J Mol Model
DOI 10.1007/s00894-013-2056-9

ORIGINAL PAPER

Discovery of novel urokinase plasminogen activator (uPA) inhibitors using ligand-based modeling and virtual screening followed by in vitro analysis

Mahmoud A. Al-Sha'er · Mohammad A. Khanfar · Mutasem O. Taha

Received: 18 August 2013 / Accepted: 28 October 2013

Abstract Urokinase plasminogen activator (uPA)—a serine protease—is thought to play a central role in tumor metastasis and angiogenesis and, therefore, inhibition of this enzyme could be beneficial in treating cancer. Toward this end, we explored the pharmacophoric space of 202 uPA inhibitors using seven diverse sets of inhibitors to identify high-quality pharmacophores. Subsequently, we employed genetic algorithm-based quantitative structure-activity relationship (QSAR) analysis as a competition arena to select the best possible combination of pharmacophoric models and physicochemical descriptors that can explain bioactivity variation within the training inhibitors (r^2_{162}=0.74, F-statistic=64.30, r^2_{LOO}=0.71, r^2_{PRESS} against 40 test inhibitors=0.79). Three orthogonal pharmacophores emerged in the QSAR equation suggesting the existence of at least three binding modes accessible to ligands within the uPA binding pocket. This conclusion was supported by receiver operating characteristic (ROC) curve analyses of the QSAR-selected pharmacophores. Moreover, the three pharmacophores were comparable with binding interactions seen in crystallographic structures of bound ligands within the uPA binding pocket. We employed the resulting pharmacophoric models and associated QSAR equation to screen the national cancer institute (NCI) list of compounds. The captured hits were tested in vitro. Overall, our modeling workflow identified new low micromolar anti-uPA hits.

Keywords Urokinase plasminogen activator · Ligand based analysis · Serine peptidase · Anticancer · Anti-inflammatory

Introduction

Urokinase-type plasminogen activator (uPA)

Urokinase-type plasminogen activator (uPA) is a serine protease that has been implicated as a key mediator of cellular invasion and tissue remodeling [1]. An inhibitor of uPA may have a therapeutic role in disease situations where uPA-driven degradation of extracellular matrix, or uPA-dependent cell migration is thought to be important including tumor growth, metastasis, angiogenesis and chronic wounds [2–7].

Evidence has also been obtained to suggest that uPA, or plasmin produced by its action, may play a role in preventing healing of chronic wounds [3, 7]. Consequently, a selective inhibitor for uPA could have therapeutic value in cancer and wound healing [1, 7].

The main focus of recent efforts towards the development of new uPA inhibitors concentrate on structure-based ligand design [8–10] and high throughput screening [11, 12]. To date, several uPA X-ray complexes are documented in the Protein Data Bank (e.g., PDB codes: 1OWD, 1OWE, 1SQO, 1SQT, 1SQA, 1CFL, 1EJN, 1OWH, 1OWK, 1OWJ, 1U6Q, 1YWH, 2OW8) with good resolution. However, although crystallographic structures are generally considered the most reliable structural information for drug design purposes, they are limited by inadequate resolution [13] and crystallization-related artifacts of the ligand–protein complex [14–16]. Moreover, crystallographic structures generally ignore structural

Electronic supplementary material The online version of this article (doi:10.1007/s00894-013-2056-9) contains supplementary material, which is available to authorized users.

M. A. Al-Sha'er
Faculty of Pharmacy, Zarqa University, Zarqa 13132, Jordan
~~e-mail: a.mahmoud@zu.edu.jo~~

M. A. Khanfar · M. O. Taha (✉)
Department of Pharmaceutical Sciences, Faculty of Pharmacy, University of Jordan, Amman, Jordan
e-mail: mutasem@ju.edu.jo

heterogeneity related to protein anisotropic motion and discrete conformational substrates [17].

The continued interest in designing new uPA inhibitors and the lack of adequate ligand-based computer-aided drug discovery efforts, which can overcome the drawbacks of structure-based design, combined with the significant induced fit flexibility observed for uPA [18], 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 uPA inhibitory scaffolds. We previously reported the use of this innovative approach towards the discovery of new inhibitory leads against glycogen synthase kinase-3β, [19] bacterial MurF [20], protein tyrosine phosphatase [21], DPP IV [22], hormone sensitive lipase [23], β-secretase [24], influenza neuraminidase [25], migration inhibitory factor [26], cyclin dependent kinase inhibitors (CDK1)[27], and heat shock protein 90 (Hsp90) [28].

Methods

Molecular modeling

Pharmacophore and QSAR modeling studies were performed using the CATALYST (HYPOGEN module) [33] and CERIUS2 software suites implemented in Discovery Studio 2.5.5 from Accelrys Inc. (San Diego, CA,, http://www.accelrys.com). Structure drawing was performed employing ChemDraw Ultra 7.0 (Cambridge Soft Corp. (http://www.cambridgesoft.Com), Cambridge, MA).

Data set and conformational analysis

The structures of 202 uPA inhibitors (**1**–**202**, Fig. 1, Table A in the electronic supplementary material) were collected from recently published literature [29–36]. Although the in vitro bioactivities of the collected inhibitors were gathered from separate articles, the fact that the bioactivities were expressed as affinity values (K_i) should minimize any discrepancies resulting from variations in bioassay procedure [26]. The logarithm transformation of K_i (μM) values were used in QSAR and pharmacophore modeling, thus linearly correlating the bioactivities with binding free energy change.

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.

The conformational space of collected each inhibitor (**1**–**202**, Fig. 1, Table A under electronic supplementary material) was explored adopting the "best conformer generation" option within CATALYST [37] 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, i.e., a conformational ensemble was generated with an energy threshold of 20 kcal/mol^{-1} from the local minimized structure at which has the lowest energy level and a maximum limit of 250 conformers per molecule [37, 38].

Generation and assessment of binding hypotheses

All 202 molecules with their associated conformational models were grouped into a spreadsheet. The biological data of the inhibitors were reported with an "Uncertainty" value of three, 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, seven structurally diverse training subsets were selected: subsets **I**, **II**, **III**, **IV**, **V**, **VI** and **VII** shown in Table B in the electronic supplementary material. The selected training sets were utilized to conduct 48 modeling runs to explore the pharmacophoric space of uPA inhibitors. Table C of the supplementary material shows the training subsets and different parameters implemented for each pharmacophore exploration run. The exploration process included altering number and type of possible binding features (hydrogen bond acceptors, hydrogen bond donors, aromatic rings, ionizable groups and hydrophobic features), feature spacing parameter (100 and 300 pm) and the maximum number of allowed features in the resulting pharmacophore hypotheses.

Pharmacophore modeling employing CATALYST proceeds through three consecutive steps: the constructive phase, subtractive phase and optimization phase (see CATALYST Modeling Algorithm under section **SM-1** in Supplementary Materials) [37–43]. In the optimization phase, CATALYST attempts to minimize a cost function for each hypotheses consisting of three terms: Weight cost, Error cost and Configuration cost (see CATALYST Cost Analysis in Assessment of Generated Binding Hypotheses under section **SM-2** in Supplementary Materials).

CATALYST-HYPOGEN cross-validates the resulting optimal pharmacophores using the Cat-Scramble module implemented in CATALYST. This validation procedure is based on Fischer's randomization test [44]. 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

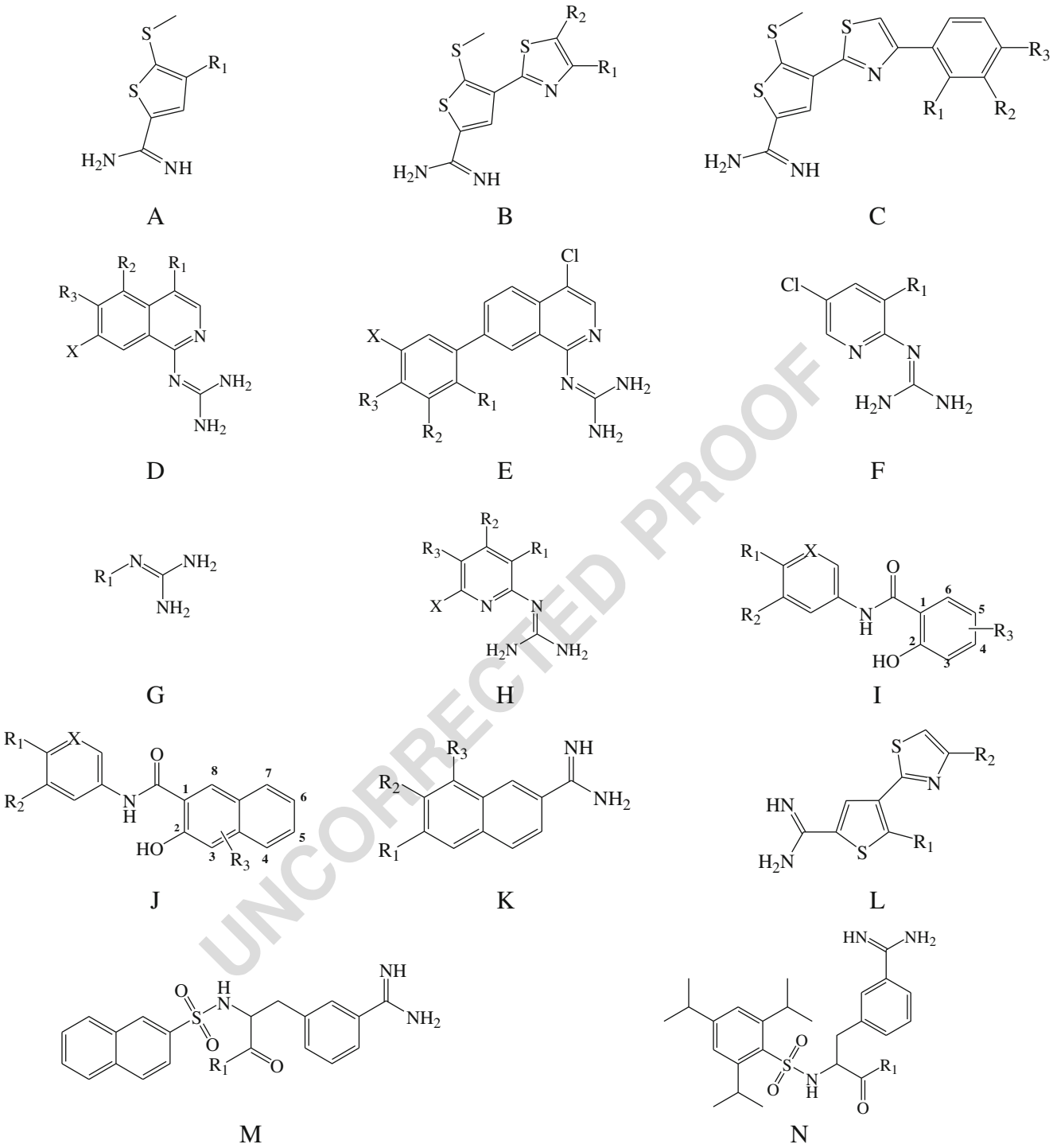


Fig. 1 Chemical scaffolds for urokinase plasminogen activator (uPA)

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 [37, 44, 45]. Based on Fischer randomization criteria; all 480 pharmacophores exceeded the 95 % significance threshold for subsequent processing. Table D under Supplementary Materials shows different cost criteria and significance levels of representative

pharmacophoric hypotheses (see pharmacophore clustering under QSAR modeling section).

QSAR modeling

The resulting pharmacophore models (480) were clustered into 45 groups utilizing the hierarchical average linkage method available in CATALYST. Subsequently, the highest-ranking representatives, as judged based on their significance F-values (calculated from correlating their fit values against the whole list of collected compounds with the corresponding molecular bioactivities) were selected to represent their corresponding clusters in subsequent QSAR modeling. Table D under Supplementary Materials 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.

QSAR modeling commenced by selecting a subset of 162 compounds from the total list of inhibitors (**1**–**202**, Fig. 1, Table A under Supplementary Materials) as a training set for QSAR modeling; the remaining 40 molecules (ca. 20 % of the dataset) were employed as an external test subset for validating the QSAR models. The test molecules were selected as follows: all 202 inhibitors were ranked according to their K_i values, and then every fifth compound was selected for the test set starting from the high-potency end. The selected test molecules should represent similar range of biological activities to that of the training set. The selected test inhibitors are marked with asterisks in Table A under Supplementary Materials.

The logarithm of measured K_i (μM) values was used in QSAR, thus correlating the data linear to the free energy change. Subsequently, we implemented genetic algorithm and multiple linear regression analyses to select optimal combination of pharmacophoric models and other physicochemical descriptors capable of self-consistent and predictive QSAR model. Section SM-3 under Supplementary Materials describes extensively the experimental details of QSAR modeling procedure [37, 46].

Addition of exclusion volumes

To account for the steric constrains of the binding pocket, we decided to complement our QSAR-selected pharmacophore models (i.e., Hypo34/2, Hypo37/3 and Hypo38/10) with exclusion volumes employing Hip-Hop-Refine module of CATALYST. Hip-Hop-Refine uses inactive training compounds to construct excluded volumes that resemble the steric constrains 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 [21–23, 37]. Subset VIII (in Table E under Supplementary Material) was used to construct exclusion spheres around Hypo34/2, Hypo37/3 and Hypo38/10. Section SM-4 under Supplementary Materials describes in details the Hip-Hop-Refine algorithm and settings implemented herein to decorate Hypo34/2, Hypo37/3 and Hypo38/10 with exclusion spheres.

The resulting sterically refined pharmacophores, as well as their unrefined versions, were validated by receiver operating characteristic curve analysis (ROC). [47–50], Theoretical and experimental details of this procedure are as shown in section SM-5 under Supplementary Material.

In silico screening for new uPA inhibitors

The sterically refined versions of Hypo34/2, Hypo37/3 and Hypo38/10 were employed as 3D search queries to screen the 3D flexible molecular database of the National Cancer Institute (NCI). The screening was done employing "Best Flexible Database Search" option implemented within CATALYST. Captured hits were filtered according to Lipinski's [51] and Veber's [49] rules. Remaining hits were fitted against Hypo34/2, Hypo37/3 and Hypo38/10 using the "best fit" option within CATALYST via implementing equation (D) in section SM-2 under Supplementary Materials. The fit values together with the relevant molecular descriptors of each hit were substituted in the optimal QSAR equation. The highest ranking molecules based on QSAR predictions were acquired and tested in vitro.

In vitro experimental studies

Materials

All chemicals used in these experiments were of reagent grade and obtained from commercial suppliers. NCI samples were kindly provided by the National Cancer Institute (http://www.cancer.gov/).

Quantification of the anti-uPA bioactivities of different hits

Bioassays were performed using the CHEMICON uPA kit for screening of uPA inhibitors [52]. The assay kit utilizes a chromogenic substrate, which is cleaved by active uPA enzyme. Addition of this substrate to a uPA-containing sample results in a colored product, detectable by its optical density at 405 nm. The assay was conducted as described in the uPA assay kit. Assay mixture (200 μL) composed of uPA (2.5 U, 2.5 μL), chromogenic substrate (L-pyroglutamyl-glycyl-L-arginine-p-nitroaniline hydrochloride, 20 μL, 2.5 mg/ml), 155 μL deionized H_2O (with or without inhibitor), and assay buffer (20 μL, pH 7.4) was mixed and incubated at 37 °C for 2 h. The absorbance of cleaved substrate was recorded at 405 nm. Tested hit concentrations ranged from 1 μM to 50 μM distributed log-linearly across the concentration range,

and at least two data points from each concentration were collected. The IC_{50} value for each experiment was obtained using nonlinear regression of the log(concentration) versus percent inhibition values (GraphPad Prism 5.0, http://www.graphpad.com). The assay conditions were validated by running positive (amiloride) and negative (deionized water) controls [52].

Results and discussion

Exploration of uPA pharmacophoric space

A total of 202 compounds were used in this study (**1**–**202**, see Fig. 1, Table A under supplementary material) [29–36]. We decided to explore the pharmacophoric space of uPA inhibitors through 48 HYPOGEN automatic runs and employing seven selected training subsets: subsets **I**–**VII** in Table B under supplementary material. The biological activity in the training subsets spanned from 3.5 to 4.0 orders of magnitude. The training compounds in these subsets were of maximal 3D diversity and continuous bioactivity spread over more than 3.5 logarithmic cycles [42]. CATALYST-HYPOGEN was restricted to explore pharmacophoric models incorporating from zero to one PosIon, one NegIon feature, from zero to three HBA, Hbic, and RingArom features, as shown in Table C under supplementary material. The input features were reasonably selected based on visual evaluation of the training compounds and comparison between the structures of potent, moderate and inactive members. Furthermore, we instructed the software to explore only four- and five-featured pharmacophores, i.e., ignore models of lesser number of features (as shown in Table C under supplementary material). The reader is referred to section Generation and Assessment of Binding Hypotheses in Methods and sections SM-1 and SM-2 under Supplementary Materials for more details about the CATALYST algorithm [38, 39, 42].

The resulting binding hypotheses from each automatic pharmacophore modeling run were ranked automatically according to their corresponding "total cost" value, which is defined as the sum of error cost, weight cost and configuration cost (see section Generation and Assessment of Binding Hypotheses in Methods and section SM-2 under Supplementary Materials) [37–42]. Error cost provides the highest contribution to total cost and is directly related to the capacity of the particular pharmacophore as 3D-QSAR model, i.e., in correlating the molecular structures to the corresponding biological responses [37, 39–43]. HYPOGEN also calculates the cost of the null hypothesis, which presumes that there is no relationship in the data and that experimental activities are distributed normally about their mean. Accordingly, the greater the difference from the null hypothesis cost (i.e., residual cost, Table D under Supplementary Materials) the more likely that the hypothesis does not reflect a chance correlation. CATALYST implements an additional validation technique based on Fisher's randomization test [45], namely, Cat.Scramble [37]. 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 Generation and Assessment of Binding Hypotheses in Methods) [37, 39–43].

Eventually, 480 pharmacophore models emerged from 48 automatic HYPOGEN runs, all of which exhibited Fisher randomization confidence levels ≥95 %. These successful models were clustered and the best representatives (45 models, see section Generation and Assessment of Binding Hypotheses under Methods and Table D under Supplementary Materials) were used in subsequent QSAR modeling.

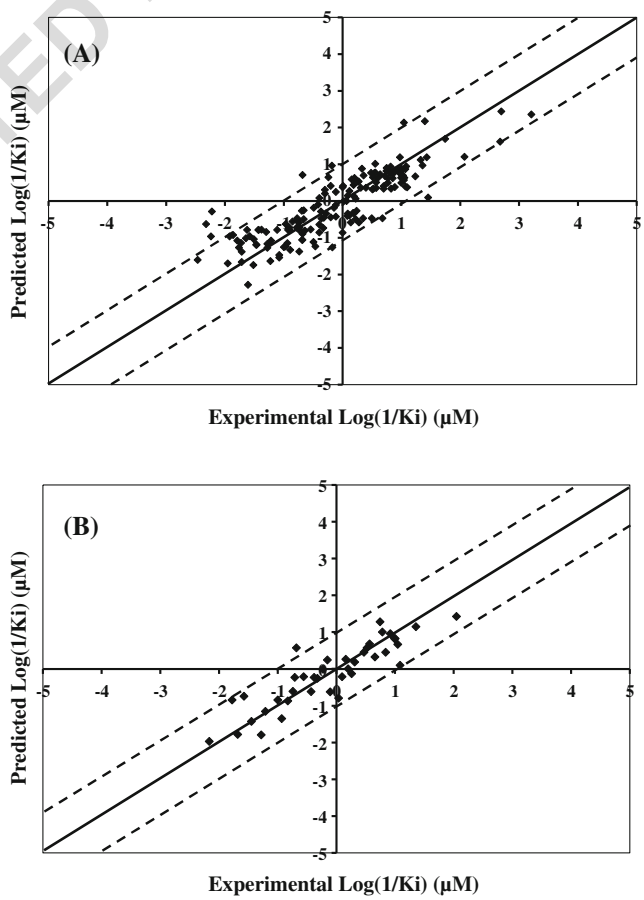


Fig. 2 Experimental versus **a** fitted (162 training compounds, r^2_{LOO}=0.71), and **b** predicted (40 test compounds, r^2_{PRESS}=0.79) bioactivities calculated from the best quantitative structure-activity relationship (QSAR) model Eq. (1). *Solid lines* Regression lines for fitted and predicted bioactivities of training and test compounds, respectively; *dotted lines* 1.0 log point error margins

Interestingly, the representative models shared comparable features and acceptable statistical success criteria.

The emergence of several statistically comparable pharmacophore models suggests the ability of uPA 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.

QSAR modeling

Despite the excellent value of pharmacophoric hypotheses in probing ligand–macromolecule recognition and as 3D search queries to search for new biologically interesting scaffolds, their predictive value as 3D-QSAR models is generally hampered by steric shielding and bioactivity-enhancing or reducing auxiliary binding groups (e.g., the biological effects of electron-donating and withdrawing substitutions) [19–28]. Moreover, our pharmacophore exploration of uPA inhibitors furnished hundreds of binding hypotheses of comparable success criteria, which makes it very hard to select any particular pharmacophore as sole representative of ligand binding within uPA. Accordingly, we were prompted 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**–**202**, Fig. 1, Table A). That is, we employed GFA-based QSAR as a competition arena to select the best pharmacophore(s), i.e., among the resulting population of binding models, and supplement it (them) with 2D descriptors to correct for the weaknesses of pharmacophore models (steric shielding and bioactivity-enhancing or reducing auxiliary binding groups). We employed a 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 (45 models) against collected uPA inhibitors (**1**–**202**, Fig. 1, Table A) were enrolled, together with around 100 other physicochemical descriptors, as independent variables in GFA-MLR-QSAR analysis [19–28, 45, 54]. We randomly selected 40 molecules (marked with asterisks in Table A under Supplementary Materials) and employed them as external test molecules for validating the QSAR models (r^2_{PRESS}). Additionally, all QSAR models were cross-validated automatically using the leave-one-out (LOO) cross-validation (see sections QSAR Modeling under Methods and section SM-3 under Supplementary Materials). [46, 54].

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.

Table 1 Pharmacophoric features and corresponding weights, tolerances and 3D coordinates of Hypo34/2, Hypo37/3 and Hypo38/10. *HBA* Hydrogen bond acceptors, *RingArom* aromatic rings, *Hbic* hydrophobic features

Model	Definition		Chemical features						
			HBA		HBA		RingArom		Hbic
Hypo34/2[a]	Weights		2.18		2.18		2.18		2.18
	Tolerances		1.60	2.20	1.60	2.20	1.60	1.60	1.60
	Coordinates	X	−5.14	−4.564	4.119	5.854	−0.9117	0.626	−0.8532
		Y	−3.338	−0.398	0.129	−1.018	−0.7359	−2.904	3.994
		Z	0.604	0.4555	−2.117	4.278	−1.906	−0.719	0.5818
			HBA		HBA		RingArom		Hbic
Hypo37/3[b]	Weights		1.806		1.806		1.806		1.806
	Tolerances		1.60	2.20	1.60	1.60	1.60	1.60	1.60
	Coordinates	X	−0.0.51	−0.662	−0.546	−0.330	0.294	1.735	−2.706
		Y	4.744	5.735	−0.0448	−0.7208	−4.394	2.561	−1.0022
		Z	0.4635	−2.364	−0.1923	2.7225	−4.946	−6.178	−0.4202
			HBA		HBA		RingArom		Hbic
Hypo38/10[b]	Weights		1.29		1.29		1.29	1.29	1.29
	Tolerances		1.6	2.2	1.6	1.60	1.6	1.6	1.6
	Coordinates	X	4.479	−0.821	−1.209	−2.949	3.058	4.221	4.573
		Y	8.359	−9.341	−14.949	−13.368	17.378	12.374	4.838
		Z	−1.982	4.907	4.779	2.915	−4.1832	−2.548	−1.419

[a] Hypo34/2: the 2nd pharmacophore hypothesis generated in the 34st HYPOGEN run (Table D under Supplementary Material)

[b] Hypo37/3: the 3th pharmacophore hypothesis generated in the 37th HYPOGEN run (Table D under Supplementary Material)

[b] Hypo38/10: the 10th pharmacophore hypothesis generated in the 38th HYPOGEN run (Table D under Supplementary Material)

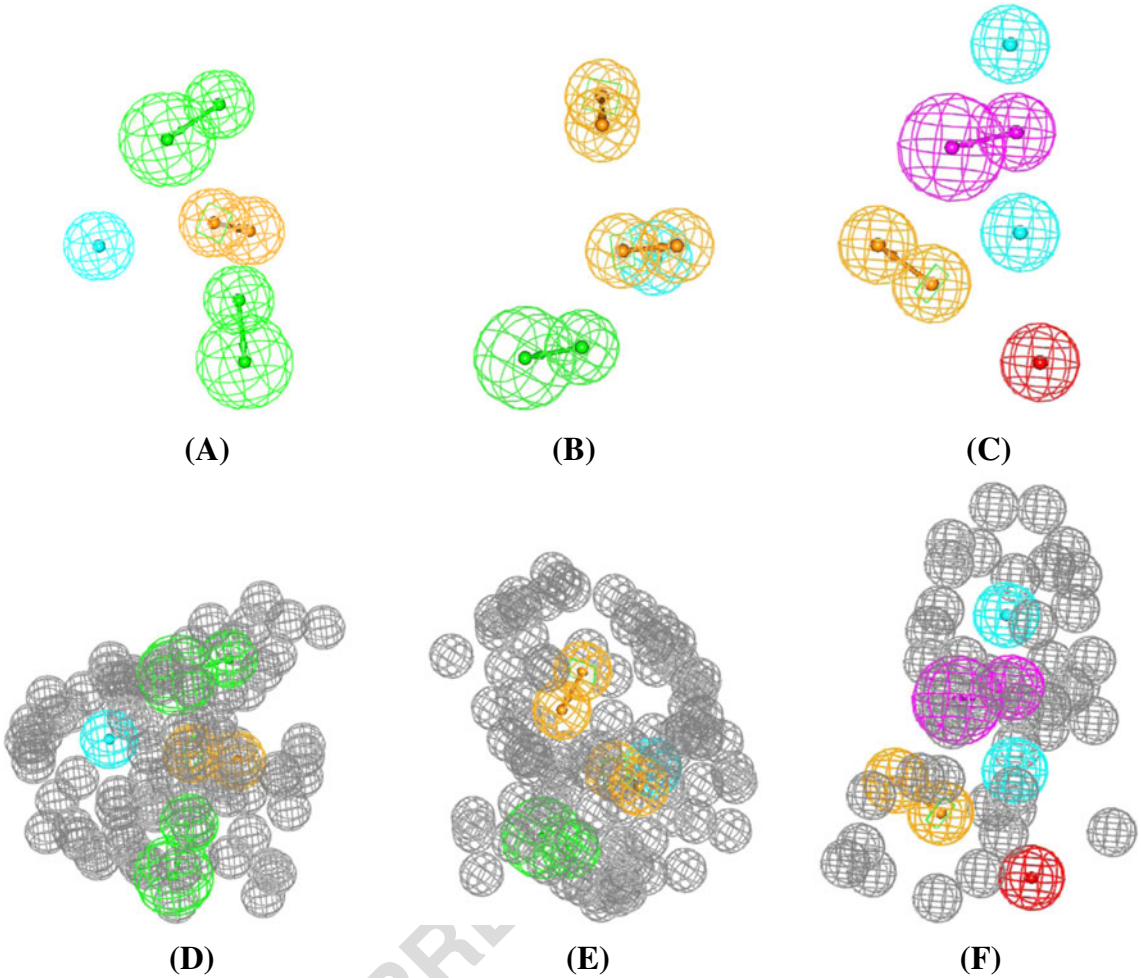


Fig. 3 Pharmacophoric features of **a** Hypo34/2, **b** Hypo37/3 and **c** Hypo38/10. *Pink vectored spheres* Hydrogen bond doner (HBD) features, *blue spheres* hydrophobic features (Hbic), *vectored orange spheres* aromatic rings (RingArom), *green vectored spheres* hydrogen bond acceptors (HBA), *red spheres* positive ionizable features (PosIon). **d**–**f** Sterically refined versions of Hypo34/2 (**d**), Hypo37/3 (**e**), and Hypo38/10 (**f**). *Gray spheres* Exclusion volumes

$$\begin{aligned}\mathrm{Log}\left(1/\mathrm{K_i}\right) = {} & -0.41[\pm0.13]-14.59[\pm2.1]\mathbf{dsN\_Count}-1.08\times10^{-2}[\pm0.01]\mathbf{dO\_Sum}\\ & +3.73[\pm0.54]\mathbf{dsN\_Sum}-6.08\times10^{-2}[\pm0.04]\mathbf{Num\_Rotatable\ Bonds}\\ & +0.13[\pm0.024]\mathbf{Hypo34/2}+0.16[\pm0.029]\mathbf{Hypo37/3}\\ & +0.18[\pm0.04]\mathbf{Hypo38/10}\end{aligned}$$

$$r^2_{\text{training}} = 0.74,\ \mathrm{F}_{\text{statistic}} = 64.3,\ r^2_{\text{LOO}} = 0.71,$$

$$r^2_{\text{PRESS-test}} = 0.79,\ \overline{r^2_{m\ \text{training}}} = 0.70,$$

$$\Delta r^2_{m\ \text{training}} = 0.021,\ Q^2_{F1} = 0.76,^{c}R^2_P = 0.72\ldots\ldots\ldots\ldots\ldots \quad (1)$$

where $r^2{}_{\text{training}}$ is the correlation coefficient against 162 training compounds, $\mathrm{F}_{\text{statistic}}$ is Fisher significance criteria, $r^2{}_{\text{LOO}}$ is the leave-one-out correlation coefficient, and $r^2{}_{\text{PRESS-test}}$ is the predictive r^2 determined for the 40 test compounds [45, 54]. $\overline{r^2_m}$ and Δr^2_{m} are the average and delta r^2_{m} values. Both are

Table 2 Receiver operating characteristic (ROC) curve analysis criteria for quantitative structure-activity relationship (QSAR)-selected pharmacophores and their sterically refined versions. *AUC* Area under the curve, *ACC* overall accuracy, *SPC* overall specificity, *TPR* overall true positive rate, *FNR* overall false negative rat

Pharmacophore model	ROC–AUC	ACC	SPC	TPR	FNR
Hypo34/2	0.75	0.97	0.98	0.77	0.02
Hypo37/3	0.83	0.97	0.98	0.92	0.02
Hypo38/10	0.99	0.97	1.00	0.13	0.003
Refined Hypo34/2	0.94	0.97	0.99	0.55	0.014
Refined Hypo37/3	0.93	0.97	0.98	0.76	0.02
Refined Hypo38/10	1.00	0.97	1.00	0.08	0.002

recently developed metrics that test the internal and external predictive capacities of a QSAR model extensively through establishing the proximity between predicted and observed response data among 162 training compounds. QSAR models

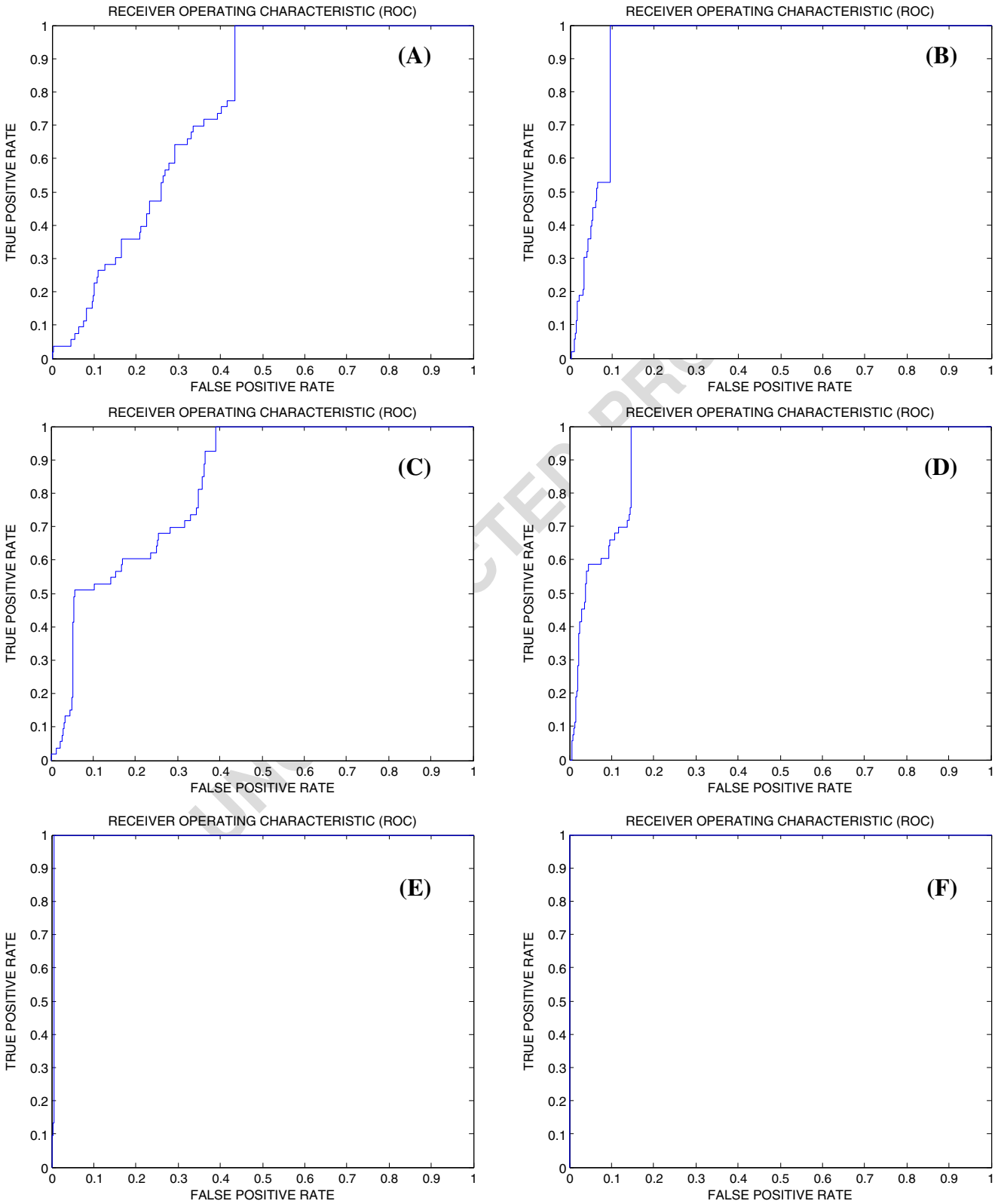


Fig. 4 Receiver operating characteristic (ROC) curves of **a** Hypo34/2, **b** sterically refined Hypo34/2, **c** Hypo 37/3, **d** sterically refined Hypo37/3, **e** Hypo38/10, **f** sterically refined Hypo38/10

of $\overline{r_m^2} > 0.5$ and $\Delta r_m^2 < 0.2$ are considered predictive and reliable [61, 62]. Q_{F1}^2 is a prediction metric proposed by Shi et al. [63] and calculated using the external testing list (40 compounds). To further establish the statistical significance of the QSAR model we performed Y randomization tests by randomly shuffling the dependent variable 100 times while keeping the independent variables as it is. ${}^{c}R_P^2$ is a metric derived from the difference between $r_{training}^2$ and average $r_{training}^2$ of random models. ${}^{c}R_P^2$ should be >0.50 for passing this test [66]. Based on these metrics, as well as others, QSAR Eq. (1) was found to pass Golbraikh and Tropsha criteria [64, 65].

The reader is refered to the Supplementary Materials (section SM-6 and Table J) to evaluate the significance of the QSAR model through extensive list of validation techniques.

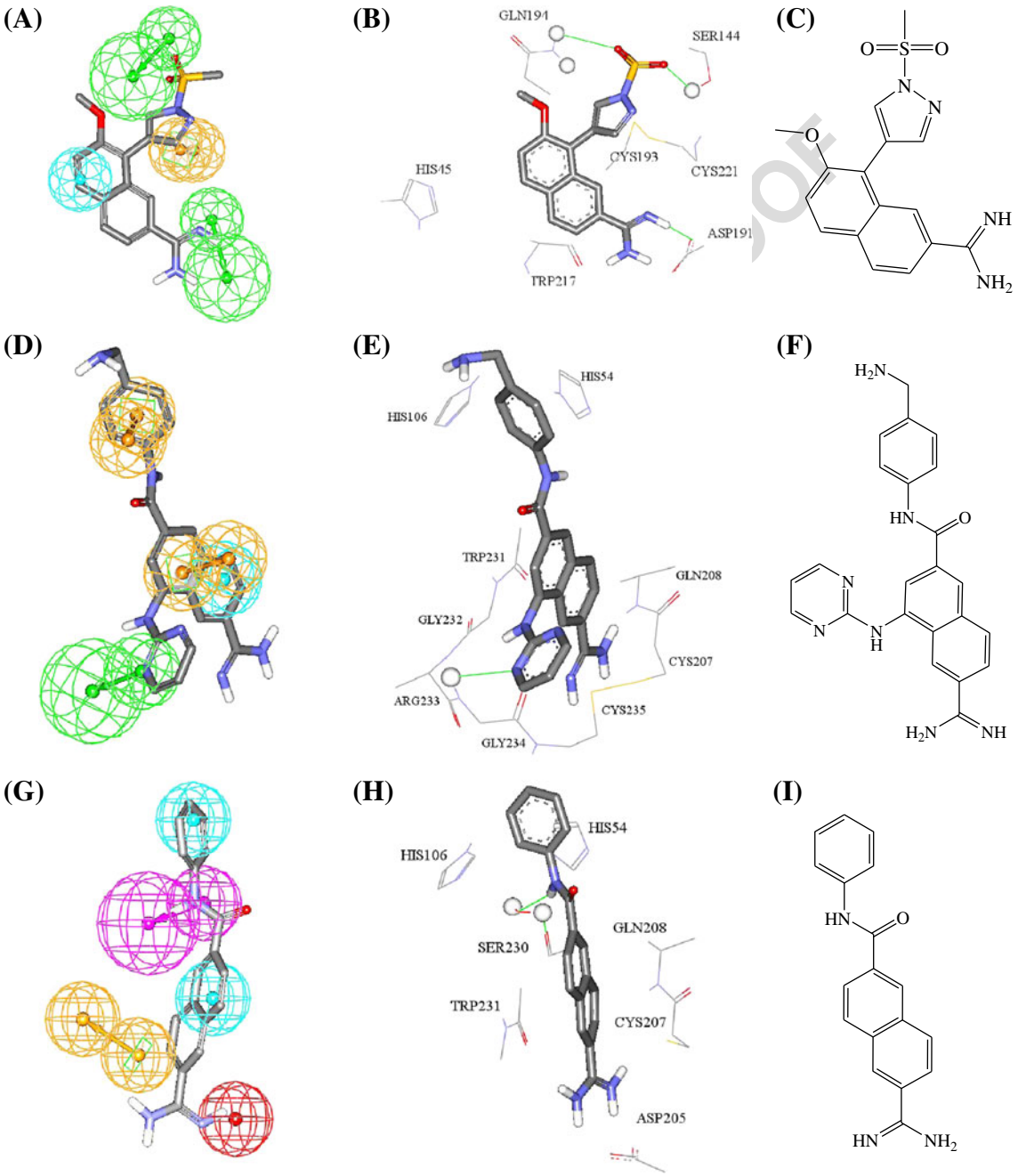


Fig. 5 **a** Mapping of compound **148** (K_i=0.63 μM, Table A under Supplementary Materials) against Hypo34/2. **b** Co-crystallized complex of **148** within uPA (PDB code: 1SQT, resolution=1.90 Å). **c** Chemical structure of **148**. **d** Mapping of compound **158** (K_i=0.0006 μM, Table A under Supplementary Materials) against Hypo37/3. **e** Co-crystallized complex of **158** within uPA (PDB code: 1SQA, resolution=2.0 Å). **f** Chemical structure of **158**. **g** Mapping of compound **142** against Hypo38/10. **h** Co-crystallized complex of **142** within uPA (PDB code: 1OWE, resolution=1.6 Å). **i** chemical structure of **142**

Hypo34/2, Hypo37/3 and Hypo38/10 (Table 1) represent the fit values of the training compounds against these pharmacophores (shown in Fig. 2) as calculated from equation (D) in Supplementary Materials [33]. **dsN_Count**, **dsN_Sum**, **dO_Sum** are electrotopological state indices related to the number of imine nitrogen (**dsN_Count** and **dsN_Sum**) and ether oxygen atoms (**dO_Sum**) in training molecules [46]. **Num_RotaTable Bonds** is the number of rotatable bonds defined as any single non-ring bond, bonded to a nonterminal heavy (i.e., non-hydrogen) atom. Amide C–N bonds are not considered because of their high rotational energy barrier [46]. Table H and Table I show the values molecular descriptors of QSAR Eq. (1) as calculated for training and testing compounds, respectively.

Emergence of three reasonably orthogonal pharmacophoric models, i.e., Hypo34/2, Hypo37/3 and Hypo38/10 (Table G under Supplementary Material shows their cross-correlation coefficient) in Eq. (1) suggests that they represent three complementary binding modes accessible to ligands within the binding pocket of uPA. Similar conclusions were reached about the binding pockets of other targets based on QSAR analysis [19–28]. Figure 3 shows Hypo34/2, Hypo37/3 and Hypo38/10. The X, Y, and Z coordinates of the three pharmacophores are illustrated in Table 1.

Interestingly, the regression slopes of the three pharmacophore models suggest they make mediocre but rather equivalent contributions to bioactivity. Nevertheless, these models illustrated excellent abilities in separating active compounds from inactive decoys in ROC analysis [47–49, 55]. Table 2 and Fig. 4 show the ROC results of our QSAR-selected pharmacophores (see SM-5 Receiver Operating Characteristic Curve Analysis under Supplementary Materials for more details).

To correlate the binding features in Hypo34/2, Hypo37/3 and Hypo38/10 with ligand-receptor binding interactions anchoring inhibitors into the binding pocket of uPA, we compared the pharmacophoric features of Hypo34/2, Hypo37/3 and Hypo38/10 with the way in which they map three co-crystallized ligands (**148**, **158** and **142**) within uPA (PDB codes: 1SQA, 1SQT and 1OWE) [34, 57] as shown in Fig. 5. Figure 5a,d,g compares how training compounds **148**, **158** and **142** (Table A under Supplementary Materials) map Hypo34/2, Hypo37/3 and Hypo38/10 with the way these ligands bind within uPA's binding pocket (PDB code: 1SQT, 1SQA and 1OWE, respectively).

From Fig. 5a and b, mapping the sulfonyl oxygen of **148** against a HBA in Hypo34/2 corresponds clearly to hydrogen bonding interaction connecting the same sulfone group with the amidic NH and OH of Gln194 and Ser144, respectively. Similarly, π-stacking interactions anchoring the pyrazole aromatic ring of **148** against the disulfide bridge of Cys221 and Cys193 seem to correspond to fitting the same pyrazole ring against the aromatic ring (RingArom) feature in Hypo34/2. Furthermore, fitting the terminal amidine group of **148** against the hydrogen bond acceptor (HBA) feature in Hypo34/2, correlates with hydrogen-bonding interactions connecting the amidino group with the carboxylate residues of Asp191. Finally, the fact that the naphthalene linker reside within a hydrophobic pocket comprised of Cys193, Trp217 and His45 correspond to fitting this group against this hydrophobic feature (Hbic) in Hypo34/2.

Figure 5d and e compare the co-crystallized pose of **158** in uPA (PDB code: 1SQA) with the way it maps Hypo37/3. Mapping the heterocyclic nitrogen atom of the pyrimidinyl ring against HBA features in Hypo37/3 corresponds to hydrogen-bonding interaction connecting this nitrogen to the peptidic NH of Gly234 (bonded to Arg233). Similarly, fitting the terminal benzylamine aromatic ring of **158** against the RingArom feature in Hypo37/3 agrees with π-stacking interactions resulting from inserting the particular aromatic ring between the imidazole rings of His54 and His106.

Similarly, mapping the naphthyl residue of **158** against Hbic and RingArom features in Hypo37/3 correlates with hydrophobic proximity between this substituent and hydrophobic side chains of Gly232, Cys207 and Cys235, and π-stacking with peptidic amides of Gln208 and Trp231.

Finally, Fig. 5g and h compare the co-crystallized pose of **142** in uPA (PDB code: 1OWE) with the way Hypo38/10 maps **142**. Mapping the amide NH of **142** against HBD feature in Hypo38/10 corresponds to hydrogen-bonding interactions connecting this nitrogen to the carbonyl oxygen of Ser230 via a bridging water molecule. Moreover, fitting the

Table 3 Numbers of captured hits by sterically refined versions of Hypo34/2, Hypo37/3 and Hypo38/10

		Pharmacophore models		
3D Database[a]	Post screening filtering[b]	Sterically-refined Hypo34/2	Sterically-refined Hypo37/3	Sterically-refined Hypo38/10
NCI	Before	8633	7771	145
	After	3402	5531	113

[a] National Cancer Institute list of available compounds (238,819 structures)

[b] Using Lipinski's [51] and Veber's [49] rules

Table 4 Predicted and experimental bioactivities of high-ranking hit molecules

Hits[a]	Name[b]	Experimental % inhibition	
		at 10 μM[c]	IC_{50} (μM)[d]
203	135,766	63	6.3
204	666,712	57	9.0
205	4,367	55	11.3
206	144,205	41	28.4
237[e]	Amiloride	42	12.3

[a] Chemical structures shown in Fig. 9

[b] NCI number

[c] Experimental percentage of inhibition determined at 10 μM inhibitor concentrations

[d] IC_{50} values experimentally determined for most active hits

[e] Reported Amiloride uPA inhibitory IC_{50}=7.0 μM. [58] Each values represents the average of duplicate measurements

naphthalene aromatic system of **142** against RingArom and Hbic features in Hypo38/10 agrees with π-stacking this ring system against the amidic backbone of Cys207 and Trp231 and its close proximity to the hydrophobic linker of Gln208. Additionally, mapping the terminal anilide ring of **142** against Hbic feature in Hypo38/10 agrees with stacking this ring between aromatic rings of His106 and His54. Finally, mapping the amidine of **142** against PosIon feature in Hypo38/10 corresponds to ionic attraction connecting this positive group with the carboxylate side chain of Asp205.

Steric refinement, virtual screening and in vitro validation

Pharmacophores serve as useful 3D QSAR models and 3D search queries; however, they lack the steric constrains necessary to define the size of the binding pocket. This liability renders pharmacophoric models rather promiscuous in some cases [25]. Therefore, we decided to complement the optimal pharmacophores with exclusion spheres employing the Hip-Hop-Refine module implemented within CATALYST [37]. Excluded volumes resemble sterically inaccessible regions within the binding site (see section SM-4: Hip-Hop-Refine

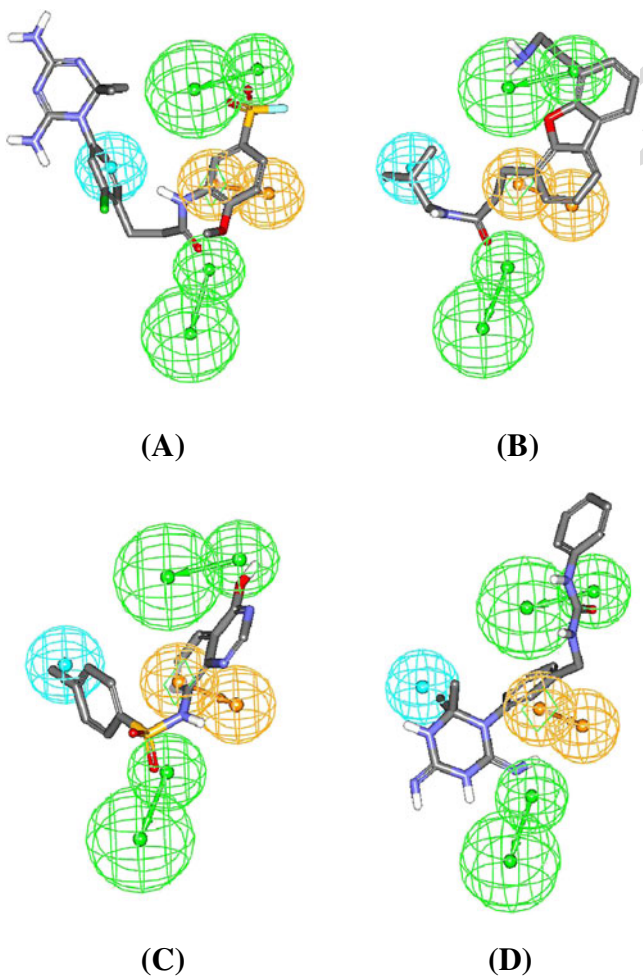


Fig. 6 **a**, **b**, **c** and **d** show Hypo34/2 fitted against active hits **203**, **204**, **205** and **206**, respectively

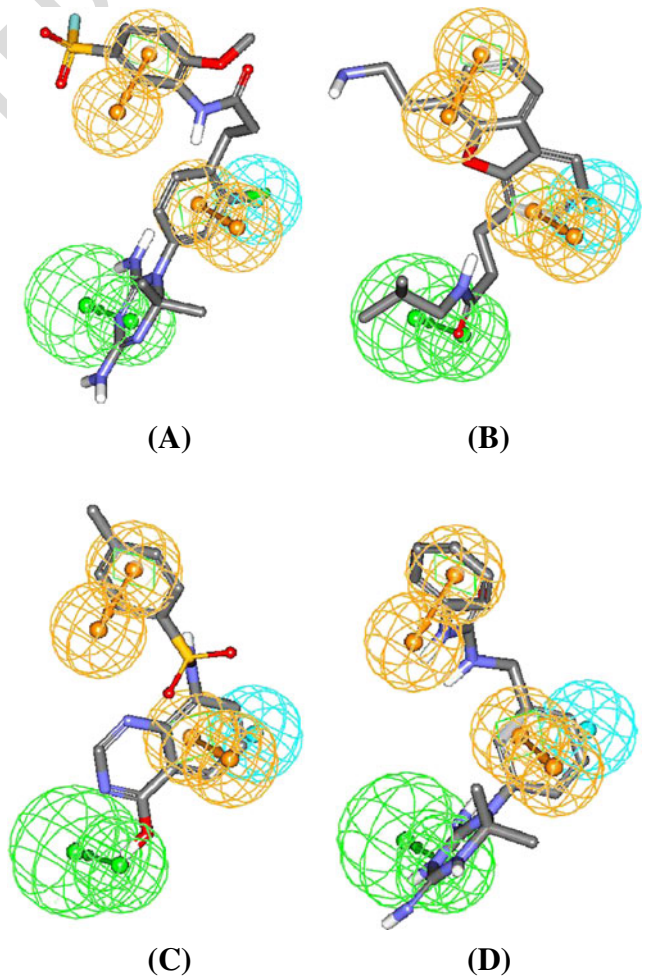
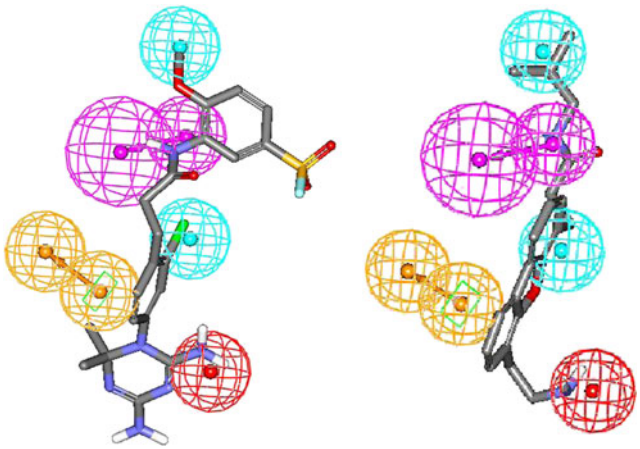


Fig. 7 **a**, **b**, **c** and **d** show Hypo37/3 fitted against hits **203**, **204**, **205** and **206**, respectively



(A) (B)

Fig. 8 **a** and **b** show Hypo38/10 fitted against active hits **203** and **204**, respectively

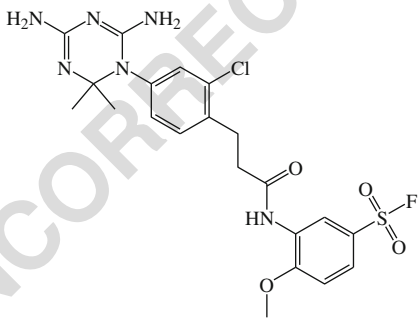
algorithm and employed settings under Supplementary Material for more details) [56].

We selected a diverse training subset for Hip-Hop-Refine modeling (subset **VIII** in Table E under supplementary material). The training compounds were selected in such a way that the bioactivities of weakly active compounds are explainable by steric clashes within the binding pocket.

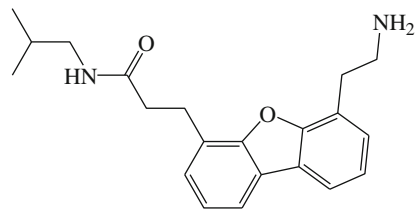
We assessed the success of steric refinement experiments through ROC analysis of the sterically refined pharmacophore versions. Table 2 shows the ROC results of the refined pharmacophores compared to their unrefined counterparts. Clearly, steric refinement improved the classification power of the three pharmacophores. This effect was particularly evident with Hypo34/2 and Hypo37/3, which had their ROC areas under the curve (AUCs) increased from 75 % and 83 % to 94 % and 93 %, respectively. However, the effect of steric refinement on the efficiency of Hypo38/10 was less drastic. This is not surprising, since this pharmacophore is inherently of superior classification power due to the presence of a PosIon features among its binding features.

Sterically refined Hypo34/2 (Fig. 3d), Hypo37/3 (Fig. 3e) and Hypo38/10 (Fig. 3f) were employed as 3D search queries against the National Cancer Institute list of compounds (NCI, 238,819 structures). Table 3 summarizes the numbers of captured hits by sterically refined versions of the pharmacophores. Subsequently, captured hits were filtered based on Lipinski's and Veber's rules, [50, 51]. The remaining hits were fitted against Hypo34/2, Hypo37/3 and Hypo38/10 and their fit values, together with other relevant molecular descriptors, were substituted in QSAR Eq. (1) to predict their anti-uPA bioactivities. The

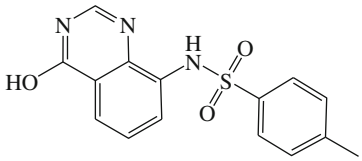
Fig. 9 Chemical structure of the most active hits



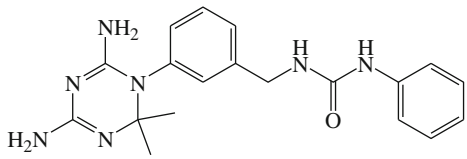
203 (NCI0135766)



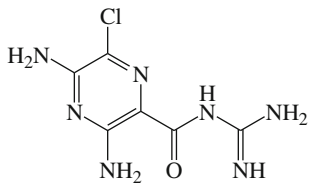
204 (NCI0666712)



205 (NCI004367)



206 (NCI0144205)



(237) Amiloride

highest-ranking hits were evaluated in vitro against human uPA [52]. Figure 9 and Table 4 shows the most active hits, while Table F under supplementary material shows other less active hits. Figures 6, 7 and 8 show how the most potent hits **203, 204, 205** and **206** map against Hypo34/2, Hypo37/3 and Hypo38/10.

Interestingly, although three of our hits shared related chemical functionalities with known anti-uPA compounds, e.g., guanidines, amidines and sulfonamides (i.e., **203**, **205** and **206**, Fig. 9), one of the hits, i.e., **204** (IC_{50}=9.0 μM, Table 4 and Fig. 9) is completely novel and represents a new class of uPA inhibitors that can be potentially optimized into interesting new drug molecules. It should be mentioned that the absence of guanidine and amidine groups from **204** should enhance the bioavailability of this class of anti-uPA agents.

Conclusions

uPA inhibitors are currently considered as potential treatments for cancer. The pharmacophoric space of uPA inhibitors was explored via seven 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 optimal QSAR model capable of explaining anti-uPA bioactivity variation across 202 collected uPA inhibitors. Three pharmacophoric models emerged in the QSAR equation suggesting the existence of more than one binding modes accessible to ligands within uPA binding pocket. The QSAR equation and the associated pharmacophoric models were validated experimentally by the identification of several uPA inhibitors retrieved via in silico screening, out of which three NCI hits illustrated superior potencies over the standard uPA inhibitor amiloride. Our results suggest that the combination of pharmacophoric exploration and QSAR analyses can be useful tool for finding new diverse uPA inhibitors.

Acknowledgments The authors thank the Deanship of Scientific Research and Hamdi-Mango Center for Scientific Research at the University of Jordan for their generous funds. We are also thankful for NCI institution for supporting us with free samples.

References

1. Rosenberg S (1999) The urokinase-type plasminogen activator and its receptor in cancer. Annu Rep Med Chem 34:121–128
2. Fazioli F, Blasi F (1994) Urokinase-type plasminogen activator and its receptor: new targets for anti-metastatic therapy? Trends Pharmacol Sci 15:25–29
3. Evans DM, Sloan-Stakleff KD (1998) Maximum effect of urokinase plasminogen activator inhibitors in the control of invasion and metastasis of rat mammary cancer. Invasion Metastasis 18:252–260
4. Stacey MC, Burnand KG, Mahmoud-Alexandroni M, Gaffney PJ, Bhogal BS (1993) Tissue and urokinase plasminogen activators in the environs of venous and ischaemic leg ulcers. Br J Surg 80:596–599
5. Palolahti M, Lauharanta J, Stephens RW, Kuusela P, Vaheri A (1993) Proteolytic activity in leg ulcer exudate. Exp Dermatol 2:29–37
6. Rogers AA, Burnett S, Moore JC, Shakespeare PG, John Chen WY (1995) Involvement of proteolytic enzymes—plasminogen activators and matrix metalloproteinases—in the pathophysiology of pressure ulcers. Wound Repair Regen 3:273–283
7. Wysocki AB, Kusakabe AO, Chang S, Tuan T-L (1999) Temporal expression of urokinase plasminogen activator, plasminogen activator inhibitor and gelatinase-B in chronic wound fluid switches from a chronic to acute wound profile with progression to healing. Wound Repair Regen 7:154–165
8. Matthews H, Ranson M, Tyndall JDA, Kelso MJ (2011) Synthesis and preliminary evaluation of amiloride analogs as inhibitors of the urokinase-type plasminogen activator (uPA). Bioorg Med Chem Lett 21:6760–6766
9. West CW, Adler M, Arnaiz D, Chen D, Chu K, Gualtieri G, Ho E, Huwea C, Light D, Phillips G, Pulk R, Sukovich D, Whitlow M, Yuan S, Bryant J (2009) Identification of orally bioavailable, non-amidine inhibitors of urokinase plasminogen activator (uPA). Bioorg Med Chem Lett 19:5712–5715
10. Pandya V, Jain M, Chakrabarti G, Soni H, Parmar B, Chaugule B, Patel J, Joshi J, Joshi N, Rath A, Raviya M, Shaikh M, Sairam KVVM, Patel H, Patel P (2011) Discovery of inhibitors of plasminogen activator inhibitor-1: structure–activity study of 5-nitro-2-phenoxybenzoic acid derivatives. Bioorg Med Chem Lett 21:5701–5706
11. Ye B, Bauer S, Buckman BO, Ghannam A, Griedel BD, Khim S-K, Lee W, Sacchi KL, Shaw KJ, Liang A, Wu Q, Zhao Z (2003) Synthesis and biological evaluation of menthol-based derivatives as inhibitors of plasminogen activator inhibitor-1 (PAI-1). Bioorg Med Chem Lett 13:3361–3365
12. Gopalsamy A, Kincaid SL, Ellingboe JW, Groeling TM, Antrilli TM, Krishnamurthy G, Aulabaugh A, Friedrichsb GS, Crandall DL (2004) Design and synthesis of oxadiazolidinediones as inhibitors of plasminogen activator inhibitor-1. Bioorg Med Chem Lett 14:3477–3480
13. Beeley RA, Sage NC (2003) GPCRs: an update on structural approaches to drug discovery. Targets 2:19–25
14. Klebe G (2006) Virtual ligand screening: strategies, perspectives and limitations. Drug Discov Today 11:580–594
15. Steuber H, Zentgraf M, Gerlach C, Sotriffer CA, Heine A, Klebe GJ (2006) 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 363:174–187
16. Stubbs MT, Reyda S, Dullweber F, Moller M, Klebe G, Dorsch D, Mederski W, Wurziger H (2002) pH-dependent binding modes observed in trypsin crystals: lessons for structure-based drug design. Chem Biol Chem 3:246–249
17. DePristo MA, de Bakker PIW, Blundell TL (2004) Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography. Structure 12:831–838
18. Mertens HD, Kjaergaard M, Mysling S, Gårdsvoll H, Jørgensen TJ, Svergun DI, Ploug M (2012) A flexible multidomain structure drives the function of the urokinase-type plasminogen activator receptor (uPAR). J Biol Chem 287:34304–34315
19. Taha MO, Bustanji Y, Al-Ghussein MAS, Mohammad M, Zalloum H, Al-Masri IM, Atallah N (2008) 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 51:2062–2077

20. Taha MO, Atallah N, Al-Bakri AG, Paradis-Bleau C, Zalloum H, Younis K, Levesque RC (2008) Discovery of new MurF inhibitors via pharmacophore modeling and QSAR analysis followed by in-silico screening. Bioorg Med Chem 16:1218–1235
21. Taha MO, Bustanji Y, Al-Bakri AG, Yousef M, Zalloum WA, Al-Masri IM, Atallah N (2007) Discovery of new potent human protein tyrosine phosphatase inhibitors via pharmacophore and QSAR analysis followed by in silico screening. J Mol Graph Model 25:870–884
22. Al-masri IM, Mohammad MK, Taha MO (2008) Discovery of DPP IV inhibitors by pharmacophore modeling and QSAR analysis followed by in silico screening. Chem Med Chem 3:1763–1779
23. Taha MO, Dahabiyeh LA, Bustanji Y, Zalloum H, Saleh S (2008) Combining ligand-based pharmacophore modeling, QSAR analysis and in-silico screening for the discovery of new potent hormone sensitive lipase inhibitors. J Med Chem 51:6478–6494
24. Al-Nadaf A, Abu Sheikha G, Taha MO (2010) Elaborate ligand-based pharmacophore exploration and QSAR analysis guide the synthesis of novel pyridinium-based potent β-secretase inhibitory leads. Bioorg Med Chem 18:3088–3115
25. Abu-Hammad AM, Taha MO (2009) Pharmacophore modeling, quantitative structure–activity relationship analysis, and shape-complemented in silico screening allow access to novel influenza neuraminidase inhibitors. J Chem Inf Model 49:978–996
26. Al-Sha'er MA, VanPatten S, Al-Abed Y, Taha MO (2013) Elaborate ligand-based modeling reveal new migration inhibitory factor inhibitors. J Mol Graph Model 42:104–114
27. Al-Sha'er MA, Taha MO (2010) Discovery of novel CDK1 inhibitors by combining pharmacophore modeling, QSAR analysis and in silico screening followed by in vitro bioassay. Eur J Med Chem 45:4316–4330
28. Al-Sha'er MA, Taha MO (2010) Elaborate ligand-based modeling reveals new nanomolar heat shock protein 90α inhibitors. J Chem Inf Model 50:1706–1723
29. Barber CG, Dickinson RP (2002) Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 2: (3-substituted-5-halo-2-pyridinyl)guanidines. Bioorg Med Chem Lett 12:185–187
30. Subasinghe NL, Illig C, Hoffman J, Rudolph MJ, Wilson KJ, Soll R, Randle T, Green D, Lewandowski F, Zhang M, Bone R, Spurlino J, DesJarlais R, Deckman I, Molloy CJ, Manthey C, Zhou Z, Sharp C, Maguire D, Crysler C, Grasberger B (2001) Structure-based design, synthesis and SAR of a novel series of thiopheneamidine urokinase plasminogen activator inhibitors. Bioorg Med Chem Lett 11:1379–1382
31. Barber CG, Dickinson RP, Fish PV (2004) Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 3: 1-Isoquinolinylguanidines. Bioorg Med Chem Lett 14:3227–3230
32. Barber CG, Dickinson RP, Horne VA (2002) Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 1: 2-pyridinylguanidines. Bioorg Med Chem Lett 12:181–184
33. Spencer JR, McGee D, Allen D, Katz BA, Luong C, Sendzik M, Squires N, Mackman RL (2002) 4-Aminoarylguanidine and 4-aminobenzamidine derivatives as potent and selective urokinase-type plasminogen activator inhibitors. Bioorg Med Chem Lett 12:2023–2026
34. Wendt MD, Geyer A, McClellan WJ, Rockway TW, Weitzberg M, Zhao X, Mantei R, Stewart K, Nienaber V, Klinghofera V, Giranda VL (2004) Interaction with the S1b-pocket of urokinase: 8-heterocycle substituted and 6,8-disubstituted 2-naphthamidine urokinase inhibitors. Bioorg Med Chem Lett 14:3063–3068
35. Rudolph MJ, Illig CR, Subasinghe NL, Wilson KJ, Hoffman JB, Randle T, Green D, Molloy CJ, Soll RM, Lewandowski F, Zhang M, Bone R, Spurlino JC, Deckman IC, Manthey C, Sharp C, Maguire D, Grasberger BL, DesJarlais RL, Zhou Z (2002) Design and synthesis of 4,5-disubstituted-thiophene-2-amidines as potent urokinase inhibitors. Bioorg Med Chem Lett 12:491–495
36. StOrzebecher J, Vieweg H, Steinmetzer T, Schweinitz A, Stubbs MT, Renatus M, WikstrOm P (1999) 3-Amidinophenylalanine-based inhibitors of urokinase. Bioorg Med Chem Lett 9:3147–3152
37. (2005) CATALYST 4.11 users' manual. Accelrys Software, San Diego
38. Sheridan RP, Kearsley SK (2002) Why do we need so many chemical similarity search methods? Drug Discov Today 7:903–911
39. Sutter J, Güner O, Hoffmann R, Li H, Waldman M (2000) In: Güner OF (ed) Pharmacophore perception, development, and use in drug design. International University Line, La Jolla, pp 501–511
40. Kurogi Y, Güner OF (2001) Pharmacophore modeling and three dimensional database searching for drug design using catalyst. Curr Med Chem 8:1035–1055
41. Poptodorov K, Luu T, Langer T, Hoffmann R (2006) In: Hoffmann RD (ed) Methods and principles in medicinal chemistry. Pharmacophores and pharmacophores searches, vol 2. Wiley-VCH, Weinheim, pp 17–47
42. Li H, Sutter J, Hoffmann R (2000) In: Güner OF (ed) Pharmacophore perception, development, and use in drug design. International University Line, La Jolla, pp 173–189
43. Bersuker IB, Bahçeci S, Boggs JE (2000) In: Güner OF (ed) Pharmacophore perception, development, and use in drug design. International University Line, La Jolla, pp 457–473
44. (2005) CERIUS2 LigandFit user manual (version 4.10). Accelrys, San Diego, pp 3–48
45. Fischer R (1966) The principle of experimentation illustrated by a psychophysical experiment. Hafner, New York, Chapter II
46. (2005) CERIUS2, QSAR users' manual, version 4.10 Accelrys, San Diego, 43–88, 221–235, 237–250
47. Kirchmair J, Markt P, Distinto S, Wolber G, Langer T (2008) 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 22:213–228
48. Irwin JJ, Shoichet BK (2005) ZINC—a free database of commercially available compounds for virtual screening. J Chem Inf Comput Sci 45:177–182
49. Triballeau N, Acher F, Brabet I, Pin J-P, Bertrand H-O (2005) 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 48:2534–2547
50. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD (2002) Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 45:2615–2623
51. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Del Rev 46:3–26
52. uPA Activity Assay kit Cat.No. ECM600. http://www.millipore.com/catalogue/item/ecm600
53. Van Drie JH (2003) Pharmacophore discovery—lessons learned. Curr Pharm 9:1649–1664 Q1
54. Ramsey LF, Schafer WD (1997) The statistical sleuth, 1st edn. Wadsworth, Belmont
55. Verdonk ML, Marcel L, Berdini V, Hartshorn MJ, Mooij WTM, Murray CW, Taylor RD, Watson P (2004) Virtual screening using protein-ligand docking: avoiding artificial enrichment. J Chem Inf Comput Sci 44:793–806
56. Clement OO, Mehl AT (2000) Pharmacophore perception, development, and use in drug design. In: Guner OF (ed) IUL biotechnology series. International University Line, La Jolla, pp 71–84
57. Wendt MD, Rockway TW, Geyer A, McClellan W, Weitzberg M, Zhao X, Mantei R, Nienaber VL, Stewart K, Klinghofer V, Giranda

VL (2004) Identification of novel binding interactions in the development of potent, selective 2-naphthamidine inhibitors of urokinase. Synthesis, structural analysis, and SAR of N-phenyl amide 6-substitution. J Med Chem 47:303–324

58. Vassalli J-D, Belin D (1987) Amiloride selectively inhibits the urokinase-type plasminogen activator. FEBS Lett 214:187–191
59. Roy K, Chakraborty P, Mitra I, Ojha PK, Kar S, Narayan R (2013) Some case studies on application of "rm2" metrics for judging quality of QSAR predictions: emphasis on scaling of response data. J Comput Chem 34:1071–1082
60. Roy K, Mitra I (2011) On various metrics used for validation of predictive QSAR models with applications in virtual screening and focused library design. Comb Chem High Throughput Screen 14: 450–474
61. Roy K, Mitra I, Kar S, Ojha PK, Das RN, Kabir H (2012) Comparative studies on some metrics for external validation of QSPR models. J Chem Inf Model 52:396–408
62. Ojha PK, Mitra I, Das RN, Roy K (2011) Further exploring rm2 metrics for validation of QSPR models dataset. Chemom Intell Lab Syst 107:194–205
63. Shi LM, Fang H, Tomg W, Wu J, Perkins R, Blair RM, Branham WS, Dial SL, Moland CL, Sheenan DM (2001) QSAR models using a large diverse set of estrogens. J Chem Inf Comput Sci 41:186–195
64. Golbraikh A, Tropsha A (2002) Beware of q2. J Mol Graph Model 20(4):269–276
65. Tropsha A (2010) Best practices for QSAR model development, validation, and exploitation. Mol Inf 29:476–488
66. Mita I, Saha A, Roy K (2010) Exploring quantitative structure–activity relationship studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants. Mol Simulat 36: 1067–1079
67. http://aptsoftware.co.in/DTCMLRWeb
68. http://www.aptsoftware.co.in/rmsquare

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES.

Q1. Ref. 53 was not cited anywhere in the text. Please provide a citation. Alternatively, delete the item from the list.

Q2. Ref. 59 was not cited anywhere in the text. Please provide a citation. Alternatively, delete the item from the list.

Q3. Ref. 60 was not cited anywhere in the text. Please provide a citation. Alternatively, delete the item from the list.

Q4. Ref. 67 was not cited anywhere in the text. Please provide a citation. Alternatively, delete the item from the list.

Q5. Ref. 68 was not cited anywhere in the text. Please provide a citation. Alternatively, delete the item from the list.