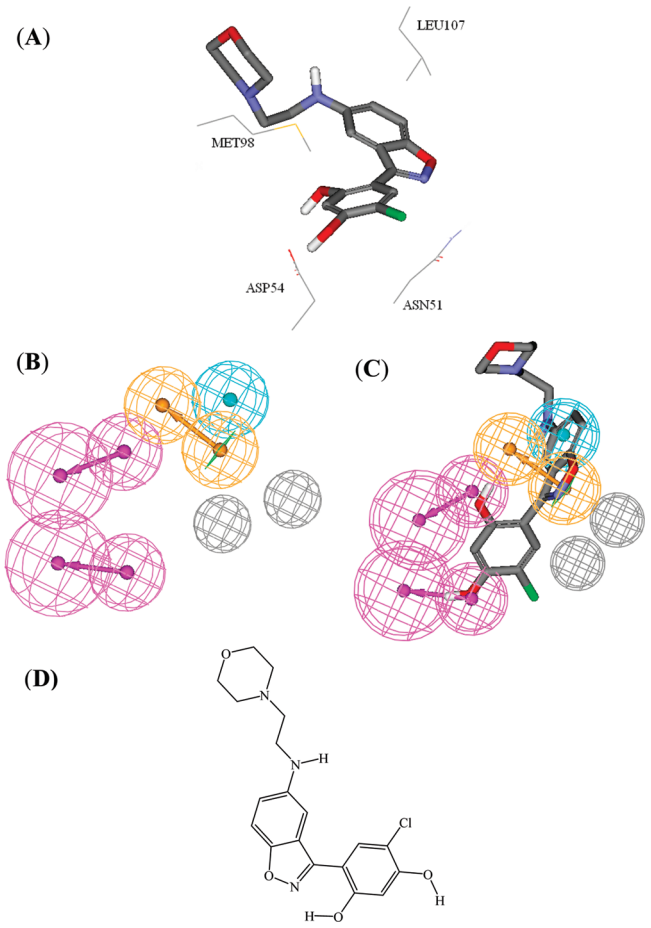


Figure 4. (A) Docked structure of training compound **83** (IC_{50} = 30.0 nM, Table 1) into Hsp90α binding pocket (PDB code 1YET, resolution 1.9 Å´). (B) Pharmacophoric features of Hypo9/1: HBD as pink vectored spheres, Hbic as blue spheres, RingArom as vectored orange spheres. (C) Hypo9/1 fitted against **30**. (D) Chemical structure of **30**.

Figures 3−5 show Hypo1/7, Hypo8/8, and Hypo9/1 and how they map **36** (IC_{50} = 35 nM), **83** (IC_{50} = 30 nM), and **4** (IC_{50} = 27 nM), respectively. The *X*, *Y*, and *Z* coordinates of the three pharmacophores are illustrated in Table 3.

The quadratic format of the pharmacophore models in eq 1 suggests that ligand/Hsp90α affinity is more sensitive to fitting the pharmacophoric models at higher fit values compared with lower values, that is, minor misalignments among attracting moieties within the complex drastically reduces ligand−Hsp90α affinities.

Emergence of the electrotopological state index SsOH in conjunction with negative regression coefficient in eq 1 suggest that presence of hydroxyl groups negatively influences affinity, probably by promoting compound hydration, that is, by competing with binding. A similar explanation holds for the emergence of AtypeN73 in association with negative slope in eq 1. Interestingly, Jurs-RPCS emerged in eq 1 in association with significant positive contribution suggesting that molecules of diffuse positive charges over wider surface areas (i.e., larger atoms or fragments) illustrate better anti-Hsp90α bioactivities compared with molecules of focused positive charges on smaller-sized fragments. This behavior is probably because of the presence of three negatively charged carboxylate moieties in the binding pocket, namely, Asp54, Asp93, and Asp102, which favor positively charged ligands albeit without detrimental interferences from hydration associated with localized charges on smaller fragments.

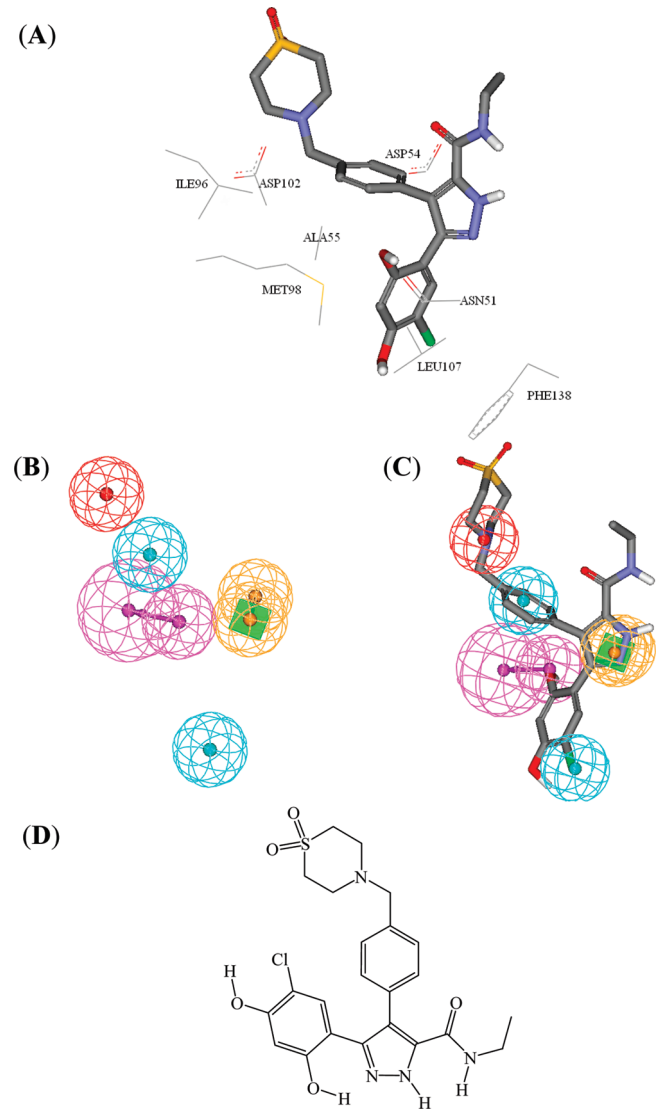


Figure 5. (A) Docked structure of training compound **4** (IC_{50} = 27.0 nM, Table 1) into Hsp90 (PDB code 1YET, resolution 1.9 Å´). (B) Pharmacophoric features of Hypo8/8: HBD as pink vectored spheres, Hbic as blue spheres, RingArom as vectored orange spheres, PosIon as red spheres. (C) Hypo8/8 fitted against **4**. (D) Chemical structure of **4**.

Table 2. The Cross Correlation between the Successful Hypotheses in Equation 1

	Hypo1/7	Hypo8/8	Hypo9/1
Hypo1/7	1		
Hypo8/8	0.409	1	
Hypo9/1	0.273	0.0508	1

2.4. Addition of Exclusion Volumes. Although ligand-based pharmacophores serve as excellent tools to probe ligand/macromolecule recognition and can serve as useful 3D QSAR models and 3D search queries, they suffer from a major drawback: they lack steric constraints necessary to define the size of the binding pocket. This liability renders pharmacophoric models rather promiscuous in some cases.[29] Therefore, we decided to complement the optimal pharmacophores with exclusion spheres employing the Hip-Hop-Refine module implemented within CATALYST.[31] Excluded

Table 3. Pharmacophoric Features and Corresponding Weights, Tolerances and 3D Coordinates of Hypo1/7, Hypo8/8, and Hypo9/1

model	definition		chemical features						
			HBD		HBD		RingArom		Hbic
Hypo1/7[a]	weights		2.32		2.32		2.32		2.32
	tolerances		1.60	2.20	1.60	2.20	1.60	1.60	1.60
	coordinates	*X*	8.776	10.955	5.975	6.665	3.946	4.752	0.519
		Y	2.619	0.570	−3.595	−6.229	−3.313	−3.649	−3.106
		Z	1.279	1.045	1.314	0.055	2.170	5.040	4.402

model	definition		chemical features						
			HBD		RingArom		Hbic	Hbic	PosIon
Hypo8/8[b]	weights		2.06		2.06		2.06	2.06	2.06
	tolerances		1.60	2.20	1.60	1.60	1.60	1.60	1.60
	coordinates	*X*	−0.134	−2.754	3.001	0.550	5.20	−1.186	−2.576
		Y	−0.661	0.136	−1.852	−3.505	−0.574	1.225	3.451
		Z	4.195	5.421	1.895	2.405	7.114	1.461	−0.699

			chemical features								
			HBD		HBD		RingArom		Hbic	ExVol	ExVol
Hypo9/1[c]	weights		1.97		1.97		1.97		1.97	1.97	1.97
	tolerances		1.60	2.20	1.60	2.20	1.60	1.60	1.60	1.20	1.20
	coordinates	*X*	−5.378	−6.332	−0.923	−2.90	0.933	1.525	2.860	−1.646	0.26
		Y	0.646	3.490	2.468	3.889	−1.142	1.712	0.060	−3.202	−3.469
		Z	−0.027	−0.003	1.735	3.488	1.490	2.199	−0.340	3.103	−2.678

[a] Hypo1/7 is the seventh pharmacophore hypothesis generated in the first HYPOGEN run (Table 1). [b] Hypo8/8 is the eighth pharmacophore hypothesis generated in the eighth HYPOGEN run (Table 1). [c] Hypo9/1 is the first pharmacophore hypothesis generated in the ninth HYPOGEN run (Table 1).

volumes resemble sterically inaccessible regions within the binding site (see section 4.1.8).[45]

Two structurally diverse training subsets were selected for Hip-Hop-Refine modeling (Table E in Supporting Information): subset **VII** for Hypo1/7 and Hypo9/1, and subset **VIII** for Hypo8/8. The training compounds were selected in such away that the bioactivities of weakly active compounds are explainable by steric clashes within the binding pocket. Figure 6B,D,F shows sterically refined versions of Hypo1/7, Hypo8/8, and Hypo9/1, respectively.

2.5. Receiver Operating Characteristic (ROC) Curve Analysis. To further validate the resulting models (both QSAR and pharmacophores), we subjected our QSAR-selected pharmacophores to receiver operating curve (ROC) analysis. In ROC analysis, the ability of a particular pharmacophore model to correctly classify a list of compounds as actives or inactives is indicated by the area under the curve (AUC) of the corresponding ROC, as well as other parameters, namely, overall accuracy, overall specificity, overall true positive rate, and overall false negative rate (see section 4.1.9 for more details).[46−49]

Table 4 and Figure 7 show the ROC results of our QSAR-selected pharmacophores. Hypo1/7 and Hypo9/1 illustrated good overall performances with AUC values of 82.5% and 82.3%, respectively. On the other hand, Hypo8/8 exhibited excellent performance with AUC value of 99.5%. This is not unexpected, because the presence of positive ionizable and HBD features in Hypo8/8 enhance its selectivity compared with Hypo1/7 and Hypo9/1.

Clearly, the performances of sterically refined versions of Hypo1/7 and Hypo9/1 improved significantly as reflected by their ROC−AUC values, which shifted from 82.50% to 94.10% and from 82.3% to 93.8%, respectively. Unsurprisingly, the performance of the refined Hypo8/8 is comparable to the unrefined counterpart (ROC−AUC shifted from 99.50% to 100%).

2.6. Comparison of Pharmacophore Model with the Active Site of Hsp90α. To further emphasize the validity of our pharmacophore/QSAR modeling approach, we compared the crystallographic structure of **36** complexed to Hsp90α (IC_{50} = 35.0 nM, PDB code 2VCI)[32] with Hypo1/7. Figure 3 shows the chemical structure of the ligand and compares its Hsp90α complex with the way it maps to Hypo1/7 employing rigid mapping, that is, fitting the ligand's bound state against the pharmacophore without conformational adjustments. Mapping the amidic NH and resorcinolic OH of **36** against two HBDs in Hypo1/7 (Figure 3C) corresponds to hydrogen-bonding interactions tying the NH and OH with the peptidic carbonyl of Gly97 and carboxylate side chain of Asp93 (Figure 3A), respectively, in the crystallographic complex. Similarly, in the crystallographic structure, the isopropyl moiety of **36** is positioned in a hydrophobic pocket comprised of the side chains of Leu107 and Phe138 (figure 3A) corresponding to mapping this group against Hbic feature in Hypo1/7 (Figure 3C). Finally, mapping the resorcinolic aromatic ring of **36** against RingArom feature in Hypo1/7 (Figure 3C) seems to correspond with stacking against the amidic side chain of Asn51 (Figure 3A) in the complex.

Furthermore, we compared the pharmacophoric features of Hypo8/8 and Hypo9/1 and how they map **83** and **4**, respectively, with optimal docked poses of the two compounds. The docking experiments were conducted employing LigandFit docking engine and through default docking settings.[60] Figures 4 and 5 show the pharmacophores, docked poses, and corresponding mapped conformers.

Mapping the two resorcinolic hydroxyls of **83** (IC_{50} = 30.0 nM, Table A in Supporting Information) with two HBDs in

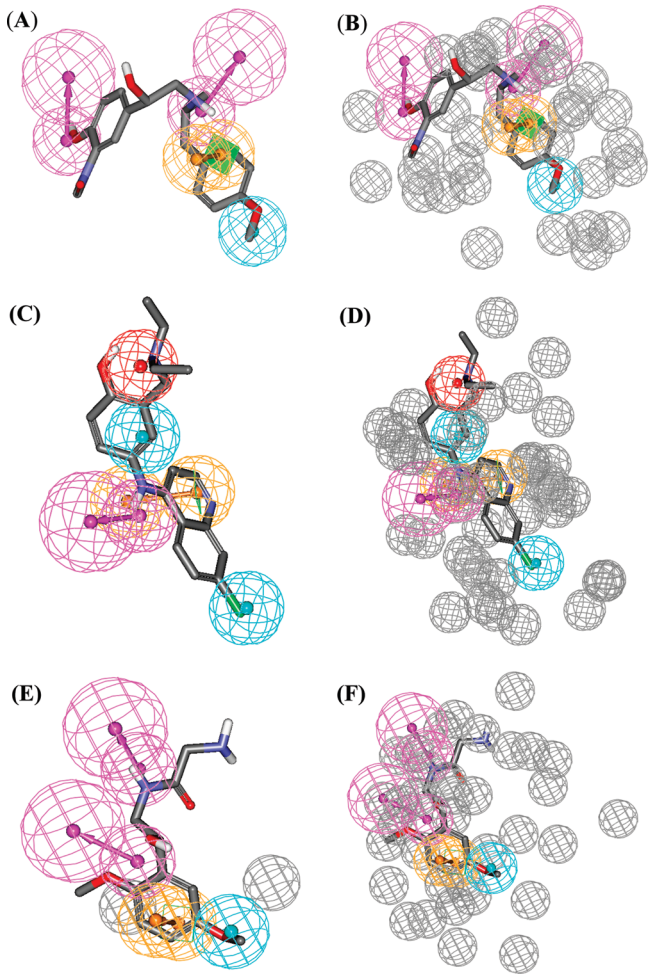


Figure 6. (A, B) Hypo1/7 fitted against hit **84** (IC_{50} = 3.3 nM, Table 1) without and with steric refinement, respectively, (C, D) Hypo8/8 mapped against **85** (IC_{50} = 5.0 nM) without and with steric refinement, respectively, and (E, F) Hypo9/1 mapped against **87** (IC_{50} = 20 nM) without and with steric refinement.

Hypo9/1 (Figure 4B,C) agrees with hydrogen-bonding interactions connecting both hydroxyls with the carboxylate side chain of Asp54 in docked pose (Figure 4A). On the other hand, positioning the benzene ring of the benzisoxazole moiety of **83** between the hydrophobic side chains of Leu107 and Met98 in the docked pose (Figure 4A) agrees with mapping this moiety against the Hbic feature in Hypo9/1 (Figure 4C). Similarly, mapping the isoxazole fragment of **83** against the RingArom feature in Hypo9/1 (Figure 4C) correlates with stacking this group against the amidic side chain of Asn51 in the docked pose (Figure 4A).

Similar analogies can be drawn between the docked conformer of **4** (IC_{50} = 27.0 nM, Table A in Supporting Information) and the way it maps to Hypo8/8, as shown in Figure 5. Mapping the amino functionality of **4** against the PosIon feature in Hypo8/8 (Figure 5C) agrees with an electrostatic attraction connecting this group and the carboxylate of Asp102 in the docked pose (Figure 5A). Likewise, positioning the resorcinolic chloro substituent of **4** within a hydrophobic pocket comprised of the side chains of Leu107 and Phe138 in the docked pose (Figure 5A) agrees with mapping this moiety against the Hbic feature in Hypo8/8 (Figure 5C). Similarly, mapping the central phenyl linker of **4** against the Hbic feature in Hypo8/8 (Figure 5C) correlates nicely with fitting this group into a hydrophobic pocket consisting of side chains of Ala55, Ile96, and Met98 (Figure 5A). Equally, mapping the resorcinolic OH of **4** against a HBD in Hypo8/8 (Figure 5C) points to a hydrogen-bonding interaction connecting this group with the peptidic carbonyl of Asn51 in the docked pose (Figure 5A). Finally, mapping the pyrazole ring of **4** against a RingArom feature in Hypo8/8 (Figure 5C) agrees with stacking this electron-deficient ring against the carboxylate side chain of Asp54 in the docked pose (Figure 5A).

Clearly from the above discussion, Hypo1/7, Hypo8/8, and Hypo9/1 represent three valid binding modes assumed by ligands within Hsp90α. Furthermore, these models point to a limited number of critical interactions required for high ligand−Hsp90α affinity in each of the binding modes. In contrast, crystallographic or docked complexes reveal many bonding interactions without highlighting critical ones. Incidentally, Figures 3A−5A only show interactions corresponding to pharmacophoric features, while other binding interactions were hidden for clarity.

2.7. *In Silico* Screening and Subsequent *in Vitro* Evaluation. Sterically refined Hypo1/7, Hypo8/8, and Hypo9/1 were employed as 3D search queries against the NCI (238 819 structures) and DAC (3002 compounds) libraries using the "Best Flexible Database Search" option implemented within CATALYST. Compounds that have their chemical groups spatially overlap (map) with corresponding features of the particular pharmacophoric model were captured (hits). Table 5 summarizes the numbers of captured hits by sterically refined versions of the pharmacophores.

NCI hits were filtered based on Lipinski's and Veber's rules,[50,51] while DAC hits were left unfiltered. We believe loosening Lipinski's and Veber's prefilters in the DAC case allows reasonable structural diversity among hit molecules particularly under the combined restrictions of three orthogonal searching pharmacophores and selection via a QSAR model. Lax in drug-likeness prefilters is not unreasonable because it is highly unexpected that such studies (i.e., *in silico* screening followed by *in vitro* evaluation) can identify drugs or even drug candidates. Nevertheless, enforcing some kind of drug-likeness prefilters should help in finding hits more amenable for subsequent optimization into leads, which explains our adherence to Lipinski's and Veber's rules in the NCI hits case.

Surviving hits were fitted against Hypo1/7, Hypo8/8, and Hypo9/1 (without exclusion volumes), and their fit values, together with other relevant molecular descriptors, were substituted in QSAR eq 1 to predict their anti-Hsp90α IC_{50} values. The highest-ranking available hits, 18 drugs and 82 NCI compounds, were evaluated *in vitro* against human recombinant Hsp90α via an established malachite green-based assay that measures the amount of inorganic phosphate released by the ATPase activity of human Hsp90α.[57,58,64] Although human Hsp90α is known to have low ATPase activity,[62,63] we decided to evaluate our hits against this chaperon because it provides a more realistic and relevant platform for future testing against cancer, that is, compared with yeast Hsp90α. Furthermore, we implemented a reported bioassay procedure that incubates Hsp90α, ATP, and the particular inhibitor over 24 h, which should compensate for the slow ATPase bioactivity of human Hsp90α. To validate our bioassay settings, that is, the use of human Hsp90α and malachite green, we determined the IC_{50} value of standard

Table 4. ROC Curve Analysis Criteria for QSAR-Selected Pharmacophores and Their Sterically Refined Versions

pharmacophore model	ROC[a]−AUC[b]	ACC[c]	SPC[d]	TPR[e]	FNR[f]
Hypo1/7	0.825	0.9260	0.9324	0.857	0.068
Hypo8/8	0.995	0.9269	0.9562	0.556	0.044
Hypo9/1	0.823	0.9268	0.9323	0.857	0.0677
refined Hypo1/7	0.941	0.9260	0.9374	0.794	0.063
refined Hypo8/8	1.00	0.9269	0.9575	0.540	0.043
refined Hypo9/1	0.938	0.9269	0.9411	0.746	0.0588

[a] Receiver operating characteristic curve. [b] Area under the curve. [c] Overall accuracy. [d] Overall specificity. [e] Overall true positive rate. [f] Overall false negative rate.

inhibitor geldanamycin under the same conditions. Our conditions determined the IC_{50} value of geldanamycin to be 273 nM, which is within the reported literature range of 282 ± 9.7 nM (determined by fluorescence polarization).[61]

Initially, hits were screened at 10 μM concentrations; subsequently, compounds of anti-Hsp90α inhibitory percentages ≥60% at 10 μM were further assessed to determine their IC_{50} values. Table 6 and Figure 8 show active hits and their corresponding estimated and experimental anti-Hsp90α bioactivities.

Clearly from Table 6 and Figure 8, formoterol (a long-acting β2-adrenergic agonist used in the management of asthma and chronic obstructive pulmonary disease) illustrated the highest anti-Hsp90α activity with IC_{50} value of 3 nM, while the aminoquinolines amodiaquine and primaquine came next with experimental IC_{50} values of 5 and 6 nM, respectively. These results agree with reported observations about unexplained antileukemic actions of formoterol[53] and the Hsp90α inhibitory properties of some miscellaneous aminoquinolines (other than amodiaquine and primaquine).[59] Interestingly, midodrine (an α1-adrenergic agonist), which illustrated an anti-Hsp90α IC_{50} of 20 nM, is not reported to possess any anticancer properties and, therefore, warrants future research in this direction. Figure 6 shows how formoterol, amodiaquine, and midodrine map to Hypo1/7, Hypo8/8, and Hypo9/1, respectively.

Finally, *in vitro* testing showed that 20 NCI high-ranking hits possess nanomolar to micromolar inhibitory profiles against Hsp90α (Table 6, Figure 8), while the rest were found to be weakly active or inactive.

It remains to be mentioned that although QSAR predictions were rather accurate with some hit compounds, for example, **102**, **107**, **108**, **109**, and **120**, they deviated by a maximum of two logarithmic cycles with most hits (Figure 9). We believe these errors are because training compounds used in QSAR and pharmacophore modeling are significantly structurally different from hit molecules (see Table 7 and section 2.8), which limits the extrapolatory potential of the QSAR equation. Furthermore, the fact that we implemented a bioassay method different from that used for training compounds can also explain part of the predicted-to-experimental differences in hit bioactivities. Additionally, QSAR-based predictions have their errors; in fact an r_{PRESS}^2 of 0.619 suggests a certain level of uncertainty in predictions.

2.8. Similarity Analysis between Training Compounds and Active Hits. We employed three library comparison methods implemented in Discovery Studio 2.5 to assess the structural similarity/diversity between the modeled compounds (**1**−**83**, Figure 1 and Table A in Supporting Information) and other reported Hsp90 inhibitors[52,67−69] (library A) compared with active hits (**84**−**121**, library B), namely, Murcko assemblies, Bayesian model, and global fingerprints.

In Murcko assemblies, the algorithm breaks the ligands of each library into unique occurrences of molecular rings, ring assemblies, bridge assemblies, chains, Murcko assemblies, or any combination of these. Murcko assemblies are contiguous ring systems plus chains that link two or more rings.[65] The two libraries are compared using a Tanimoto similarity of the assemblies based on the fragments that are common and unique to each library.[66]

On the other hand, in the Bayesian model approach, two Bayesian models were built, one to learn library A and one to learn library B. Finally, it scores all ligands using both models. A distance is computed as eq 2:

$$\text{Distance} = \text{ScoreAA} + \text{ScoreBB} - \text{ScoreAB} - \text{ScoreBA} \quad (2)$$

where ScoreAA is the average score of library A molecules scored by the Bayesian model that learned library A molecules, while ScoreBB is the average score of library B molecules scored by the Bayesian model that learned library B. ScoreAB and ScoreBA are the average scores of libraries A and B molecules scored by the Bayesian models that learned libraries B and A, respectively. The higher the distance, the more dissimilar the libraries are.[66]

Finally, the global fingerprint comparison algorithm generates a global fingerprint for all ligands in the training list and all ligands in the hits list and then computes a Tanimoto similarity coefficient between the two libraries.[66]

Table 7 shows the results of the three similarity/diversity assessment procedures. Clearly, the three methods suggest minimal structural similarity between known Hsp90 inhibitors and our active hits.

3. CONCLUSIONS

Hsp90α inhibitors are currently considered to be potential treatments for cancer. The pharmacophoric space of Hsp90α inhibitors was explored via six diverse sets of inhibitors and using CATALYST-HYPOGEN to identify high-quality binding model(s). Subsequently, genetic algorithm and multiple linear regression analysis were employed to access an optimal QSAR model capable of explaining anti-Hsp90α bioactivity variation across 83 collected Hsp90α inhibitors. Three orthogonal pharmacophoric models emerged in the QSAR equation suggesting the existence of at least three distinct binding modes accessible to ligands within the Hsp90α binding pocket. The QSAR equation and the associated pharmacophoric models were experimentally validated by the identification of several Hsp90α inhibitors

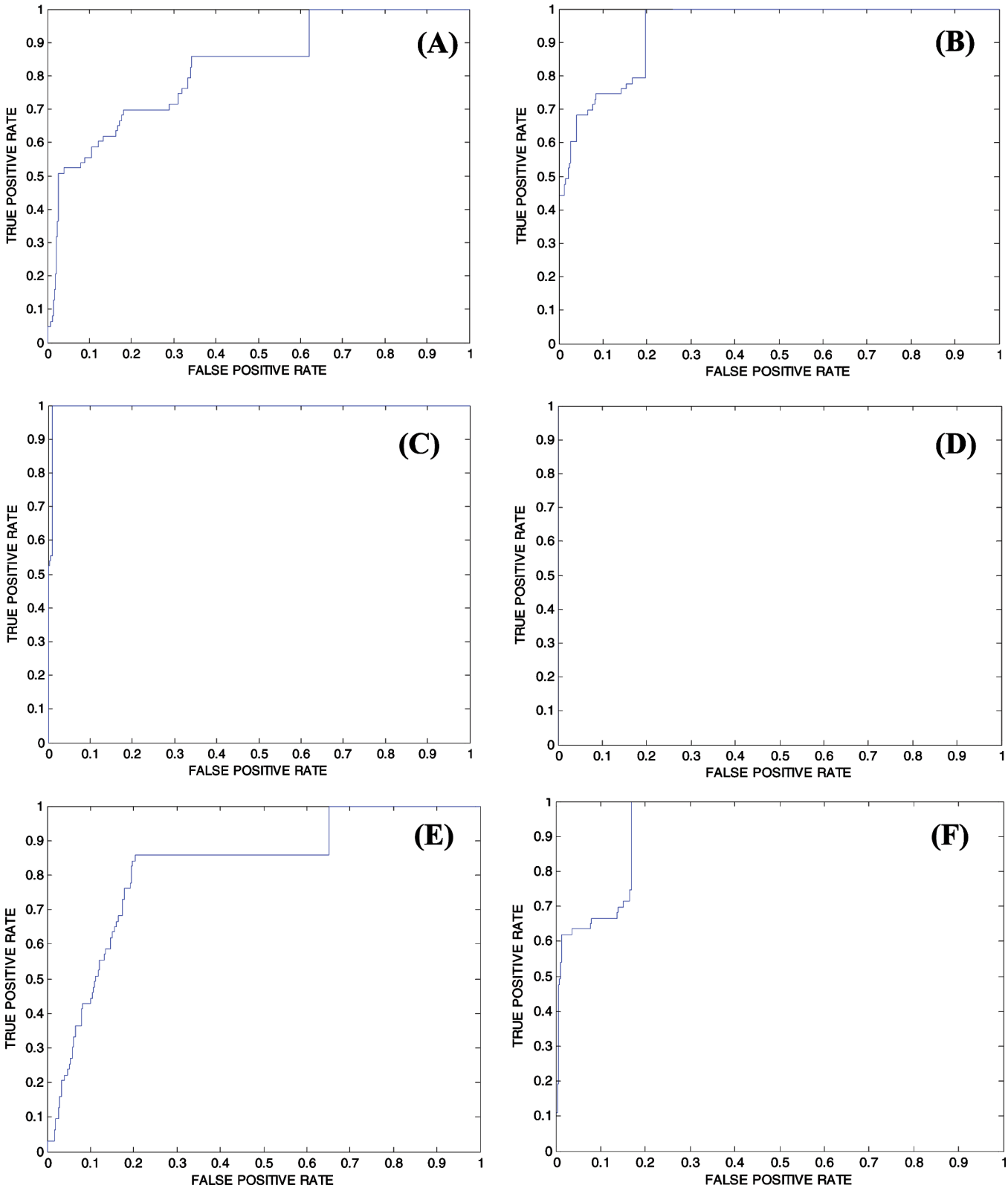


Figure 7. ROC curves of (A) Hypo1/7, (B) sterically refined Hypo1/7, (C) Hypo 8/8, (D) sterically refined Hypo8/8, (E) Hypo9/1, and (F) sterically refined Hypo9/1.

retrieved via *in silico* screening, out of which four established drugs illustrated nanomolar potencies: amodiaquine, primaquine, midodrine, and formoterol. Our results suggest that the combination of pharmacophoric exploration and QSAR analyses can be a useful tool for finding new diverse Hsp90α inhibitors.

4. EXPERIMENTAL SECTION

4.1. Molecular Modeling. *4.1.1. Software and Hardware.* The following software packages were utilized in the present research: CATALYST (Version 4.11), Accelrys Inc. (www.accelrys.com), San Diego, CA; CERIUS2 (Version 4.10), Accelrys Inc. (www.accelrys.com), San Diego, CA; CS

Table 5. Numbers of Captured Hits by Sterically-Refined Versions of Hypo1/7, Hypo8/8, and Hypo9/1

		pharmacophore models		
3D database[a]	post-screening filtering[b]	sterically refined Hypo1/7	sterically refined Hypo8/8	sterically refined Hypo9/1
NCI	before	3392	623	1766
	after	2380	175	1360
DAC[c]		42	9	115

[a] NCI is National Cancer Institute list of available compounds (238 819 structures); DAC is our in-house list of established drugs and agrochemicals (3002 structures). [b] Using Lipinski's and Veber's rules. Two Lipinski's violations were tolerated. [c] No post-screening filtering.

ChemDraw Ultra 6.0, Cambridge Soft Corp. (http://www.cambridgesoft.com), Cambridge, MA; Discovery Studio 2.5, Accelrys Inc. (www.accelrys.com), San Diego, CA.

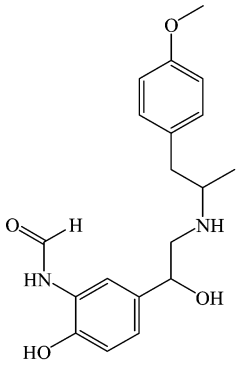
Pharmacophore and QSAR modeling studies were performed using CATALYST (HYPOGEN module) and CERIUS2 software suites from Accelrys Inc. (San Diego, California, www.accelrys.com) installed on a Silicon Graphics Octane2 desktop workstation equipped with a dual 600 MHz MIPS R14000 processor (1.0 GB RAM) running the Irix 6.5 operating system. Structure drawing was performed employing ChemDraw Ultra 6.0, which was installed on a Pentium 4 PC.

4.1.2. Data Set. The structures of 83 Hsp90α inhibitors (**1**−**83**, Table A in Supporting Information, Figure 1) were collected from recently published literature.[32−34] Although the *in vitro* bioactivities of the collected inhibitors were gathered from three separate articles, they were determined employing the same bioassay methodologies. The bioactivities were expressed as the concentrations of the test compounds that inhibited the activity of Hsp90α by 50% (IC_{50}, μM). The logarithm of measured IC_{50} (μM) values was used in the three-dimensional quantitative structure−activity analysis (3D-QSAR), thus correlating the data linearly to the free energy change.

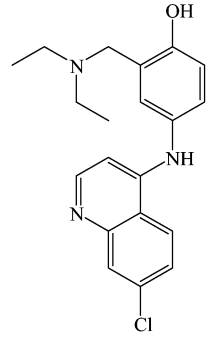
Table 6. Predicted and Experimental Bioactivities of High-Ranking Hit Molecules

		best fit values[c]			predicted	experimental
hits[a]	name[b]	Hypo1/7	Hypo8/8	Hypo9/1	IC_{50} (nM)[d]	IC_{50} (nM)[e]
84	formoterol	7.32	0	5.28	290	3 (0.98)[f,g]
85	amodiaquine	5.31	5.81	4.42	802	5 (0.99)[g]
86	primaquine	7.90	4.51	6.33	42	6 (0.96)[g]
87	midodrine	7.77	0.52	6.26	134	20 (0.99)[g]
88	terfenadine	8.70	0	4.70	351	3467 (0.96)[g]
89	famotidine	4.95	0	5.23	1510	4970 (0.82)[g]
90	isoxsuprine	4.91	0	5.46	578	59%[h]
91	diaveridine	4.85	0	5.53	1010	52%[h]
92	bromodiolone	7.61	0	5.74	507	50%[h]
93	betaxolol	1.70	0	5.22	2330	50%[h]
94	bumbuterol	3.55	0	5.17	3280	49%[h]
95	clobenprobit	8.58	0	5.39	82	49%[h]
96	bicalutamide	4.74	0	5.82	1350	47%[h]
97	nicardipine	0	7.61	3.81	5799	43%[h]
98	chlorhexidine	8.16	0.29	5.94	104	39%[h]
99	chlorbufam	4.62	0	3.46	3120	36%[h]
100	trimethoprim	4.53	0	5.09	1640	26%[h]
101	atenolol	0.54	0	5.32	5420	3%[h]
102	NSC159885	5.74	0	6.76	59	66 (0.99)
103	NSC335504	6.16	0	6.29	2130	87 (0.99)
104	NSC14710	6.14	9.02	7.84	5	100 (0.92)
105	NSC74819	8.81	0	6.45	40	125 (0.92)
106	NSC674003	8.47	0	6.49	61	128 (0.97)
107	NSC88001	4.62	0	7.95	139	137 (0.99)
108	NSC340303	7.29	0	7.42	285	149 (0.95)
109	NSC191952	3.5	0	2.99	548	272 (0.98)
110	NSC139029	7.5	0	2.99	0.2	629 (0.99)
111	NSC56628	8.14	4.63	4.39	7	648 (0.90)
112	NSC56621	6.26	9.08	8.06	10	759 (0.93)
113	NSC657681	8.69	0	4.87	4	781 (0.99)
114	NSC1761	3.99	0.818	2.91	14	783 (0.99)
115	NSC680410	7.55	0	6.2	3	947 (0.99)
116	NSC138756	5.4	0	6.34	4	970 (0.99)
117	NSC380303	6.3	0	6.6	56	983 (0.94)
118	NSC201560	6.67	0	6.05	355	1038 (0.90)
119	NSC56642	8.34	4.91	4.76	7	1170 (0.99)
120	NSC128439	8.27	0	5.64	1165	1410 (0.97)
121	NSC608048	6.39	0	3.99	12	1740 (0.99)
GMC	geldanamycin[i]					273 (0.93)

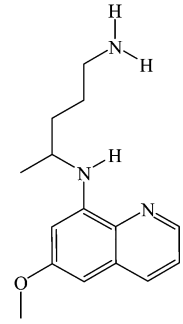
[a] Chemical structures shown in Figure 8. [b] Chemical or generic name or NCI number. [c] Fit values calculated against respective hypotheses using eq C in Supporting Information. [d] Predicted IC_{50} in nM according to QSAR eq 1. [e] Experimental IC_{50} values determined at 10, 1, 0.1, and occasionally at 0.01 and 0.001 μM inhibitor concentrations. [f] Values in brackets represent the correlation coefficients of the corresponding dose−response line. [g] These values represent average results of duplicate measurements. [h] Percent inhibition at 10 μM inhibitor concentration. [i] Reported IC_{50} = 281 nM.[61]



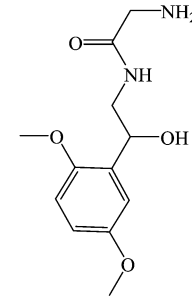
84



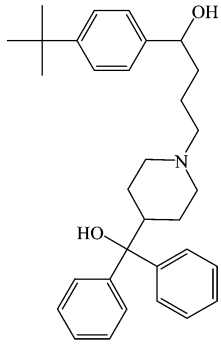
85



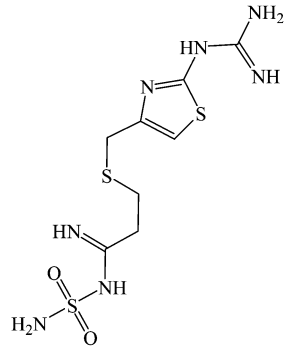
86



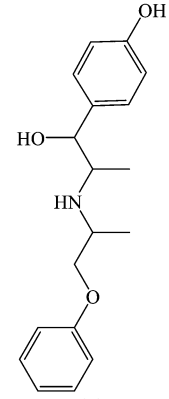
87



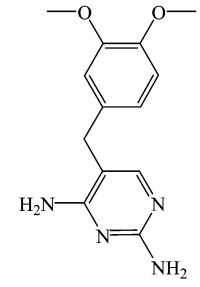
88



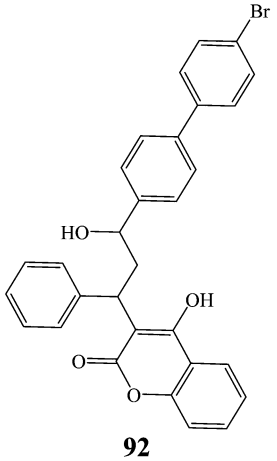
89



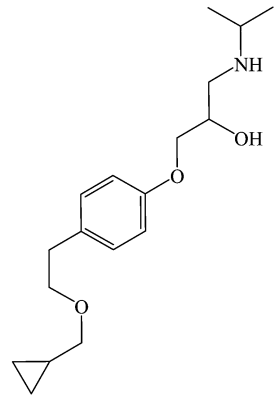
90



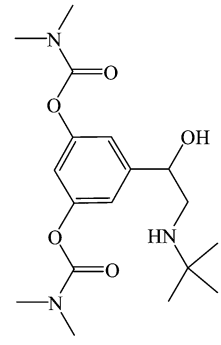
91



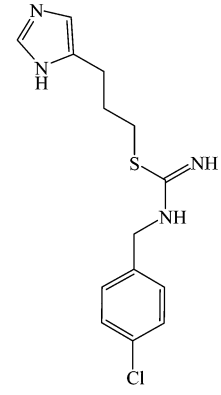
92



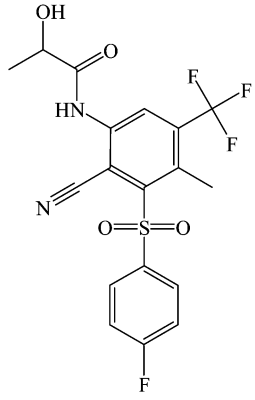
93



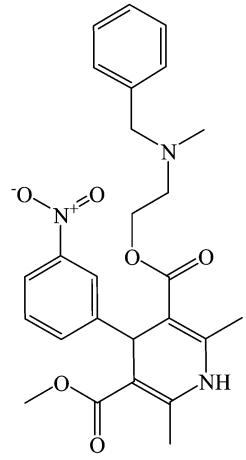
94



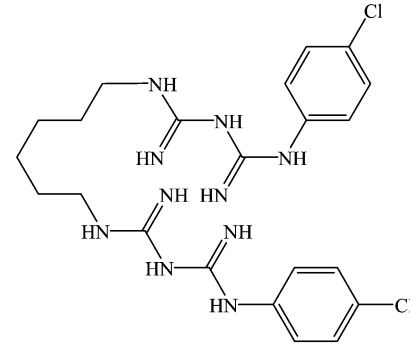
95



96



97



98

Figure 8

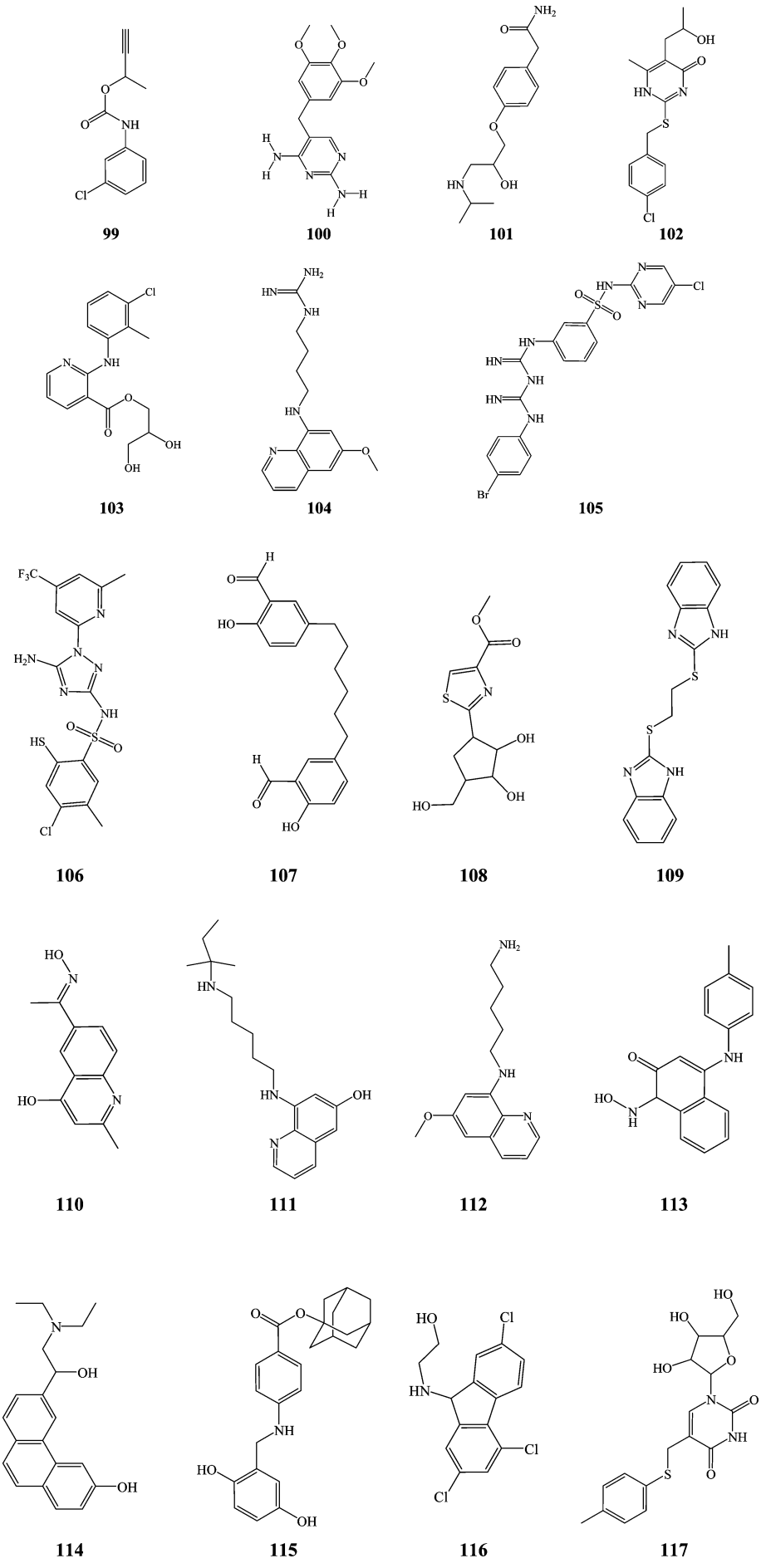
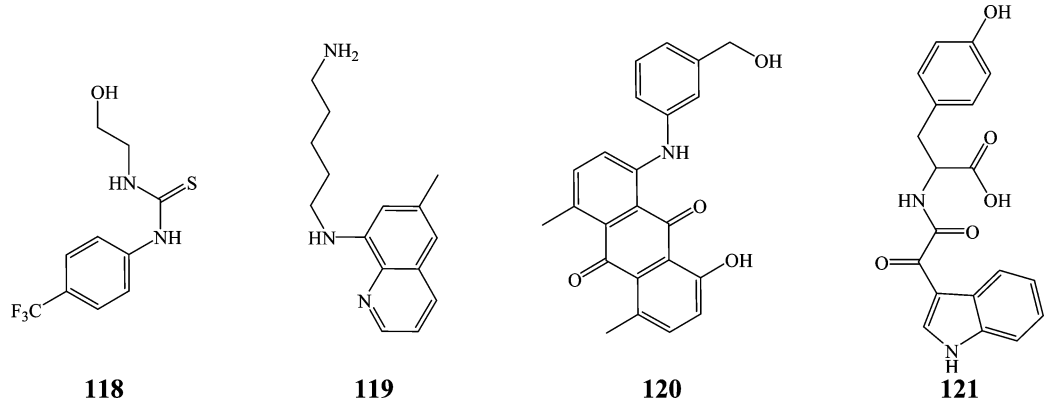

99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117

Figure 8



118 119 120 121

Figure 8. Chemical structures of active hits.

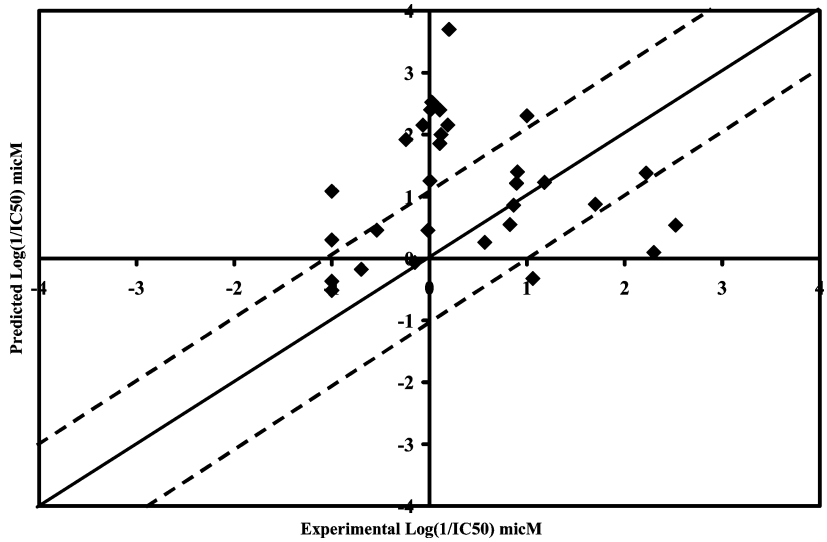


Figure 9. Experimental versus predicted bioactivities of captured hits as calculated from QSAR model eq 1. The solid lines are the regression lines for the fitted and predicted bioactivities of training and test compounds, respectively, whereas the dotted lines indicate 1.0 log point error margins. Only hit compounds having known IC_{50} values were included in this plot. Hits that illustrate Hsp90 inhibition percentage of 50% (±1%) at 10 μM were assumed to have IC_{50} values of 10 μM.

Table 7. Results of Similarity Analysis between Training Compounds and Active Hits

Murcko assemblies[a]		Bayesian model[b]		global fingerprints[b,c]	
number of total assemblies	81	average LibA score of library A ligands	5.23	number of total global fingerprint bits	1729
number of common assemblies	0	average LibB score of library A ligands	−66.14	number of common global fingerprint bits	179
number of assemblies only in library A[d]	53	average LibA score of library B ligands	−26.85	number of global fingerprint bits only in library A	856
number of assemblies only in library B[e]	28	average LibB score of library B ligands	26.58	number of global fingerprint bits only in library B	694
similarity score between the two libraries	0	Bayesian distance between the two libraries	124.8	similarity score between the two libraries	0.104

[a] See section 2.8 and ref 65. [b] See section 2.8 and ref 66. [c] Done by implementing the fingerprint descriptor FCFC_6, which correspond to functional-class extended-connectivity fingerprint count up to diameter 6.[66] [d] Library A list includes all training and testing compounds employed in pharmacophore and QSAR modeling (**1**−**83**, Figure 1 and Table A in Supporting Information) in addition to 84 Hsp90 inhibitors reported in the literature,[67−69] including geldanamycin,[67] 17-allylamino-17-demethoxygeldamycin,[32] and radicicol.[67] The total library contains 167 inhibitors (an sd file of this list is provided as Supporting Information). [e] Library B includes active hits (**84**−**121**, Figure 8 and Table 6).

In a few cases where the IC_{50} values of some compounds were expressed as being higher than 20 μM, we assumed that their IC_{50} values were equal to 200 μM (4 logarithmic cycles away from the most potent inhibitor (i.e., **56**, **79**, and **82**). These assumptions are necessary to allow pharmacophore modeling, statistical correlation, and QSAR analysis. The logarithmic transformation of IC_{50} values should minimize any potential errors resulting from such assumptions.

The two-dimensional (2D) chemical structures of the inhibitors were sketched using ChemDraw Ultra and saved in MDL-molfile format. Subsequently, they were imported into CATALYST, converted into corresponding standard 3D structures, and energy-minimized to the closest local minimum using the molecular mechanics CHARMm force field implemented in CATALYST. The resulting 3D structures were utilized as starting conformers for CATALYST conformational analysis.

4.1.3. Conformational Analysis. The molecular flexibilities of the collected compounds were taken into account by considering each compound as a collection of conformers representing different areas of the conformational space accessible to the molecule within a given energy range. Accordingly, the conformational space of each inhibitor (**1**−**83**, Figure 1 and Table A in Supporting Information) was explored adopting the "best conformer generation" option within CATALYST[31] based on the generalized CHARMm force field implemented in the program. Default parameters were employed in the conformation generation procedure of training compounds and screened libraries (NCI and DAC), that is, a conformational ensemble was generated with an energy threshold of 20 kcal/mol from the local minimized structure at which has the lowest energy level and a maximum limit of 250 conformers per molecule.[31]

4.1.4. Pharmacophoric Hypotheses Generation. All 83 molecules with their associated conformational models were regrouped into a spreadsheet. The biological data of the inhibitors were reported with an "Uncertainty" value of 3, which means that the actual bioactivity of a particular inhibitor is assumed to be situated somewhere in an interval ranging from one-third to three-times the reported bioactivity value of that inhibitor.[39,40] Subsequently, six structurally diverse training subsets, sets **I**, **II**, **III**, **IV**, **V**, and **VI** in Table B of the Supporting Information, were carefully selected from the collection for pharmacophore modeling. Typically, CATALYST requires informative training sets that include at least 16 compounds of evenly spread bioactivities over at least three and a half logarithmic cycles. Lesser training lists could lead to chance correlation and thus faulty models.

The selected training sets were utilized to conduct 24 modeling runs to explore the pharmacophoric space of Hsp90α inhibitors (Table C in Supporting Information). The exploration process included altering the interfeature spacing parameter (100 and 300 pm) and the maximum number of allowed features in the resulting pharmacophore hypotheses, that is, they were allowed to vary from 4 to 5 for first and second runs and from 5 to 5 for third and fourth runs of each training set, as shown in Table C of the Supporting Information.

Pharmacophore modeling employing CATALYST proceeds through three successive phases: the constructive phase, subtractive phase, and optimization phase (see CATALYST Modeling Algorithm in Supporting Information).[31,36,38−41]

4.1.5. Assessment of the Generated Hypotheses. When generating hypotheses, CATALYST attempts to minimize a cost function consisting of three terms: Weight cost, Error cost and Configuration cost (see CATALYST Cost Analysis in Assessment of Generated Binding Hypotheses in Supporting Information).

An additional approach to assess the quality of CATALYST-HYPOGEN pharmacophores is to cross-validate them using the Cat-Scramble program implemented in CATALYST. This validation procedure is based on Fischer's randomization test.[42] In this validation test, we selected a 95% confidence level, which instructs CATALYST to generate 19 random spreadsheets by the Cat-Scramble command. Subsequently, CATALYST-HYPOGEN is challenged to use these random spreadsheets to generate hypotheses using exactly the same features and parameters used in generating the initial unscrambled hypotheses. Success in generating pharmacophores of comparable cost criteria to those produced by the original unscrambled data reduces the confidence in the training compounds and the unscrambled original pharmacophore models.[31,42,54] Based on Fischer randomization criteria; only 221 pharmacophores exceeded the 85% significance threshold for subsequent processing (clustering and QSAR analyses).

4.1.6. Clustering of the Generated Pharmacophore Hypotheses. The successful models (221) were clustered into 25 groups utilizing the hierarchical average linkage method available in CATALYST. Subsequently, the highest-ranking representatives, as judged based on their significance *F*-values, were selected to represent their corresponding clusters in subsequent QSAR modeling. Table 1 shows information about representative pharmacophores including their pharmacophoric features, success criteria, and differences from corresponding null hypotheses. The table also shows the corresponding Cat. Scramble confidence levels for each representative pharmacophore.

4.1.7. QSAR Modeling. A subset of 67 compounds from the total list of inhibitors (**1**−**83**, Table A in Supporting Information and Figure 1) was utilized as a training set for QSAR modeling; the remaining 16 molecules (ca. 20% of the data set) were employed as an external test subset for validating the QSAR models. The test molecules were selected as follows: the 83 inhibitors were ranked according to their IC_{50} values, and then every fifth compound was selected for the test set starting from the high-potency end. This selection considers the fact that the test molecules must represent a range of biological activities similar to that of the training set. The selected test inhibitors are marked with double asterisks in Table A of the Supporting Information.

The logarithm of measured $1/IC_{50}$ (μM) values was used in QSAR, thus correlating the data linearly to the free energy change. The chemical structures of the inhibitors were imported into CERIUS2 as standard 3D single conformer representations in SD format. Subsequently, different descriptor groups were calculated for each compound employing the C2.DESCRIPTOR module of CERIUS2. The calculated descriptors included various simple and valence connectivity indices, electrotopological state indices, and other molecular descriptors (logarithm of partition coefficient, polarizability, dipole moment, molecular volume, molecular weight, molecular surface area, etc.).[43] The training compounds were fitted (using the best-fit option in CATALYST)[31] against the representative pharmacophores (25 models, Table 1), and their fit values were added as additional descriptors. The fit value for any compound is obtained automatically via eq 5.[31]

Genetic function approximation (GFA) was employed to search for the best possible QSAR regression equation capable of correlating the variations in biological activities of the training compounds with variations in the generated descriptors, that is, multiple linear regression modeling (MLR). GFA techniques rely on the evolutionary operations of "crossover and mutation" to select optimal combinations of descriptors (i.e., chromosomes) capable of explaining bioactivity variation among training compounds from a large pool of possible descriptor combinations, that is, chromosome populations. However, to avoid overwhelming GFA−MLR

with a large number of poor descriptor populations, we removed lowest-variance descriptors (20%) prior to QSAR analysis.

Each chromosome is associated with a fitness value that reflects how good it is compared with other solutions. The fitness function employed herein is based on Friedman's "lack-of-fit" (LOF).[43]

Our preliminary diagnostic trials suggested the following optimal GFA parameters: explore linear, quadratic, and spline equations at mating and mutation probabilities of 50%; population size = 500; number of genetic iterations = 30 000; lack-of-fit (LOF) smoothness parameter = 1.0. However, to determine the optimal number of explanatory terms (QSAR descriptors), it was decided to scan and evaluate all possible QSAR models resulting from 4 to 10 explanatory terms.

All QSAR models were validated employing leave-one-out cross-validation (r_{LOO}^2), bootstrapping (r_{BS}^2), and predictive r^2 (r_{PRESS}^2) calculated from the test subsets. The predictive r_{PRESS}^2 is defined as in eq 3:

$$r_{PRESS}^2 = SD - PRESS/SD \quad (3)$$

where SD is the sum of the squared deviations between the biological activities of the test set and the mean activity of the training set molecules and PRESS is the squared deviations between predicted and actual activity values for every molecule in the test set.

4.1.8. Addition of Exclusion Volumes. To account for the steric constraints of the binding pocket, we decided to decorate Hypo1/7, Hypo8/8, and Hypo9/1 with exclusion volumes employing the Hip-Hop-Refine module of CATALYST. Hip-Hop-Refine uses inactive training compounds to construct excluded volumes that resemble the steric constraints of the binding pocket. It identifies spaces occupied by the conformations of inactive compounds and free from active ones. These regions are then filled with excluded volumes.[25−27]

Since each pharmacophore resembles a separate binding mode, it was decided to select two separate training subsets for constructing appropriate exclusion spheres. Subset **VII** was used to construct exclusion spheres around Hypo1/7 and Hypo9/1, while subset **VIII** was used to construct exclusion spheres around Hypo8/8, as in Table E of the Supporting Information.

In Hip-Hop-Refine, the user defines how many molecules must map completely or partially the particular hypothesis via the Principal and Maximum Omitted Features (MaxOmitFeat) parameters. Active compounds are normally assigned a MaxOmitFeat parameter of zero and principal value of 2 to instruct the software to consider all their chemical moieties in pharmacophore modeling and to fit them against all the pharmacophoric features of a particular hypothesis. On the other hand, inactive compounds are allowed to miss one (or two) features by assigning them a MaxOmitFeat of 1 (or 2) and principal value of zero.

We decided to consider 20 μM as an appropriate activity/inactivity threshold. Accordingly, inhibitors of IC_{50} values <20 μM were regarded as "actives" and were assigned principal and MaxOmitFeat values of 2 and 0, respectively, On the other hand, inhibitors of IC_{50} > 20 μM were considered inactive and were assigned principal values of zero.[31] However, each inactive compound was carefully evaluated to assess whether its low potency is attributable to missing one or more pharmacophoric features, that is, compared with active compounds, or related to possible steric clashes within the binding pocket, or due to both factors and accordingly the MaxOmitFeat parameter was set to 1 or 2. Hip-Hop-Refine was configured to allow a maximum of 100 exclusion spheres to be added to the generated pharmacophoric hypotheses. Table E, Supporting Information, shows the training compounds employed in this step and their corresponding principal and MaxOmitFeat parameters.

4.1.9. Receiver Operating Characteristic (ROC) Curve Analysis. Selected pharmacophore models (i.e., Hypo1/7, Hypo8/8, and Hypo 9/1) were validated by assessing their abilities to selectively capture diverse Hsp90α active compounds from a large testing list of actives and decoys.

The testing list was prepared as described by Verdonk and co-workers.[46,47] Briefly, decoy compounds were selected based on three basic one-dimensional (1D) properties that allow the assessment of distance (D) between two molecules (e.g., i and j): (1) the number of hydrogen-bond donors (NumHBD); (2) number of hydrogen-bond acceptors (NumHBA); (3) count of nonpolar atoms (NP, defined as the summation of Cl, F, Br, I, S, and C atoms in a particular molecule). For each active compound in the test set, the distance to the nearest other active compound is assessed by their Euclidean distance (eq 4):

$$D(i,j) = \sqrt{(\mathrm{NumHBD}_i - \mathrm{NumHBD}_j)^2 + (\mathrm{NumHBA}_i - \mathrm{NumHBA}_j)^2 + (\mathrm{NP}_i - \mathrm{NP}_j)^2} \quad (4)$$

The minimum distances are then averaged over all active compounds (Dmin). Subsequently, for each active compound in the test set, around 30 decoys were randomly chosen from the ZINC database.[48] The decoys were selected in such a way that they did not exceed Dmin distance from their corresponding active compound.

To diversify active members in the list, we excluded any active compound having zero distance ($D(i,j)$) from other active compound(s) in the test set. Active testing compounds were defined as those possessing Hsp90α affinities ranging from 0.0006 to 6.0 μM. The test set included 33 active compounds and 989 ZINC decoys.

The test set (1022 compounds) was screened by each particular pharmacophore employing the "Best flexible search" option implemented in CATALYST, while the conformational spaces of the compounds were generated employing the "Fast conformation generation option" implemented in CATALYST. Compounds missing one or more features were discarded from the hit list. *In silico* hits were scored employing their fit values as calculated by eq C in Supporting Information.

The ROC curve analysis describes the sensitivity (Se or true positive rate, eq 5) for any possible change in the number of selected compounds (n) as a function of (1 − Sp). Sp is defined as specificity or true negative rate (eq 6).[47,49]

$$Se = \frac{\text{Number of selected actives}}{\text{Total number of actives}} = \frac{TP}{TP + FN} \quad (5)$$

$$\mathrm{Sp} = \frac{\text{Number of discarded inactives}}{\text{Total number of inactives}} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{6}$$

where TP is the number of active compounds captured by the virtual screening method (true positives), FN is the number of active compounds discarded by the virtual screening method, TN is the number of discarded decoys (presumably inactives), and FP is the number of captured decoys (presumably inactive).[47,49]

If all molecules scored by a virtual screening (VS) protocol with sufficient discriminatory power are ranked according to their score (i.e., fit values), starting with the best-scored molecule and ending with the molecule that got the lowest score, most of the actives will have a higher score than the decoys. Since some of the actives will be scored lower than decoys, an overlap between the distribution of active molecules and decoys will occur, which will lead to the prediction of false positives and false negatives.[47,49] The selection of one score value as a threshold strongly influences the ratio of actives to decoys and therefore the validation of a VS method. The ROC curve method avoids the selection of a threshold by considering all Se and Sp pairs for each score threshold.[47,49] A ROC curve is plotted by setting the score of the active molecule as the first threshold. Afterward, the number of decoys within this cutoff is counted and the corresponding Se and Sp pair is calculated. This calculation is repeated for the active molecule with the second highest score and so forth, until the scores of all actives are considered as selection thresholds.

The ROC curve representing ideal distributions, where no overlap between the scores of active molecules and decoys exists, proceeds from the origin to the upper-left corner until all the actives are retrieved and Se reaches the value of 1. In contrast to that, the ROC curve for a set of actives and decoys with randomly distributed scores tends toward the $\mathrm{Se} = 1 - \mathrm{Sp}$ line asymptotically with increasing number of actives and decoys.[47,49] The success of a particular virtual screening workflow can be judged from the following criteria (shown in Table 4): (1) Area under the ROC curve (AUC).[47,49,55] In an optimal ROC curve an AUC value of 1 is obtained; however, random distributions cause an AUC value of 0.5. Virtual screening that performs better than a random discrimination of actives and decoys retrieves an AUC value between 0.5 and 1, whereas an AUC value lower than 0.5 represents the unfavorable case of a virtual screening method that has a higher probability to assign the best scores to decoys than to actives.[47,49] (2) Overall accuracy (ACC) describes the percentage of correctly classified molecules by the screening protocol (eq 7). Testing compounds are assigned a binary score value of zero (compound not captured) or one (compound captured).[47,49,56]

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{N} = \frac{A}{N}\mathrm{Se} + \left(1 - \frac{A}{N}\right)\mathrm{Sp} \tag{7}$$

where N is the total number of compounds in the testing database and A is the number of true actives in the testing database. (3) Overall specificity (SPC) describes the percentage of discarded inactives by the particular virtual screening workflow. Inactive test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless to their individual fit values.[47,49,56] (4) Overall true positive rate (TPR or overall sensitivity) describes the fraction percentage of captured actives from the total number of actives. Active test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless of their individual fit values.[47,49,56] (5) Overall false negative rate (FNR or overall percentage of discarded actives) describes the fraction percentage of active compounds discarded by the virtual screening method. Discarded active test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless of their individual fit values.[47,49,56]

4.1.10. In Silico Screening for New Hsp90α Inhibitors. The sterically refined versions of Hypo1/7, Hypo8/8, and Hypo 9/1 were employed as 3D search queries to screen two 3D flexible structural databases, NCI and DAC. The screening was done employing the "Best Flexible Database Search" option implemented within CATALYST. NCI hits were filtered according to Lipinski's[50] and Veber's[51] rules. Remaining hits were fitted against the three pharmacophores using the "best fit" option within CATALYST. The fit values together with the relevant molecular descriptors of each hit were substituted in the optimal QSAR eq 1. The highest ranking molecules based on QSAR predictions were acquired and tested *in vitro*. Table 6 shows active hits and their QSAR predictions and experimental bioactivities.

4.2. *In Vitro* Experimental Studies. *4.2.1. Materials.* All of the chemicals used in these experiments were of reagent grade and were obtained from commercial suppliers. Hit compounds formoterol, amodiaquine, primaquine, midodrine, terfenadine, famotidine, isoxsuprine, diaveridine, bromodiolone, betaxolol, bumbuterol, clobenprobit, bicalutamide, atenolol, nicardipine, chlorhexidine, chlorbufam, and trimethoprim were purchased from Sigma-Aldrich (Germany) and were of purity >98%. NCI samples were kindly provided by the National Cancer Institute and were of purity >95%. Recombinant human Hsp90α (BIOQUOTE, York, U.K.), ATP 100× solution (BIOQUOTE, York, U.K.), geldanamycin (BIOQUOTE, York, U.K.), Quantichrome ATPase/GTPase Kit (BioAssay Systems, Hayward, CA), water for bioanalysis (Sigma, St. Louis, MO), DMSO for bioanalysis (Sigma, St. Louis, MO).

4.2.3. Preparation of Hit Compounds for In Vitro Assay. The tested compounds were provided as dry powders in variable quantities (5−500 mg). They were initially dissolved in DMSO to give stock solutions of 0.2 M. Subsequently, they were diluted to the required concentrations with deionized water for enzymatic assay.

4.2.4. Quantification of Hsp90α Activity in a Spectrophotometric Assay. The ATPase activity of Hsp90α was quantified by colorimetric measurement of released inorganic phosphate. Bioassays were performed by mixing Hsp90α solution (6 μL, 25 μg/mL in assay buffer), 24 μL of assay buffer, and 5 μL of the particular tested compounds to yield final inhibitor concentrations of 100, 10, 1, and 0.1 μM per well (in some cases 0.01 μM). The final concentration of DMSO did not exceed 1.0%. The mixtures were incubated for 30 min at 37 °C in an ELISA plate shaker, and then ATP solutions (5 μL, 4 mM in assay buffer) were added to each mixture. Blank was prepared as above except 5 μL of distilled water was used instead of inhibitor solution. The mixtures were equilibrated to 37 °C and incubated for 24 h.

The enzymatic reaction was terminated by the addition of 80 μL of malachite green ammoniummolybdate−tween 20 solution in 0.27 M H_2SO_4 and 10 μL of 34% sodium citrate. Color was allowed to develop at room temperature for 30 min, and sample absorbances were determined at $\lambda_{max} = 620$ nm using a plate reader (Bio-Tek instruments ELx 800, Winooski, VT). Inhibition of Hsp90α was calculated as percent activity of the uninhibited ATPase control. Geldanamycin was tested as positive control,[61] while negative controls were prepared by adding the substrate after reaction termination.[57,58,64]

ACKNOWLEDGMENT

This project was partially sponsored by the Faculty of Graduate Studies (This work is part of the Ph.D. Thesis of Mahmoud A. Al-Sha'er). The authors thank the Deanship of Scientific Research and Hamdi-Mango Center for Scientific Research at the University of Jordan for their generous funds.

Supporting Information Available: CATALYST modeling algorithm, CATALYST cost analysis in assessment of generated binding hypotheses, structures of Hsp90 inhibitors used in modeling, training subsets employed, training sets and CATALYST run parameters employed, Descriptor values of Hsp90α training and test inhibitors, and training subsets VII and VIII used for adding excluded spheres. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Prodromou, C.; Pearl, L. H. Structure and Functional Relationships of Hsp90. *Curr. Cancer Drug Targets* **2003**, *3*, 301–323.
(2) Edwards, C. R. W., Bouchier, I. A. D., Eds., Oncology. *Davidson's Principles and Practice of Medicine*, 16th ed.; ELBS publisher: Edinburgh, Churchill Livingstone, 1991; Chapter 7.
(3) Solit, D. B.; Rosen, N. Hsp90: A Novel Target for Cancer Therapy. *Curr. Top. Med. Chem.* **2006**, *6*, 1205–1214.
(4) Chiosis, G.; Rodina, A.; Moulick, K. Emerging Hsp90 Inhibitors: From Discovery to Clinic. *Anti-Cancer Agents Med. Chem.* **2006**, *6*, 1–8.
(5) Chiosis, G.; Lucas, B.; Shtil, A.; Huezoa, H.; Rosen, N. Development of a Purine-Scaffold Novel Class of Hsp90 Binders that Inhibit the Proliferation of Cancer Cells and Induce the Degradation of Her2Tyrosine Kinase. *Bioorg. Med. Chem.* **2002**, *10*, 3555–3564.
(6) Neckers, L. Using Natural Product Inhibitors to Validate Hsp90 as a Molecular Target in Cancer. *Curr. Top. Med. Chem.* **2006**, *6*, 1163–1171.
(7) Xiao, L.; Lu, X.; Ruden, D. M. Effectiveness of Hsp90 Inhibitors as Anti-Cancer Drugs. *Mini-Rev. Med. Chem.* **2006**, *6*, 1137–1143.
(8) Hideyuki, O.; Masami, K.; Yuichi, T.; Yumiko, U.; Jun, F.; Takayuki, N.; Shiro, S.; Makoto, S.; Shunichi., I.; Yoshinori, Y.; Endang, S. R.; Yutaka, K.; Michio, I. Conformational Significance of EH21A1-A4, Phenolic Derivatives of Geldanamycin, for Hsp90 Inhibitory Activity. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1577–1580.
(9) Kasibhatla, S. R.; Hong, K.; Biamonte, M. A.; Busch, D. J.; Karjian, P. L.; Sensintaffar, J. L.; Kamal, A.; Lough, R. E.; Brekken, J.; Lundgren, K.; Grecko, R.; Timony, G. A.; Ran, Y.; Mansfield, R.; Fritz, L. C.; Ulm, E.; Burrows, F. J.; Boehm, M. F. Rationally Designed High-Affinity 2-Amino-6-halopurine Heat Shock Protein 90 Inhibitors That Exhibit Potent Antitumor Activity. *J. Med. Chem.* **2007**, *50*, 2767–2778.
(10) Hwangseo, P.; Yun-Jung, K.; Ji-Sook, H. A Novel Class of Hsp90 Inhibitors Isolated by Structure-Based Virtual Screening. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6345–6349.
(11) Barril, X.; Beswick, M.; Collier, A.; Drysdale, M.; Dymock, B.; Fink, A.; Grant, K.; Howes, R.; Jordan, A.; Massey, A. 4-Amino Derivatives of the Hsp90 Inhibitor CCT018159. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2543–2548.
(12) Barril, X.; Brough, P.; Drysdale, M.; Hubbard, R. E.; Massey, A.; Surgenor, A.; Wright, L. Structure-Based Discovery of a New Class of Hsp90 Inhibitors. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 5187–5191.
(13) Jeffrey, R. H.; Chang, P.; Andrew, M. P.; Aaron, R. K.; Michael, D. W.; Xilu, W.; Christopher, L. L.; Jamey, C. M.; Kerry, M. S.; Russell, A. J.; Jun, C.; Paul, L. R.; Sha, J.; Stephen, K. T.; Edward, D. M.; Sarah, A. D.; Uri, S. L.; Jean, M. S.; Karl, A. W.; Diane, M. B.; Stephen, W. F.; Steven, W. E.; Philip, J. H. Discovery and Design of Novel HSP90 Inhibitors Using Multiple Fragment-based Design Strategies. *Chem. Biol. Drug.* **2007**, *70*, 1–12.
(14) Avila, C.; Hadden, M. K.; Ma, Z.; Kornilayevc, B. A.; Yeb, Q.-Z.; Blagg, B. S. J. High-Throughput Screening for Hsp90 ATPase Inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3005–3008.
(15) Sakkiah, S.; Thangapandian, S.; John, S.; Won, 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.
(16) Chen, C. Y.-C. Bioinformatics, Chemoinformatics, And Pharmainformatics Analysis of HER2/HSP90 Dual-Targeted Inhibitors. *J. Taiwan Inst. Chem. E.* **2010**, *41*, 143–149.
(17) Beeley; Nigel, R. A.; Sage, C. GPCRs: An Update on Structural Approaches to Drug Discovery. *Targets* **2003**, *2*, 19–25.
(18) Klebe, G. Virtual Ligand Screening: Strategies, Perspectives and Limitations. *Drug Discovery Today* **2006**, *11*, 580–594.
(19) Steuber, H.; Zentgraf, M.; Gerlach, C.; Sotriffer, C. A.; Heine, A.; Klebe, G. J. Expect the Unexpected or Caveat for Drug Designers: Multiple Structure Determinations Using Aldose Reductase Crystals Treated under Varying Soaking and Co-crystallisation Conditions. *Mol. Biol.* **2006**, *363*, 174–187.
(20) Stubbs, M. T.; Reyda, S.; Dullweber, F.; Moller, M.; Klebe, G.; Dorsch, D.; Mederski, W.; Wurziger, H. pH-Dependent Binding Modes Observed in Trypsin Crystals: Lessons for Structure-Based Drug Design. *ChemBioChem* **2002**, *3*, 246–249.
(21) DePristo, M. A.; de Bakker, P. I. W.; Blundell, T. L. Heterogeneity and Inaccuracy in Protein Structures Solved by X-ray Crystallography. *Structure* **2004**, *12*, 831–838.
(22) Neckers, L.; Mollapour, M.; Tsutsumi, S. The Complex Dance of the Molecular Chaperone Hsp90. *Trends Biochem. Sci.* **2009**, *34*, 223–226.
(23) Taha, M. O.; Bustanji, Y.; Al-Ghussein, M. A. S.; Mohammad, M.; Zalloum, H.; Al-Masri, I. M.; Atallah, N. Pharmacophore Modeling, Quantitative Structure-Activity Relationship Analysis and In Silico Screening Reveal Potent Glycogen Synthase Kinase-3β Inhibitory Activities for Cimetidine, Hydroxychloroquine and Gemifloxacin. *J. Med. Chem.* **2008**, *51*, 2062–2077.
(24) Taha, M. O.; Atallah, N.; Al-Bakri, A. G.; Paradis-Bleau, C.; Zalloum, H.; Younis, K.; Levesque, R. C. Discovery of New MurF Inhibitors via Pharmacophore Modeling and QSAR Analysis followed by in-silico screening. *Bioorg. Med. Chem.* **2008**, *16*, 1218–1235.
(25) Taha, M. O.; Bustanji, Y.; Al-Bakri, A. G.; Yousef, M.; Zalloum, W. A.; Al-Masri, I. M.; Atallah, N. Discovery of New Potent Human Protein Tyrosine Phosphatase Inhibitors via Pharmacophore and QSAR Analysis Followed by *in Silico* Screening. *J. Mol. Graphics Modell.* **2007**, *25*, 870–884.
(26) Al-masri, I. M.; Mohammad, M. K.; Taha, M. O. Discovery of DPP IV Inhibitors by Pharmacophore Modeling and QSAR Analysis Followed by *in Silico* Screening. *Chem. Med. Chem.* **2008**, *3*, 1763–1779.
(27) Taha, M. O.; Dahabiyeh, L. A.; Bustanji, Y.; Zalloum, H.; Saleh, S. Combining Ligand-Based Pharmacophore Modeling, QSAR Analysis and *In-Silico* Screening for the Discovery of New Potent Hormone Sensitive Lipase Inhibitors. *J. Med. Chem.* **2008**, *51*, 6478–6494.
(28) Al-Nadaf, A.; Abu Sheikha, G.; Taha, M. O. Elaborate Ligand-Based Pharmacophore Exploration and QSAR Analysis Guide the Synthesis of Novel Pyridinium-based Potent β-Secretase Inhibitory Leads. *Bioorg. Med. Chem.* **2010**, *18*, 3088–3115.
(29) Abu-Hammad, A. M.; Taha, M. O. Pharmacophore Modeling, Quantitative Structure-Activity Relationship Analysis, And Shape-Complemented in Silico Screening Allow Access to Novel Influenza Neuraminidase Inhibitors. *J. Chem. Inf. Model.* **2009**, *49*, 978–996.
(30) Abu Khalaf, R.; Abu Sheikha, G.; Bustanji, Y.; Taha, M. O. Discovery of New Cholesteryl Ester Transfer Protein Inhibitors via Ligand-Based Pharmacophore Modeling and QSAR Analysis Followed by Synthetic Exploration. *Eur. J. Med. Chem.* **2010**, *45*, 1598–1617.
(31) CATALYST 4.11 Users' Manual; *Accelrys Software Inc.*: San Diego, CA, 2005.
(32) Brough, P. A.; Aherne, W.; Barril, X.; Borgognoni, J.; Boxall, K.; Cansfield, J. E.; Cheung, K.-M. J.; Collins, I.; Davies, N. G. M.; 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-Diarylisoxazole Hsp90 Chaperone Inhibitors: Potential Therapeutic Agents for the Treatment of Cancer. *J. Med. Chem.* **2008**, *51*, 196–218.

(33) McDonald, E.; Jones, K.; Brough, P. A.; Drysdale, M. J.; Workman, P. Discovery and Development of Pyrazole-Scaffold Hsp90 Inhibitors. *Curr. Top. Med. Chem.* **2006**, *6*, 1193–1203.
(34) 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.
(35) Van Drie, J. H. Pharmacophore Discovery— Lessons Learned. *Curr. Pharm. Des.* **2003**, *9*, 1649–1664.
(36) Poptodorov, K.; Luu, T.; Langer, T.; Hoffmann, R. In *Methods and Principles in Medicinal Chemistry. Pharmacophores and Pharmacophores Searches*; Hoffmann, R. D., Ed.; Wiley-VCH: Weinheim, Germany, 2006; Vol. 2, pp 17−47.
(37) Sheridan, R. P.; Kearsley, S. K. Why Do We Need so Many Chemical Similarity Search Methods. *Drug Discovery Today* **2002**, *7*, 903–911.
(38) Li, H.; Sutter, J.; Hoffmann, R. In *Pharmacophore Perception, Development, and Use in Drug Design*; Güner, O. F., Ed.; International University Line: La Jolla, CA, 2000; pp 173−189.
(39) Sutter, J.; Güner, O.; Hoffmann, R.; Li, H.; Waldman, M. In *Pharmacophore Perception, Development, and Use in Drug Design*; Güner, O. F., Ed.; International University Line: La Jolla, CA, 2000; pp 501−511.
(40) Kurogi, Y.; Güner, O. F. Pharmacophore Modeling and Three Dimensional Database Searching for Drug Design Using Catalyst. *Curr. Med. Chem.* **2001**, *8*, 1035–1055.
(41) Bersuker, I. B.; Bahçeci, S.; Boggs, J. E. In *Pharmacophore Perception, Development, and Use in Drug Design*; Güner, O. F., Ed.; International University Line: La Jolla, CA, 2000, pp 457−473.
(42) Fischer, R. *The Principle of Experimentation Illustrated by a Psycho-Physical*, 8th ed.; ExpeHafner Publishing Co., Hafner Publishing: New York, 1966; Chapter II.
(43) *CERIUS2, QSAR Users' Manual*, version 4.10; Accelrys Inc.: San Diego, CA, 2005; pp43−88, 221−235, 237−250.
(44) Ramsey, L. F.; Schafer, W. D. *The Statistical Sleuth*, 1st ed.; Wadsworth Publishing Company: Belmont, CA, 1997.
(45) Clement, O. O.; Mehl, A. T. Pharmacophore Perception, Development, and Use in Drug Design. In *IUL Biotechnology Series*; Guner, O. F., Ed.; International University Line: La Jolla, CA, 2000; pp 71−84.
(46) Verdonk, M. L.; Marcel, L.; Berdini, V.; Hartshorn, M. J.; Mooij, W.T. M.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual Screening Using Protein−Ligand Docking: Avoiding Artificial Enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793–806.
(47) Kirchmair, J.; Markt, P.; Distinto, S.; Wolber, G.; Langer, T. Evaluation of the Performance of 3D Virtual Screening Protocols: RMSD Comparisons, Enrichment Assessments, and Decoy Selection−What Can We Learn from Earlier Mistakes. *J. Comput.-Aided. Mol. Des.* **2008**, *22*, 213–228.
(48) Irwin, J. J.; Shoichet, B. K. ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Comput. Sci.* **2005**, *45*, 177–182.
(49) Triballeau, N.; Acher, F.; Brabet, I.; Pin, J.-P.; Bertrand, H.-O. Virtual Screening Workflow Development Guided by the "Receiver Operating Characteristic" Curve Approach. Application to High-Throughput Docking on Metabotropic Glutamate Receptor Subtype 4. *J. Med. Chem.* **2005**, *48*, 2534–2547.
(50) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Delivery Rev.* **2001**, *46*, 3–26.
(51) Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623.
(52) Ganesh, T.; Thepchatri, P.; Li, L.; Du, Y.; Fu, H.; Snyder, J. P.; Sun, A. Synthesis and SAR Study of *N*-[4-Hydroxy-3-(2-hydroxynaphthalene-1-yl)phenyl]-arylsulfonamides: Heat Shock Protein 90 (Hsp90) Inhibitors with Submicromolar Activity in an *in Vitro* Assay. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4982–4987.
(53) Matsuda, M. M.; Moynet, D.; Molimard, M.; Dumazet, H. F.; Marit, G.; Reiffers, J.; Mossalayi, M. D. Long-Acting beta2-Adrenergic Formoterol and Salmeterol Induce the Apoptosis of B-Chronic Lymphocytic Leukaemia Cells. *Br. J. Hamaetol.* **2004**, *124*, 141–150.
(54) Krovat, E. M.; Langer, T. Non-Peptide Angiotensin II Receptor Antagonists: Chemical Feature Based Pharmacophore Identification. *J. Med. Chem.* **2003**, *46*, 716–726.
(55) Jacobsson, M.; Liden, P.; Stjernschantz, E.; Bostroem, H.; Norinder, U. Improving Structure-Based Virtual Screening by Multivariate Analysis of Scoring Data. *J. Med. Chem.* **2003**, *46*, 5781–5789.
(56) Gao, H.; Williams, C.; Labute, P.; Bajorath, J. Binary Quantitative Structure Activity Relationship (QSAR) Analysis of Estrogen Receptor Ligands. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 164–168.
(57) Lanzetta, P. A.; Alvarez, L. J.; Reinach, P. S.; Candia, O. A. An Improved Assay for Nanomole Amounts of Inorganic Phosphate. *Anal. Biochem.* **1979**, *100*, 95–97.
(58) Christopher, A.; Boris, A. K.; Brian, S. J. Development and Optimization of a Useful Assay for Determining Hsp90_s Inherent ATPase Activity. *Bioorg. Med. Chem.* **2006**, *14*, 1134–1142.
(59) Ganesh, T.; Min, J.; Thepchatri, P.; Dua, Y.; Li, L.; Lewis, I.; Wilson, L.; Fu, H.; Chiosis, G.; Dingledine, R.; Liotta, D.; Snyder, J. P.; Sun, A. Discovery of Aminoquinolines As a New Class of Potent Inhibitors of Heat Shock Protein 90 (Hsp90): Synthesis, Biology, and Molecular Modeling. *Bioorg. Med. Chem.* **2008**, *16*, 6903–6910.
(60) *CERIUS2 4.10 LigandFit User Manual*; Accelrys Inc.: San Diego, CA, 2000.;Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. LigandFit: A Novel Method for the Shape-Directed Rapid Docking of Ligands to Protein Active Sites. *J. Mol. Graphics Modell.* **2003**, *21*, 289–307.
(61) Howes, R.; Barril, X.; Dymock, B. W.; Grant, K.; NorthWeld, C. J.; Robertson, A. G. S.; Surgenor, A.; Wayne, J.; Wright, L.; James, K.; Matthews, T.; Cheung, K.-M.; McDonald, E.; Workman, P.; Drysdale, M. J. A Xuorescence Polarization Assay for Inhibitors of Hsp90. *Anal. Biochem.* **2006**, *350*, 202–213.
(62) Obermann, W. M. J.; Sondermann, H.; Russo, A. A.; Pavletich, N. P.; Hart, F. U. In Vivo Function of Hsp90 Is Dependent on ATP Binding and ATP Hydrolysis. *J. Cell Biol.* **1998**, *143*, 4901–910.
(63) Richter, K.; Moser, S.; Hagn, F.; Friedrich, R.; Hainz, O.; Heller, M.; Schlee, S.; Kessler, H.; Reinstein, J.; Buchner, J. Intrinsic Inhibition of the Hsp90 ATPase Activity. *J. Biol. Chem.* **2006**, *281*, 11301–11311.
(64) Rowlands, M. G.; Newbatt, Y. M.; Prodromou, C.; Pearl, L. H.; Workman, P.; Aherne, W. High-Throughput Screening Assay for Inhibitors of Heat-Shock Protein 90 ATPase Activity. *Anal. Biochem.* **2004**, *327*, 176–183.
(65) Bemis, G. W.; Murcko, M. A. The Properties of Known Drugs. 1. Molecular Frameworks. *J. Med. Chem.* **1996**, *39*, 2887–2893.
(66) *Discovery Studio 2.5.5 User Guide*, Accelrys Inc.: San Diego, CA, 2010.
(67) Wang, M.; Shen, G.; Blagg, B. S. J. Radanamycin, a Macrocyclic Chimera of Radicicol and Geldanamycin. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2459–2462.
(68) Chiosis, G.; Lucas, B.; Shtil, A.; Huezoa, H.; Rosen, N. Development of a Purine-Scaffold Novel Class of Hsp90 Binders that Inhibit the Proliferation of Cancer Cells and Induce the Degradation of Her2Tyrosine Kinase. *Bioorg. Med. Chem.* **2002**, *10*, 3555–3564.
(69) Llauger, L.; He, H.; Kim, J.; Aguirre, J.; Rosen, N.; Peters, U.; Davies, P.; Chiosis, G. Evaluation of 8-Arylsulfanyl, 8-Arylsulfoxyl, and 8-Arylsulfonyl Adenine Derivatives as Inhibitors of the Heat Shock Protein 90. *J. Med. Chem.* **2005**, *48*, 2892–2905.

CI100222K