European Journal of Medicinal Chemistry 45 (2010) 4316–4330

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Discovery of novel CDK1 inhibitors by combining pharmacophore modeling, QSAR analysis and *in silico* screening followed by *in vitro* bioassay

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

Drug Discovery Unit, Department of Pharmaceutical Sciences, Faculty of Pharmacy, University of Jordan, Queen Rania St, Amman, Jordan

ARTICLE INFO

Article history:
Received 1 April 2010
Received in revised form
5 May 2010
Accepted 25 June 2010
Available online 30 June 2010

Keywords:
CDK1
Cancer
Pharmacophore modeling
QSAR
Genetic algorithm
MLR
In silico screening

ABSTRACT

Cyclin-dependent kinase 1 (CDK1) is a valid anticancer target. With this in mind we applied a modeling workflow by combining pharmacophore modeling and QSAR analysis followed by *in silico* screening towards the discovery of novel inhibitory CDK1 scaffolds. Virtual screening identified 10 low micromolar inhibitory leads: 8 from the National Caner Institute (NCI) list of compounds and 2 from our *in house* list of established drugs and agrochemicals. The most potent NCI hit illustrated anti-CDK1 IC_{50} value of 0.83 μM, while the drug hit isoxsuprine illustrated anti-CDK1 IC_{50} value of 2.9 μM and the agrochemical hit foramsulfuran showed $IC_{50} = 3.6$ μM. These results demonstrate that our virtual screening protocol is able to identify novel anti-CDK1 leads for subsequent development into potential anticancer agents.

1. Introduction

Cyclin-dependent kinases (CDK) are key regulators in mammalian cell cycle. Regulation of CDKs occurs through cyclin production, destruction, relocation, inhibitory and activating phoshorylation events. Each cyclin associates with one or two CDKs, and most CDKs associate with one or two cyclins [1].

Although genetic and RNA studies suggest CDK2 and CDK4 may not be essential for cell cycle progression [2,3], similar experiments have shown that CDK1 kinase activity is critical for cell cycle progression through mitosis. In fact, CDK1 inhibitors effectively arrested tumor cell growth prompting great recent interest in the discovery and development of new CDK1 inhibitors [3].

Several potent anti-CDK1 ATP-competitive small organic molecules were shown to have antiproliferative and pro-apoptotic properties. Examples include flavopiridol ($IC_{50} = 0.03$ μM), butyrolactone ($IC_{50} = 0.6$ μM), olomoucine ($IC_{50} = 7.0$ μM) and kenpaullone ($IC_{50} = 0.4$ μM) shown in Fig. 1 [4].

The unavailability of satisfactory high resolution crystallographic structures for CDK1 [4,5] confined modeling-related discovery projects and virtual screening campaigns to either CDK2-based docking studies (due to the availability of high resolution crystallographic structures for this homologous protein [60,61]) or ligand-based approaches, particularly quantitative structure–activity relationship analysis (QSAR) against CDK1 [6,19] or CDK2 [62,63]. Still, successful docking into CDK2 is not necessarily indicative of corresponding binding into CDK1 [64]. Furthermore, despite the excellent predictive potentials of QSAR and three-dimensional (3D) QSAR methodologies (e.g., CoMFA and CoMSIA), they generally lack the ability to act as effective search queries to mine virtual three-dimensional (3D) databases for new hits [7,8].

The continued interest in the development of new CDK1 inhibitors combined with the lack of adequate CDK1 crystallographic structures and adequate computer-aided drug discovery efforts in this area, prompted us to explore the possibility of developing ligand-based three-dimensional (3D) pharmacophore (s) integrated within self-consistent QSAR model for CDK1 inhibitors. The pharmacophore model(s) can be used as 3D search query (ies) to mine 3D libraries for new CDK1 inhibitors, while the QSAR model helps to predict the biological activities of the captured compounds and therefore prioritize them for *in vitro* evaluation [9–17]. We previously reported the use of this innovative approach towards the discovery of new inhibitory leads against glycogen synthase kinase 3β (GSK-3β) [10], dipeptidyl peptidase [11], factor Xa [12], bacterial MurF [13], protein tyrosine phosphatase 1B (PTP 1B) [14], influenza neuraminidase [15], cholesteryl ester transfer protein inhibitors [16], β-secreatase inhibitors [17], hormone sensitive lipase (HSL) [45].

* Corresponding author. Tel.: +962 6 5355000x23305; fax: +962 6 5339649.
E-mail address: mutasem@ju.edu.jo (M.O. Taha).

0223-5234/$ – see front matter © 2010 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2010.06.034

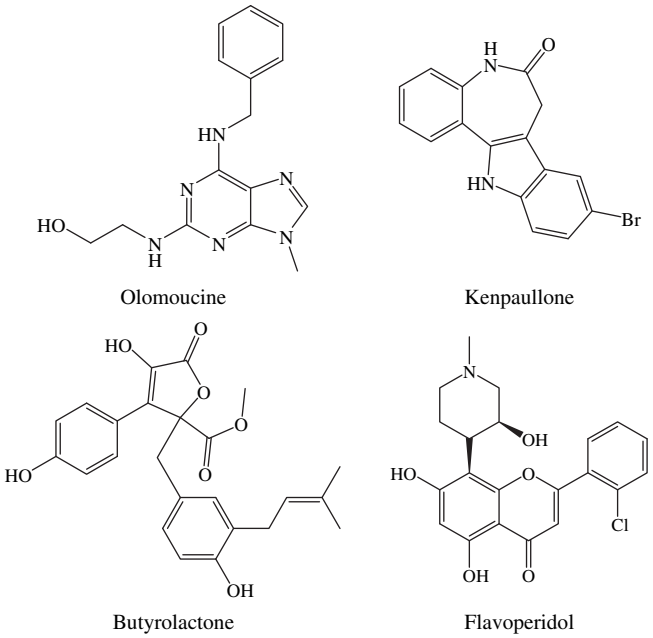


Fig. 1. Chemical structures of known CDK1 inhibitors.

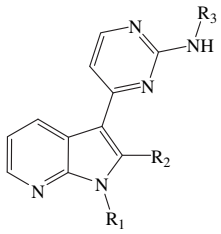
We employed the HYPOGEN module from the CATALYST software package to construct numerous plausible binding hypotheses for CDK1 inhibitors [18]. Subsequently, genetic function algorithm (GFA) and multiple linear regression (MLR) analysis were employed to search for an optimal QSAR that combines high-quality binding pharmacophores with other molecular descriptors and capable of explaining bioactivity variation across a collection of diverse CDK1 inhibitors. The optimal pharmacophores were further validated by evaluating their ability to successfully classify a list of compounds as actives or inactives by assessing their receiver-operating characteristic (ROC) curves. Subsequently, the optimal pharmacophores were complemented with exclusion spheres to enhance their ROC profiles. Thereafter, the resulting refined pharmacophores were used as 3D search queries to screen several available virtual molecular databases for new CDK1 inhibitors [20].

CATALYST models drug–receptor interactions using information derived from the ligand structures [21–28]. HYPOGEN identifies a 3D array of a maximum of five chemical features common to active training ligands that provides relative alignment for each input molecule consistent with binding to a proposed common receptor site. The chemical features considered can be hydrogen-bond donors and acceptors (HBD and HBA), aliphatic and 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 that cover a specified energy range for each input molecule [12,18,23–25,29–32]. CATALYST pharmacophores with exclusion constraints have been used as 3D queries for database searching and in 3D-QSAR studies [32,33].

2. Results and discussion

CATALYST enables automatic pharmacophore construction by using collections 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 it.

Table 1
Chemical structure of CDK1 training set.



Compound	R_1	R_2	R_3	IC_{50} (nM)
1	H	H	Phenyl-	54
2	H	H	2-OH-phenyl-	139
3	H	H	2-CH_3O-phenyl-	44
4[a]	H	H	2F-phenyl-	46
5	H	H	2-CF_3-phenyl-	56
6	H	H	2-CH_3CH_2-phenyl-	68
7	H	H	2-Cl-phenyl-	9
8[a]	H	H	2-Br-phenyl-	13
9	H	H	2-CH_3-phenyl-	3
10	H	H	3-CH_3-phenyl-	57
11	H	H	4-CH_3-phenyl-	40
12	H	H	Cyclohexyl-	14
13	H	H	Cyclohexyl-CH_2–	151
14	H	H	Trans-4-NH_2-cyclohexyl-	2
15	H	H	4-OH(CH_2CH_2)-2-CH_3-phenyl-	0.6
16	H	H	4-[Morpholin-4-yl-(CH_2)$_2$]-2-CH_3-phenyl-	1.9
17[a]	H	H	4-[Piperidin-1-yl-(CH_2)$_2$]-2-CH_3-phenyl-	1
18	H	H	4-[Pyrrolidin-1-yl-(CH_2)$_2$]-2-CH_3-phenyl-	0.6
19	Methyl	H	Cyclohexyl-	10,000
20	Acetyl	H	Cyclohexyl-	31
21[a]	Methanesulfonyl	H	Cyclohexyl-	128
22[a]	2-Dimethylaminoethyl	H	Cyclohexyl-	10,400
23	H	CH_3	Trans-4-Aminocyclohexyl-	280

[a] These compounds were employed as the external test subset in QSAR modeling.

A total of 23 compounds were used in this study (**1**–**23**, Table 1) [34]. Three training subsets were selected from this collection (Table 2). The biological activities of training compounds spanned over 3.5–4.0 orders of magnitude. Genetic algorithm and multiple linear regression statistical analysis were subsequently employed to select an optimal combination of pharmacophore(S) and complementary descriptors capable of explaining bioactivity variations among the collected inhibitors [34].

2.1. Conformational coverage

The 2D structures of the inhibitors were imported into CATALYST and converted automatically into plausible 3D single

Table 2
Training subsets employed in exploring the pharmacophoric space of CDK1 inhibitors, numbers correspond to compounds in Table 1.

Training set	Most active subset[a]	Intermediate subset	Least active subset[b]
I	15, 16, 17, 18	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 10, 13, 23	19, 22
II	9, 14, 16, 17, 18	1, 3, 6, 10, 11, 12, 13, 20, 21	19, 22
III	9, 14, 15, 16, 17, 18	1, 3, 6, 10, 11, 12, 13, 23	19, 22

[a] Potency categories as defined by Eqs. (3) and (4).
[b] Potency categories as defined by Eqs. (3) and (4).

Table 3
Training sets and CATALYST run parameters employed in exploring CDK1 pharmacophoric space.

Run number	Training set[a]	Number of compounds	Selected features[b]	Run parameters[c]
1	**I**	19	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3)	Max–Min: 5–5 Spacing 100
2		19	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3)	Max–Min: 5–5 Spacing 300
3		19	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3)	Max–Min: 4–5 Spacing 100
4		19	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3)	Max–Min: 4–5 Spacing 300
5	**II**	16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1)	Max–Min: 5-5 Spacing 100
6		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1)	Max–Min: 5–5 Spacing 300
7		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1), ExVol (0–10)	Max–Min: 5–5 Spacing 100
8		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1), ExVol (0–10)	Max–Min: 5–5 Spacing 300
9		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1)	Max–Min: 4–5 Spacing 100
10		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1)	Max–Min: 4–5 Spacing 300
11		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1), ExVol (0–10)	Max–Min: 4–5 Spacing 100
12		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1), ExVol (0–10)	Max–Min: 4–5 Spacing 300
13	**III**	16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1)	Max–Min: 5–5 Spacing 100
14		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1)	Max–Min: 5–5 Spacing 300
15		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1), ExVol (0–10)	Max–Min: 5–5 Spacing 100
16		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1), ExVol (0–10)	Max–Min: 5–5 Spacing 300
17		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1)	Max–Min: 4–5 Spacing 100
18		16	Hbic (0–3), HBA (0–3), HBD (0–3), RingArom (0–3), PosIon (0–1)	Max–Min: 4–5 Spacing 300

[a] The letters correspond to training subsets in Table 2.
[b] HBA: Hydrogen-Bond Acceptor, HBD: Hydrogen-Bond Donor, RingArom: Ring Aromatic, Hbic: Hydrophobic, PosIon:PosIonizable, ExVol: Exclusion Volumes. Numbers in brackets refer to the allowed ranges of corresponding features in each pharmacophore modeling run.
[c] Max–Min refer to the allowed range of pharmacophoric features in each model. Spacing refer to the maximum interfeature distance in picometers. Other parameters were set to their default values.

conformer representations. The resulting 3D structures were used as starting points for conformational analysis and in the determination of various molecular descriptors for QSAR modeling.

The conformational space of each inhibitor was comprehensively sampled utilizing the poling algorithm employed within the CONFIRM module of CATALYST and via the "Best" module to ensure extensive sampling of conformational space. Efficient conformational coverage guarantees minimum conformation-related noise during pharmacophore generation and validation stages [35].

2.2. Exploration of CDK1 pharmacophoric space

CATALYST-HYPOGEN enables automatic pharmacophore construction by using a collection of at least 16 molecules with bioactivities spanning over 3.5 orders of magnitude [24–28]. Therefore, the fact that we have a list of 23 CDK1 inhibitors of evenly spread bioactivities over more than 3.5 orders of magnitude prompted us to employ HYPOGEN algorithm to identify as many pharmacophoric binding modes assumed by these inhibitors as possible.

HYPOGEN implements an optimization algorithm that evaluates 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 in Experimental) [24]. The number of evaluated models 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 to exceed 17 (corresponding to 2^{17} hypotheses to be assessed by CATALYST) to guarantee thorough analysis of all models [25].

The extent of the investigated pharmacophoric space is a function of training compounds, selected input chemical features and other CATALYST control parameters [10–17,35]. Restricting the extent of explored pharmacophoric space should improve the efficiency of optimization via allowing effective evaluation of limited number of pharmacophoric models. However, extensive restrictions on the investigated 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 CDK1 inhibitors within reasonable "boundaries" through eighteen HYPOGEN automatic runs and employing three carefully selected training subsets (i.e., from the collected compounds): subsets **I**, **II** and **III** in Table 2. 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 2) 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 Experimental).

Guided by our reasonably restricted pharmacophoric exploration concept, we instructed HYPOGEN to explore only 4- and 5-featured pharmacophores, i.e., ignore models of lesser number of features (Table 3). The later restriction has the advantage of narrowing the investigated pharmacophoric space and representing the feature-rich nature of CDK1 ligands.

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 in Experimental for details about pharmacophore validation) [24–28].

Eventually, 180 pharmacophore models emerged from 18 automatic HYPOGEN runs, out of which only 91 models illustrated confidence levels 85% (Fisher scrambling criteria, see Section 4.1.5 in Experimental) [36]. Successful models were clustered and their best representatives (18 models, see Section 4.1.6 under Experimental) were used in subsequent QSAR modeling [37–39]. Table 4 shows the success criteria of the best representatives. Emergence of several statistically comparable pharmacophore models in the QSAR equation suggests the existence of multiple binding modes accessible for ligands within CDK1 binding pocket. Therefore, it is quite challenging to select any particular pharmacophore hypothesis as a sole representative of the binding process [24–28,40].

Table 4
Success criteria of representative pharmacophoric hypotheses (cluster centres).

RUN[a]	Hypotheses[b]	Pharmacophoric features in generated hypotheses	Config. cost	Total cost	Null hypo	Residual cost[c]	R^2[d]	Global R^2[e]	F[f]	%Cat. scramble
1	2[g]	HBA(F), HBD, 3 × HBic	16.19	85.5	118.5	33.0	0.96	0.46	5.65	95%
	8[g]	**2 × HBD, 3 × HBic**[h]	**16.18**	**87.2**	**118.5**	**31.32**	**0.94**	**0.70**	**20.19**	**95%**
3	6[g]	HBA(F), HBD, HBic, RingArom	17.75	88.8	118.5	29.67	0.95	0.51	7.410	95%
4	10[g]	HBD, 2 × HBic, RingArom	15.36	91.5	118.5	27.02	0.94	0.17	0.646	95%
9	10[g]	2 × HBA(F), HBD, PosIon	17.91	85.8	104.3	18.472	0.86	0.47	6.01	90%
10	3[g]	2 × HBA(F), HBD, PosIon	17.91	85.8	104.3	18.47	0.86	0.60	11.52	90%
11	1[g]	HBA(F), 2 × HBD, HBic, 5 × ExVol	17.91	76.4	104.3	27.91	0.97	0.73	24.13	85%
13	9[g]	HBA(F), HBD, 2 × HBic, PosIon	16.02	85.8	111.8	26.034	0.88	0.52	7.67	90%
14	8[g]	2 × HBD, HBD, 3 × HBic	12.41	82.1	111.8	29.74	0.87	0.55	9.15	90%
	9[g]	HBD, 3 × HBic, PosIon	12.41	82.3	111.8	29.53	0.86	0.30	2.04	90%
15	1[g]	HBA(F), 3 × HBic, PosIon, 7 × ExVol	16.02	75.5	111.8	36.37	0.97	0.52	7.67	95%
17	5[g]	HBD, HBic, 2 × RingArom	17.91	82.8	111.8	29.01	0.91	0.55	8.87	95%
	8[g]	**2 × HBD, HBic, RingArom**[h]	**17.91**	**83.0**	**111.8**	**28.83**	**0.91**	**0.69**	**18.83**	**95%**
18	2[g]	HBA(F), HBD, HBic, RingArom	15.38	80.8	111.8	31.08	0.91	0.46	5.52	95%

[a] Correspond to runs in Table 3.
[b] High ranking representative hypotheses (of their corresponding clusters, see Section 4.1.6).
[c] Difference between total cost and the cost of the corresponding null hypotheses.
[d] Correlation coefficients between pharmacophore-based bioactivity estimates and bioactivities of corresponding training compounds (subsets in Table 2).
[e] Correlation coefficients between pharmacophore-based bioactivity estimates and bioactivities of all collected compounds.
[f] Fisher statistic calculated based on the linear regression between the fit values of all collected inhibitors (**1**–**23**, Table 1) against pharmacophore hypothesis (employing the "best-fit" option and Eq. (6)) and their respective anti-CDK1 bioactivities (log ($1/IC_{50}$) values).
[g] ranking of hypotheses as generated by CATALYST in each automatic runs.
[h] Bolded pharmacophores appeared in the best QSAR equations.

2.3. QSAR modeling

Despite the significance of pharmacophoric hypotheses in understanding ligand-macromolecule affinity and as 3D search queries, their predictive value as 3D-QSAR models is generally limited by steric shielding and bioactivity-modulating auxiliary groups (electron-donating or -withdrawing functionalities) [10–17,27]. This point combined with the fact that our pharmacophore exploration of CDK1 inhibitors identified numerous successful binding hypotheses (as in Table 4) prompted us to employ classical QSAR analysis to search for the best combination of orthogonal pharmacophores and other structural descriptors (connectivity, topological, 2D, etc.) capable of explaining bioactivity variation across the collected list (**1**–**23**, Table 1), i.e., to allow different pharmacophores to compete within QSAR context. Eqs. (1) and (2) show our best-performing QSAR models. Figs. 2 and 3 show the corresponding scatter plots of experimental versus estimated bioactivities for the training and external testing sets.

$$\begin{aligned}\mathrm{Log}(1/\mathrm{IC}_{50}) &= 2.45 - 0.61[2 - \mathrm{AtypeC8}] - 38.76[\mathrm{LUMO} + 0.38] \\ &\quad + 4.82[\mathrm{Hypo1/8} - 9.99]\; r^2_{18} = 0.95, F = 79.72, \\ &\quad n = 18, r^2_{\mathrm{BS}} = 0.945, r^2_{\mathrm{LOO}} = 0.884, r^2_{\mathrm{PRESS}} = 0.864 \end{aligned} \quad (1)$$

$$\begin{aligned}\mathrm{Log}(1/\mathrm{IC50}) &= 1.25 + 2.33[\mathrm{Hypo17/8} - 7.47] - 37.79[\mathrm{LUMO} \\ &\quad + 0.38] + 0.14\mathrm{AtypeH52}\; r^2_{18} = 0.94, F = 72.79, \\ &\quad n = 18, r^2_{\mathrm{BS}} = 0.94, r^2_{\mathrm{LOO}} = 0.889, r^2_{\mathrm{PRESS}} = 0.85 \end{aligned} \quad (2)$$

where, r^2_{18} is the correlation coefficient, F is Fisher statistical parameter, n is the number of observations, r^2_{BS} is the bootstrapping regression coefficient, r^2_{LOO} is the leave-one-out correlation coefficient and r^2_{PRESS} is the predictive r^2 determined for 5 randomly selected external test compounds. AtypeH52 and AtypeC8 are part of the thermodynamic AlogP_atypes family of descriptors and they encode the hydrophobic contributions of certain hydrogen and carbon atoms in LogP [41,42]. LUMO is the energy of the lowest unoccupied molecular orbital calculated employing semiempircal quantum-mechanical methods (MOPAC) [43,44]. Hypo1/8 and Hypo17/8 represent the fit values of the training compounds against the 8th pharmacophores generated

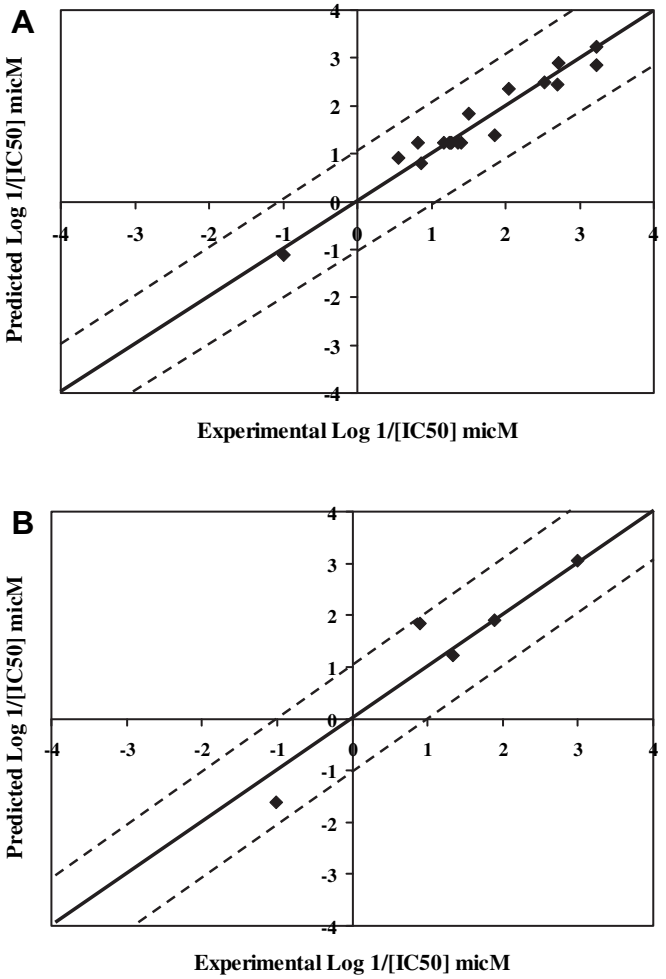


Fig. 2. Experimental versus (A) fitted (18 compounds, $r^2_{\mathrm{LOO}} = 0.884$) and (B) predicted (5 compounds, $r^2_{\mathrm{PRESS}} = 0.864$) bioactivities (expressed in nM concentrations) calculated from QSAR model Eq. (1). The solid lines are the regression lines for the fitted and predicted bioactivities of training and test compounds, respectively, whereas the dotted lines indicate the 1.0 log point error margins.

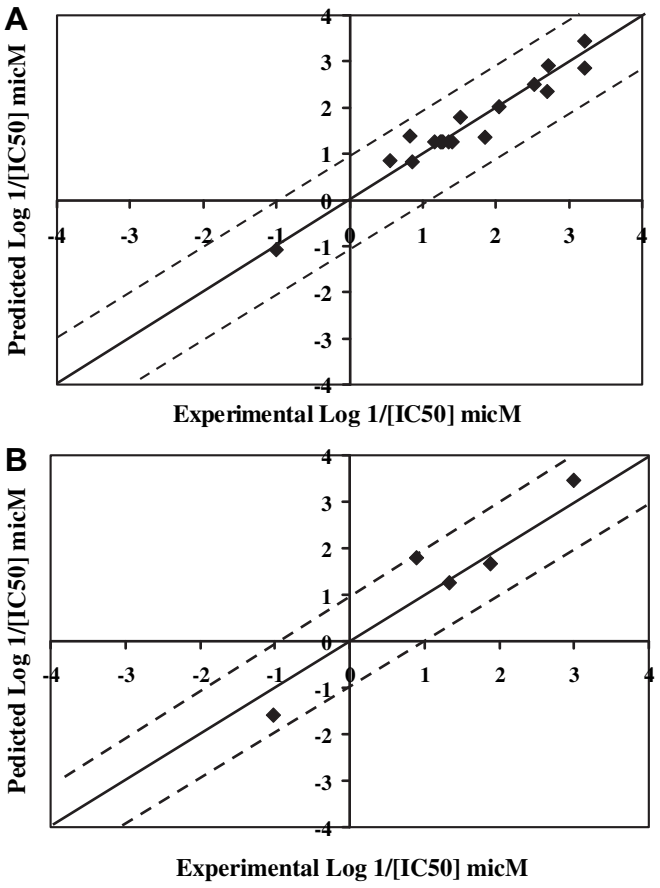


Fig. 3. Experimental versus (A) fitted (18 compounds, $r^2_{LOO} = 0.889$) and (B) predicted (5 compounds, $r^2_{PRESS} = 0.850$) bioactivities (expressed in nM concentrations) calculated from QSAR model Eq. (2). The solid lines are the regression lines for the fitted and predicted bioactivities of training and test compounds, respectively, whereas the dotted lines indicate the 1.0 log point error margins.

in the 1st and 17th modeling runs, respectively (Fig. 4, Tables 4 and 5).

Several descriptors emerged in Eqs. (1) and (2) in spline format. The spline terms employed herein are "truncated power splines" and are denoted by bolded brackets (**[ ]**). For example, $\mathbf{[}f(x) - a\mathbf{]}$ equals zero if the value of $(f(x) - a)$ is negative; otherwise, it equals $(f(x) - a)$ [45].

Emergence of two optimal QSAR models exhibiting two separate pharmacophore models (Hypo1/8 and Hypo17/8) suggests the existence of two binding modes accessible to ligands within the binding pocket of CDK1 [12]. Fig. 4 shows the two pharmacophores and how they map to the most potent training inhibitor **18** (IC_{50} = 0.6 nM, Table 1), while Fig. 5 shows how Hypo1/8 and Hypo17/8 fit two known CDK1 inhibitors outside the training list (i.e., butyrolactone and flavoperidol). Table 5 shows the X, Y, and Z coordinates of the two pharmacophores.

Interestingly, Hypo17/8 and Hypo1/8 emerged in Eqs. (1) and (2) in spline format suggesting that any of these binding modes contributes to ligand/CDK1 affinity only if the particular pharmacophore maps the ligand above the corresponding spline threshold. Accordingly, the ability of certain ligand to fit Hypo1/8 or Hypo17/8 will impact its actual affinity to CDK1 only if it fits the pharmacophore model with values >9.99 or 7.47, respectively. These values represent the corresponding spline intercepts associated with the pharmacophores in Eqs. (1) and (2). The spline cutoffs associated with each pharmacophore resemble fairly high ligand/pharmacophore mapping values. Accordingly, one can conclude that ligand binding to CDK1 is sensitive to slight misalignments among the attracting moieties within the complex, such that lowering the fit value below 9.99 or 7.47, nullifies any affinity gains from mapping the corresponding pharmacophores.

On the other hand, emergence of LUMO in both equation associated with positive spline intercepts and significant negative regression slopes suggests that high molecular electrophilicities (i.e., negative LUMO values < -0.38) nullify any LUMO contributions to bioactivity, while milder molecular electrophilicities (LUMO values > -0.38) profoundly undermine anti-CDK1 bioactivities due to the associated significant negative regression coefficient. Overall, it seems that extreme electrophilicity favors anti-CDK1 bioactivity. Similarly, emergence of two hydrophobic descriptors in the QSAR equations, namely, AtypeC8 and AtypeH52, suggest that molecular hydrophobicity enhances anti-CDK1 properties.

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 constrains necessary to define the size of the binding pocket. This liability renders pharmacophoric models rather promiscuous. Therefore, we decided to complement our QSAR-selected pharmacophores with exclusion spheres employing HipHop-Refine module within CATALYST. Excluded volumes resemble sterically inaccessible regions within the binding site [21].

HipHop-Refine requires a list of training compounds together with two qualitative descriptors that characterize the way by which each training compound contributes in defining the exclusion space, i.e., Principal and MaxOmitFeat (see Section 4.1.8 in Experimental for more details) [46,47]. Table 6 lists the collected compounds together with their HipHop-Refine parameters. Fig. 6C and F shows the sterically-refined versions of Hypo1/8 (50 added exclusion volumes) and Hypo17/8 (51 added exclusion volumes), respectively, and how they map one of the potent hits.

2.5. Receiver-operating characteristic (ROC) curve analysis and exclusion spheres

To further validate the resulting models (both QSAR and pharmacophores), we subjected our QSAR-selected pharmacophores to receiver-operating characteristic (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: overall accuracy, overall specificity, overall true positive rate and overall false negative rate (see Section 4.1.9 under Experimental for more details) [48].

Table 7 and Fig. 7 show the ROC results of our QSAR-selected pharmacophores. Hypo17/8 illustrated poor overall performance with an AUC value of 47.8%. On the other hand, Hypo1/8 exhibited mediocre performance with AUC value of 77.8%. However, the sterically-refined versions of both models illustrated much better ROC profiles. The fact that exclusion spheres resemble sterically forbidden regions in the binding pocket should sharpen the discriminatory power of the pharmacophore models, as seen in Table 7 and Fig. 7. The ROC results of the sterically-refined versions of Hypo1/8 and Hypo17/8 improved significantly compared to their unrefined versions as reflected by their ROC–AUCs, which shifted from 77.8% and 47.8% to 95.5% and 73.2% for Hypo17/8 and Hypo1/8, respectively.

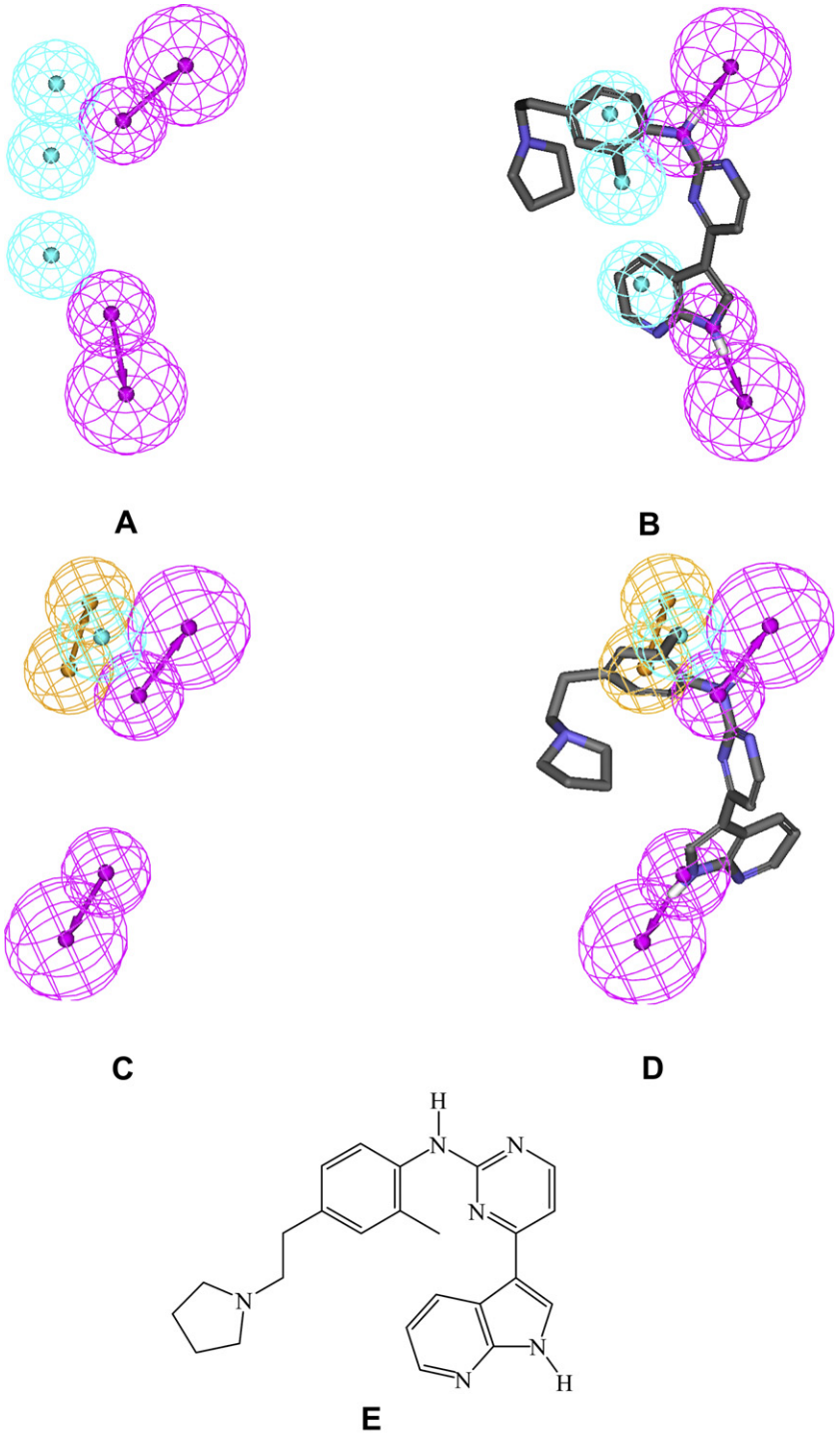


Fig. 4. (A) Pharmacophoric features of Hypo1/8: HBD as pink vectored spheres, Hbic as blue spheres, (B) Hypo1/8 mapped against training compound 18 (IC50 = 0.6 nM, Table 1), (C) Pharmacophoric features of Hypo17/8: HBD as pink vectored spheres, Hbic as blue spheres and RingArom as orange vectored spheres, (D) Hypo17/8 mapped against training compound 18, (E) Chemical structure of training compound **18** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

2.6. *In silico screening of databases*

We employed the sterically-refined versions of Hypo1/8 and Hypo17/8 as 3D search queries against the national cancer institute list of compounds (NCI, includes 238,819 compounds) and our *in house* built list of established drugs and agrochemicals (DAC, 3002 compounds). Hits are defined as those compounds that have their chemical groups spatially overlap with corresponding features within the particular pharmacophoric model.

Table 8 shows the number captured hits by each pharmacophore model. NCI hits were subsequently filtered by Lipinski's [49] and Veber's criteria [50]. However, DAC hits were left without subsequent filtration. Surviving hits were then fitted against Hypo1/8 or Hypo17/8 (fit values determined by Eq. (6)) and their fit

Table 5
Three-dimensional properties of QSAR-selected pharmacophores.

Model	definitions		Chemical features						
			HBD		HBD		Hbic	Hbic	Hbic
Hypo 1/8[a]	Weights		2.080		2.080		2.080	2.080	2.080
	Tolerances		1.60	2.20	1.60	2.20	1.60	1.60	1.60
	Coordinate	X	−1.958	−4.680	1.246	2.160	1.080	−1.300	1.200
		Y	−1.826	−2.340	−4.275	−5.880	−1.200	0.900	−1.200
		Z	−2.027	−3.260	3.956	6.320	−2.080	−2.440	3.200
			HBD		HBD		RingArom		Hbic
Hypo 17/8[b]	Weights		2.120		2.120		2.120		2.120
	Tolerances		1.60	2.20	1.60	2.20	1.60	1.60	1.60
	Coordinate	X	−5.307	−2.664	−9.151	−11.482	−7.205	−5.593	−5.599
		Y	2.545	3.915	−3.355	−5.195	3.958	4.899	5.078
		Z	−1.569	−1.945	−2.191	−2.612	−3.064	−5.413	−1.847

[a] Hypo1/8: the 8th pharmacophore hypothesis generated in the 1st HYPOGEN run (Table 4).
[b] Hypo17/8: the 8th pharmacophore hypothesis generated in the 17th HYPOGEN run (Table 4).

values were substituted in corresponding QSAR Eqs (1) or (2) to determine their predicted bioactivities. However, in order to minimize the impact of any possible extrapolatory QSAR prediction errors on decisions regarding which hits merit subsequent *in vitro* testing [7,44], bioactivity predictions were employed merely to rank the corresponding hits and to prioritize subsequent *in vitro* testing [11–17]. Out of the 88 highest-ranking hits acquired for experimental validation, 10 were found to possess significant inhibitory activities against CDK1 with IC_{50} ranging from 0.55 to 4.9 μM Fig. 8 shows the structures of the most potent hits, while Table 9 summarizes their bioactivities. Table A in Supplementary Materials lists the bioactivites of the remaining hits. The apparent discrepancies between predicted and experimental bioactivities of hit compounds are probably related to the limited number and diversity of training compounds (18), which should restrict the extrapolatory potential of the QSAR models.

Potent hits included 8 compounds from the NCI list and two from DAC. The most potent NCI hit (**25**) illustrated IC_{50} value of

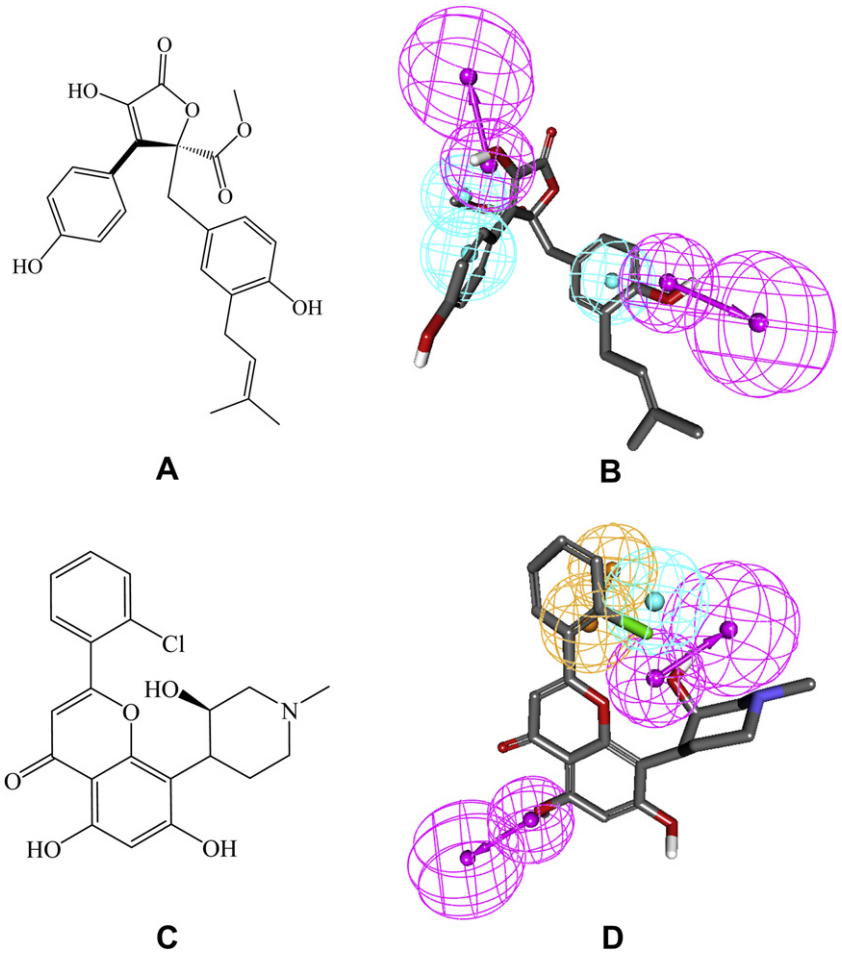


Fig. 5. (A) Chemical structure of butyrolactone (IC_{50} = 0.6 μM), (B) Hypo1/8 mapped against butyrolactone, (C) Chemical structure of flavoperidol (IC_{50} = 0.03 μM), (D) Hypo17/8 mapped against flavoperidol.

Table 6
The training compounds used for adding excluded spheres to Hypo1/8 and Hypo17/8 through HIPHOP-REFINE and their corresponding run parameters.

Compound[a]	Principal value	MaxOmitFeat[b]
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	1	1
8	1	0
9	2	0
10	1	1
11	1	1
12	1	1
13	1	1
14	2	0
15	2	0
16	2	1
17	2	1
18	2	0
19	0	2
20	1	1
21	1	1
22	0	2
23	1	1

[a] Compounds' numbers as in Table 1.
[b] MaxOmitFeat: Maximum omitted features.

0.83 μM, while the drug hit isoxsuprine (**33**) yielded IC_{50} value of 2.9 μM and the herbicide hit foramsulfuran (**34**) gave anti-CDK1 IC_{50} of 3.6 μM. The aminoquinolines **25** and **31** are close to known kinase inhibitors (i.e., Pim-1 kinase [65] and CSF-1R kinase [66]). In fact, other aminoquinolines were also reported to possess potent anti-CDK1 activities [67]. However, the remaining hits seem to represent completely novel scaffolds in the area of CDK1 inhibition.

Fig. 6 shows how the potent hit **25** (IC_{50} = 0.83 μM) maps both Hypo1/8 and Hypo17/8. Clearly from the figure, both the quinolino 4-amino substitutent and the terminal hydroxyl group of **25** are critical for anti-CDK1 bioactivity as they map the two HBDs in both pharmacophore models. A similar conclusion can be drawn for the quinoline ring, i.e., required for mapping two Hbics in Hypo1/8, RingArom and Hbic in Hypo17/8. However, the ethylamine side chain of **25** is only mapped by an Hbic feature in Hypo1/8. Figure A in Supplementary Data shows how Hypo1/8 and Hypo17/8 fit the remaining hits.

3. Conclusions

This work includes elaborate pharmacophore exploration of CDK1 inhibitors utilizing CATALYST-HYPOGEN. QSAR analysis was employed to select the best combination of molecular descriptors and pharmacophore models capable of explaining bioactivity variation across an informative list of training compounds. Two novel CDK1 inhibitory pharmacophore models emerged in the optimal QSAR equation. Successful pharmacophores were complemented with exclusion spheres to optimize their receiver-operating characteristic (ROC) curve profiles. Optimal sterically-refined models were used as 3D search queries to screen the NCI, drugs and agrochemicals databases for new CDK1 inhibitors. Ten compounds, out of 88 tested, illustrated low micromolar potencies against CDK1, which equal an overall enrichment rate of 11.4%. The results suggest that the combination of pharmacophore modeling and QSAR analysis can be a useful tool for finding potential anti-CDK1 agents.

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), USA.
- CERIUS2 (Version 4.10), Accelrys Inc. (www.accelrys.com), USA.
- CS ChemDraw Ultra 7.01, Cambridge Soft Corp. (http://www.cambridgesoft.com), USA.

Pharmacophore modeling and QSAR analysis 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 600 MHz MIPS R14000 processor (1.0 GB RAM) running the Irix 6.5 operating system.

4.1.2. Data set

The structures of 23 CDK1 inhibitors (Table 1) were collected from published literature [34]. The *in vitro* bioactivities of the collected inhibitors were expressed as the concentration of the test compound that inhibited the activity of CDK1 by 50% (IC_{50}). The logarithm of measured IC_{50} (μM) values was used in pharmacophore modeling and QSAR analysis, thus correlating the data linear to the free energy change. The logarithmic transformation of IC_{50} values should minimize any potential errors resulting from this assumption [16].

4.1.3. Conformational analysis

The conformational space of each inhibitor (**1**–**23**, Table 1) was explored adopting the "best conformer generation" option within CATALYST which is based on the generalized CHARMm force field implemented in the program. Conformational ensembles were generated with an energy threshold of 20 kcal/mol from the local minimized structure and a maximum limit of 250 conformers per molecule [18,21,22,24–28].

4.1.4. Generation of pharmacophoric hypotheses

All 23 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 [24–28]. Three structurally diverse training subsets (Table 2) were carefully selected from the collection for pharmacophore modeling. Training subsets **I**, **II** and **III** were utilized to conduct four, eight and six modeling runs, respectively, to explore the pharmacophoric space of CDK1 inhibitors. Different hypotheses were generated by altering the interfeature spacing and number of allowed features in the resulting pharmacophores (Table 3).

Pharmacophore modeling employing CATALYST proceeds through three successive phases: the constructive phase, subtractive phase and optimization phase [24–28]. During the constructive phase, CATALYST generates common conformational alignments among the most active training compounds. Only molecular alignments based on a maximum of five chemical features are considered. The program identifies a particular compound as being within the most active category if it satisfies Eq. (3) [24–28].

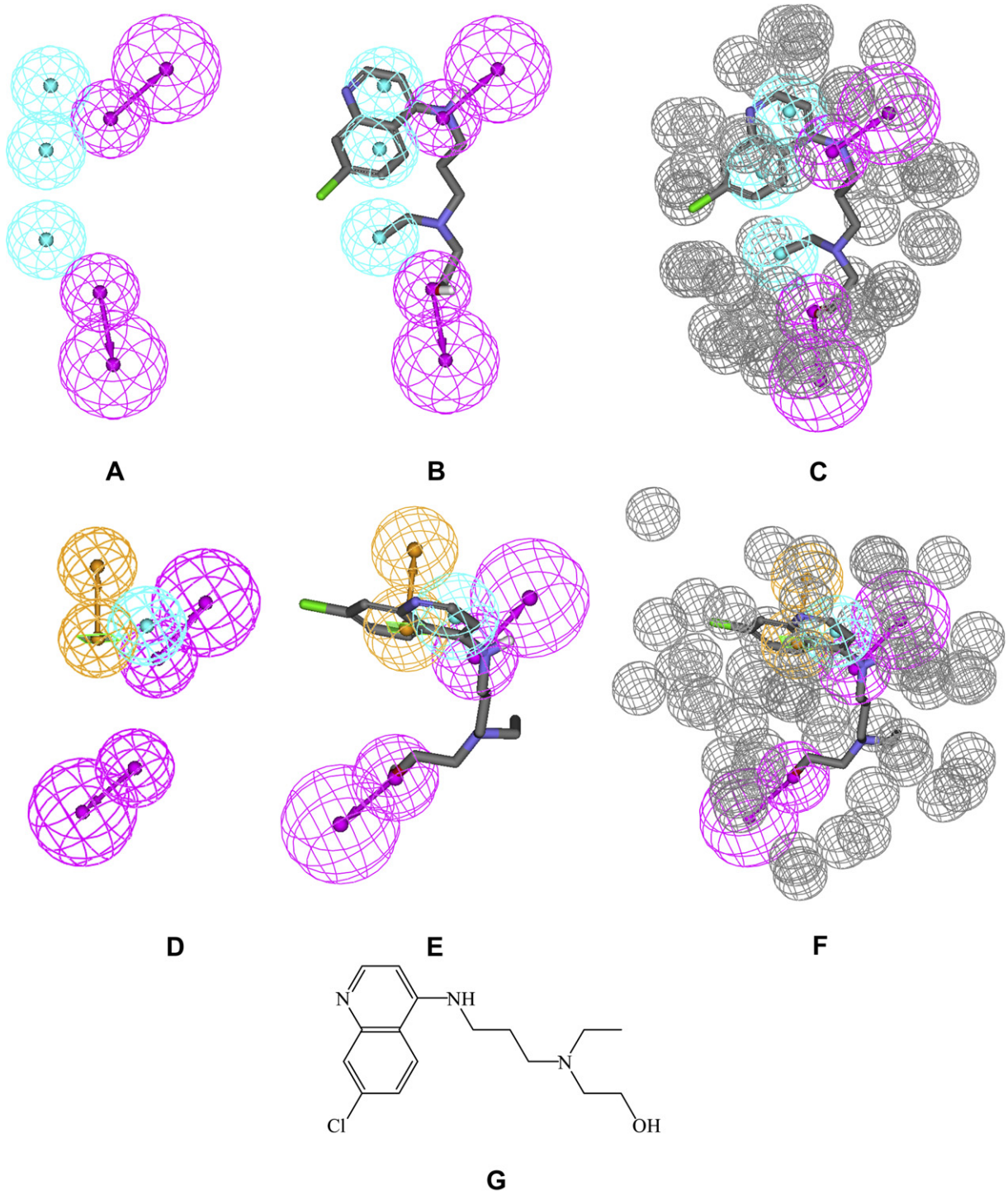


Fig. 6. (A) Hypo1/8, (B) Hypo1/8 mapped against hit compound **25** (IC50 = 0.83 mM, Table 9), (C) Sterically-refined Hypo1/8 mapped against **25**, (D) Hypo17/8, (E) Hypo17/8 mapped against **25**, (F) Sterically-refined Hypo17/8 mapped against **25**, and (G) Chemical structure of **25**.

$$(\mathrm{MAct} \times \mathrm{UncMAct}) - (\mathrm{Act}/\mathrm{UncAct}))0.0 \quad (3)$$

where "MAct" is the activity of the most active compound in the training set, "Unc" is the uncertainty of the compounds and "Act" is the activity of the training compounds under question. In the subsequent subtractive phase, CATALYST eliminates some hypotheses that fit inactive training compounds. A particular training compound is defined as being inactive if it satisfies Eq. (4) [24–28]:

$$\mathrm{Log}(\mathrm{Act}) - \log(\mathrm{MAct}) > \mathrm{BS} \quad (4)$$

where, "BS" is the bioactivity spread (equals 3.5 by default).

In the optimization phase, CATALYST applies fine perturbations in the form of vectored feature rotation, adding new feature and/or removing a feature, to selected hypotheses that survived the subtractive phase to find new models of enhanced bioactivity-to-mapping correlations. Eventually, CATALYST selects the highest-ranking models (10 by default) and presents them as the optimal

Table 7
Performance of QSAR-selected pharmacophores and their sterically-refined versions as 3D search queries.

Pharmacophore Model	ROC[a]–AUC[b]	ACC[c]	SPC[d]	TPR[e]	FNR[f]
Hypo1/8	0.778	0.932	0.986	0.192	0.0139
Hypo17/8	0.478	0.933	0.952	0.663	0.048
Refined Hypo1/8	0.955	0.932	0.993	0.096	0.007
Refined Hypo17/8	0.732	0.930	0.980	0.215	0.015

[a] ROC: receiver-operating characteristic curve.
[b] AUC: area under the curve.
[c] ACC: overall accuracy.
[d] SPC: overall specificity.
[e] TPR: overall true positive rate.
[f] FNR: overall false negative rate.

Table 8
Numbers of captured hit compounds by sterically-refined Hypo1/8 and Hypo17/8.

3D Database	Post-screening filtering[b]	Number of captured hits from pharmacophore models	
		Sterically-refined Hypo1/8	Sterically-refined Hypo17/8
NCI[a]	Before	2895	6083
	After	1953	4308
DAC[c]		14	22

DAC: the list of established drugs and agrochemicals (3002 structures).
[a] NCI: national cancer institute list of available compounds (238,819 structures).
[b] Post-screening filtering employing Lipinski's and Veber's rules. Two Lipinsk's violations were tolerated.
[c] This list of compounds was virtually screened without post-screening filtering.

pharmacophore hypotheses resulting from the particular automatic modeling run [21].

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 [21,24–28]. Weight cost is a value that increases as the feature weight in a model deviates from an ideal value of 2. The deviation between the estimated activities of the training set and their experimentally determined values adds to the error cost [21,24–28]. The activity of any compound can be estimated from a particular hypothesis through Eq. (5) [21].

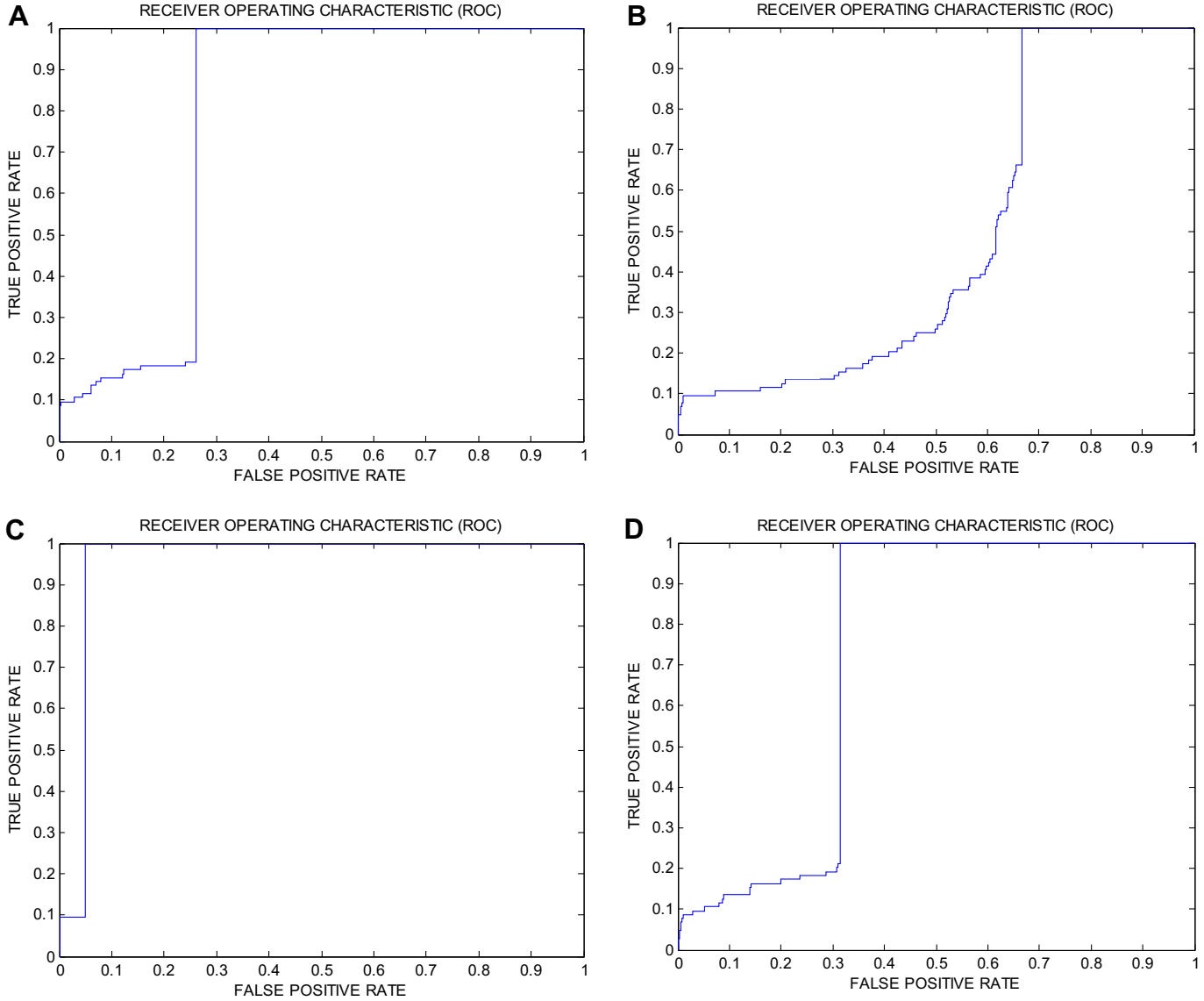


Fig. 7. ROC curves of: (A) Hypo1/8, (B) Hypo17/8, (C) Sterically-refined Hypo1/8, and (D) Sterically-refined Hypo17/8.

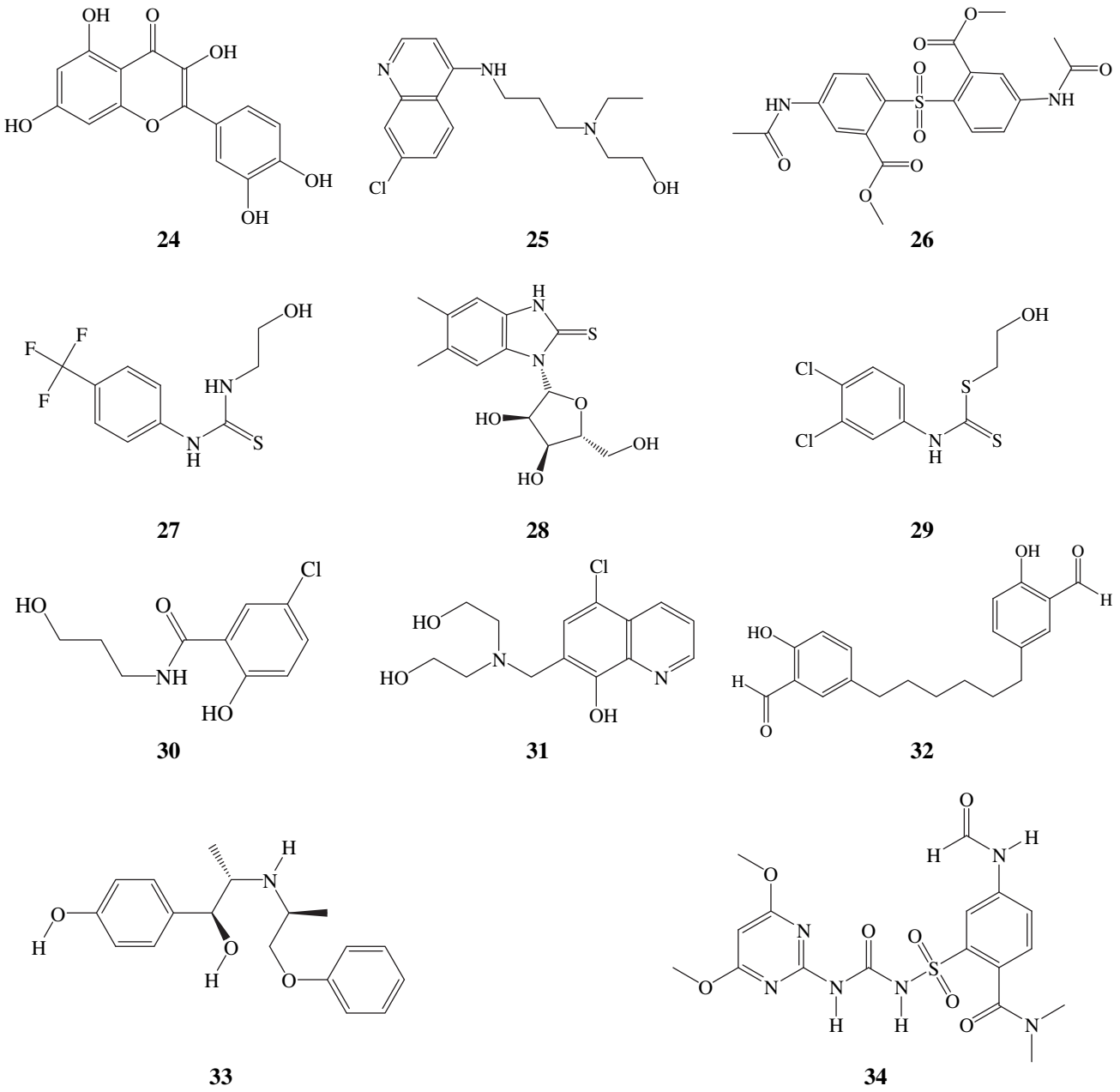


Fig. 8. Chemical structure of active compounds as CDK1 inhibitors.

$$\text{Log(Estimated activity)} = I + \text{Fit} \quad (5)$$

where, I = the intercept of the regression line obtained by plotting the log of the biological activity of the training set compounds against the Fit values of the training compounds. The Fit value for any compound is obtained automatically employing Eq. (6) [24–26].

$$\text{Fit} = \Sigma\,\text{mapped hypothesis features} \times W\left[1 - \Sigma(\text{disp/tol})^2\right] \quad (6)$$

where, Σ mapped hypothesis features represents the number of pharmacophore features that successfully superimpose (i.e., map or overlap with) corresponding chemical moieties within the fitted compound, W is the weight of the corresponding hypothesis feature spheres. This value is fixed to 1.0 in CATALYST-generated models. disp is the distance between the center of a particular pharmacophoric sphere (feature centroid) and the center of the corresponding superimposed chemical moiety of the fitted compound; tol is the radius of the pharmacophoric feature sphere (known as Tolerance, equals to 1.6 Å by default). Σ $(\text{disp/tol})^2$ is the summation of $(\text{disp/tol})^2$ values for all pharmacophoric features that successfully superimpose corresponding chemical functionalities in the fitted compound [24–26].

The third term, i.e., the configuration cost, penalizes the complexity of the hypothesis, i.e., the configuration cost. This is a fixed cost, which is equal to the entropy of the hypothesis space. The more the numbers of features (a maximum of five) in a generated hypothesis, the higher is the entropy with subsequent increase in this cost. The overall cost (total cost) of a hypothesis is calculated by summing over the three cost factors. However, error cost is the main contributor to total cost [10–17].

Table 9
High ranking hit molecules with their fit values, corresponding QSAR estimates based on Eqs. (1) and (2), and their corresponding *in vitro* anti-CDK1 bioactivities.

Hit	Name[a]	Fit values against[b]		QSAR estimated IC_{50} (μM)[c]		*In vitro* IC_{50} (μM)[d]
		Hypo1/8	Hypo17/8	Eq. (1)	Eq. (2)	
24	Quercetin[e]	–	–	–		0.45 (0.99)[f]
25	NCI0010488	9.3	7.4	0.059	0.022	0.83 (0.91)
26	NCI0141089	3.5	4.0	0.059	0.056	0.91 (0.99)
27	NCI0201560	–	5.5	–	0.056	1.40 (0.99)
28	NCI0122331	–	6.1	–	0.056	2.10 (0.83)
29	NCI0306919	–	6.5	–	0.056	2.50 (0.85)
30	NCI0058329	–	7.5	–	0.188	2.53 (0.93)
31	NCI0130803	2.4	6.7	0.059	0.056	3.23 (0.80)
32	NCI0088001	6.6	–	0.160	–	4.90 (0.84)
33	Isoxsuprine	–	5.9	–	0.578	2.90 (0.98)
34	Foramsulfuran	5.0	5.7	0.46	0.664	3.60 (0.99)

[a] Hits shown in Fig. 8.
[b] Best-fit values against each binding hypothesis calculated by Eq. (6).
[c] QSAR estimates from Eqs. (1) and (2).
[d] *In vitro* enzyme inhibition. Each value represents the average of at least three measurements.
[e] Standard Inhibitor with reported IC_{50} value range of 0.45–2.26 μM [58,59].
[f] This value represents the correlation coefficient of the corresponding dose-response line over three log cycles of inhibitor concentrations.

CATALYST 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 4), the more likely that the hypothesis does not reflect a chance correlation [21,22,24–28]. In a successful automatic modeling run, CATALYST ranks the generated models according to their total costs [24–26].

An additional approach to assess the quality of CATALYST-HYPOGEN pharmacophores is to cross-validate them using the Cat-Scramble algorithm implemented in CATALYST [18,27]. This validation procedure is based on Fisher's randomization test [36]. In this validation test, we selected a 95% confidence level, which instruct 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 [46]. 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 [28]. Based on Fisher randomization criteria, only 91 pharmacophores achieved ≥85% significance threshold, and were therefore, submitted for subsequent processing (clustering and QSAR analysis).

4.1.6. Clustering of the generated pharmacophore hypotheses

The resulting models (91) were clustered into 18 groups utilizing the hierarchical average linkage method available in CATALYST. Subsequently, the highest-ranking representatives, as judged based on their fit-to-bioactivity correlation values, were selected to represent their corresponding clusters in subsequent QSAR modeling. Table 4 shows the representative pharmacophores features, success criteria, the corresponding Cat.scramble confidence levels and differences from corresponding null hypotheses [11–17].

4.1.7. QSAR modeling

A set of 18 compounds was employed as the training set for QSAR modeling. The remaining 5 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: the 23 inhibitors were ranked according to their IC_{50} values, then every fifth compound was selected for the test set starting from the high-potency end (Table 1).

The chemical structures of the inhibitors were imported into CERIUS2 as standard 3D single conformers (of the lowest energy within the conformational ensemble generated by CATALYST) representations in SD format. Subsequently, 100 molecular descriptors were calculated for each compound employing the C2. DESCRIPTOR module of CERIUS2. The calculated descriptors included various simple and valence connectivity indices, electro-topological state indices, single point quantum-mechanical descriptors (via the AM1 model) and other molecular descriptors [38]. Furthermore, the training compounds were fitted against the 18 representative CDK1 pharmacophore hypotheses, generated by the CATALYST-HYPOGEN automatic runs (shown in Table 4), and their fit values (produced by the best-fit command within CATALYST via Eq. (6)) were added as additional molecular descriptors [18].

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, i.e., multiple linear regression modeling (MLR) [11–13].

Our preliminary diagnostic trials suggested the following optimal GFA parameters: Explore linear equations at mating and mutation probabilities of 50%; population size = 500; number of genetic iterations = 30 000 and 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 3 to 6 explanatory terms. All QSAR models were validated employing leave-one-out cross-validation (r^2_{LOO}), bootstrapping (r^2_{BS}) [12] and predictive r^2 (r^2_{PRESS}) calculated from the test subset. The predictive r^2_{PRESS} is defined as Eq. (7):

$$r^2_{PRESS} = SD - PRESS/SD \quad (7)$$

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, PRESS is the squared deviations between predicted and actual activity values for every molecule in the test set [12].

4.1.8. Addition of exclusion volumes

To account for the steric constrains of the binding pocket we decided to decorate Hypo1/8, and Hypo17/8 with exclusion volumes employing HipHop-Refine module of Catalyst. HipHop-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 [11–17].

All 23 collected compounds were used for constructing appropriate exclusion regions around Hypo1/8 and Hypo17/8. In HipHop-Refine the user defines how many molecules must map completely or partially to the hypothesis via the principal and maximum omitted features (MaxOmitFeat) parameters. Active compounds are normally assigned MaxOmitFeat parameter of zero and principal value of 2 to instruct the software to consider all their chemical moieties in modeling and to fit them against all 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 [11–17].

We decided to consider 10 μM as an appropriate activity/inactivity threshold. Accordingly, inhibitors of IC_{50} values ≤10 μM were regarded as "actives" and were assigned principal and MaxOmitFeat values of 2 (or 1) and 0 (or 1), respectively, On the other hand, inhibitors of $IC_{50} \geq 10$ μM were considered inactive and were

assigned principal values of zero and MaxOmitFeat parameter of 2. HipHop-Refine was configured to allow a maximum of 100 exclusion spheres to be added to the generated pharmacophoric hypotheses. Table 6 shows the training compounds in this step and their corresponding principal and MaxOmitFeat parameters [32].

4.1.9. Receiver-operating characteristic (ROC) curve analysis

QSAR-selected pharmacophore models (i.e., Hypo1/8 and Hypo17/8) were validated by assessing their abilities to selectively capture diverse CDK1 active compounds from a large testing list of actives and decoys.

The testing list was prepared as described by Verdonk and coworkers [51,52]. 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) and (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. (8)):

$$D(i,j) = \sqrt{(\mathrm{NumHBD}_i - \mathrm{NumHBD}_j)^2 + (\mathrm{NumHBA}_i - \mathrm{NumHBA}_j)^2 + (\mathrm{NP}_i - \mathrm{NP}_j)^2} \quad (8)$$

The minimum distances are then averaged over all active compounds ($D_{\min}$). Subsequently, for each active compound in the test set, around 37 decoys were chosen from the ZINC database [53]. The decoys were selected in such a way that they did not exceed $D_{\min}$ 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 ROC test set. Active testing compounds were defined as those possessing CDK1 affinities ranging from 0.0006 to 2.0 μM. The test set included 26 active compounds and 965 ZINC decoys.

The test set (991 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. (6).

The ROC curve analysis describes the sensitivity (Se or true positive rate, Eq. (9)) 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. (10)) [54].

$$Se = \frac{\text{Number of Selected Actives}}{\text{Total Number of Actives}} = \frac{TP}{TP+FN} \quad (9)$$

$$Sp = \frac{\text{Number of Discarded Inactives}}{\text{Total Number of Inactives}} = \frac{TN}{TN+FP} \quad (10)$$

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), while FP is the number of captured decoys (presumably inactives).

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 [16,54,55]. 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 [16]. A ROC curve is plotted by setting the score of the active molecule as the first threshold. Afterwards, 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 [16].

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 towards the Se = 1 − Sp line asymptotically with increasing number of actives and decoys [16]. The success of a particular virtual screening workflow can be judged from the following criteria (Table 7):

1) Area under the ROC curve (AUC) [16]. 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 [16,54,55].
2) Overall Accuracy (ACC): describes the percentage of correctly classified molecules by the screening protocol (Eq. (11)). Testing compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) [54].

$$\mathrm{ACC} = \frac{TP+TN}{N} = \frac{A}{N}\cdot Se + \left(1-\frac{A}{N}\right)\cdot Sp \quad (11)$$

where, N is the total number of compounds in the testing database, 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.
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 to their individual fit values.
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 to their individual fit values [16].

4.1.10. In silico screening of databases for new CDK1 inhibitors

Sterically-refined Hypo1/8 and Hypo17/8 were employed as 3D search queries against the NCI and DAC libraries using the "Best Flexible Database Search" option implemented within Catalyst. NCI hits were filtered based on Lipinski's and Veber's rules [49,50]. The remaining hits were fitted against Hypo1/8 and Hypo17/8 using the "best-fit" approach implemented within CATALYST. Subsequently, the fit values together with the relevant molecular descriptors were substituted in QSAR Eqs. (1) and (2) to predict anti-CDK1 IC_{50} values. The highest-ranking 88 hits were subsequently tested *in vitro*.

4.2. CDK1 inhibition assay

In vitro assay was conducted employing Omnia® Ser/Thr peptide 7 and CDK1/CycB kit (BioVision, Linda Vista Avenue, USA). The kit performs kinase activity assays employing recombinant human CDK1/CycB kinase. The kinase assay uses 8-hydroxy-5-(*N*,*N*-demethylsulfonamido)-2-methylquinoline (SOX) incorporated into the substrate peptide. Upon phosphorylation of the peptide by CDK1, Mg^{++} is chelated to form a bridge between the SOX moiety and the phosphate group that is added to the serine, threonine, or tyrosine residues within the peptide, resulting in an increase in fluorescence when the kinase reaction mixture is excited at 360 nm. Emission is measured at 485 nm [56,57].

4.2.1. Quantification of CDK1 activity

The assay procedure can be described briefly as follows. An aliquot of 15 μL of CDK1 (1 μg/ml) was added to 5 μL of testing sample. Then 20 μL of the kit's master mix (includes donor molecule, acceptor molecule and assay buffer) was added, mixed well, and completed to 40 μL with the provided assay buffer. After incubation at 30 °C for 5 min, fluorescence intensity (Excitation λ: 360 nm; Emission λ: 485 nm) was read in a FLX800TBI Microplate Fluorimeter (BioTek Instruments, Winooski, USA) every 30 min for 2 h. The percentage of inhibition was calculated as in Eq. (12).

$$\%\text{Inhibition} = \left(1.0 - \frac{RF_{\text{Inhibitor}}}{RF_{\text{Enzyme}}}\right) \times 100\% \quad (12)$$

where, $RF_{\text{Inhibitor}}$ and RF_{Enzyme} are the rates of fluorescence increase due to CDK1 kinase reaction in the presence and absence of a particular inhibitor, respectively. IC_{50} was determined from three inhibitor concentrations: 0.1, 1.0 and 10 μM [12].

4.2.2. Preparation of tested compounds

Commercially available compounds, i.e., **33** (Isoxsuprine) and **34** (Foramsulfuran), were purchased from Sigma–Aldrich and were of HPLC-based purity >98%. NCI hits **25** and **26** illustrated >95% purity based on combustion analysis. Tested compounds were initially dissolved in DMSO to yield 10 mM stock solutions and subsequently diluted to the required concentrations using distilled deionized water. The final concentration of DMSO was adjusted to 0.1%. The percentage of residual activity of CDK1 was determined for each compound by comparing the activity of CDK1 in the presence and absence of the tested compound. Positive controls were tested to assess the degree of CDK1 inhibition by 0.1% DMSO. CDK1 was not affected by DMSO. Negative controls as kinase buffer were used as background [12].

Acknowledgments

This project was partially sponsored by the Faculty of Graduate Studies (Ph.D. Thesis of Mahmoud A.Al-Sha'er). The authors wish to thank the Deanship of Scientific Research and Hamdi-MangoCenter for Scientific Research at the University of Jordan for their generous funds. The authors are also indebted to national cancer institute for freely providing hit molecules for evaluation.

Appendix. Supplementary data

The supplementary data associated with this article can be found in the on-line version at doi:10.1016/j.ejmech.2010.06.034.

References

[1] A.W. Murray, Recycling the cell cycle: cyclins revisited, Cell 116 (2004) 221–234.
[2] S.S. Mader, The Cell, Chapter One, Biology, fourth ed. WCB Publisher, New York, 1993.
[3] S. Chen, L. Chen, N.T. Le, C. Zhao, A. Sidduri, J.P. Lou, C. Michoud, L. Portland, N. Jackson, J.-J. Liu, F. Konzelmann, F. Chi, C. Tovar, Q. Xiang, Y. Chen, Y. Wen, L. T. Vassilevet, Synthesis and activity of quinolinyl-methylene-thiazolinones as potent and selective cyclin-dependent kinase 1 inhibitors, Bioorg. Med. Chem. Lett. 17 (2007) 2134–2138.
[4] K.S. Kim, J.S. Sack, J.S. Tokarski, L. Qian, S.T. Chao, L. Leith, Y.F. Kelly, R.N. Misra, J.T. Hunt, S.D. Kimball, W.G. Humphreys, B.S. Wautlet, J.G. Mulheron, K. R. Webster, Thio- and oxoflavopiridols, cyclin-dependent kinase 1-selective inhibitors: synthesis and biological effects, J. Med. Chem. 43 (2000) 4126–4134.
[5] C.E. Arris, F.T. Boyle, A.H. Calvert, N.J. Curtin, J.A. Endicott, E.F. Garman, A. E. Gibson, B.T. Golding, S. Grant, R.J. Griffin, P. Jewsbury, L.N. Johnson, A. M. Lawrie, D.R. Newell, M.E.M. Noble, E.A. Sausville, R. Schultz, W. Yu, Identification of novel purine and pyrimidine cyclin-dependent kinase inhibitors with distinct molecular interactions and tumor cell growth inhibition profiles, J. Med. Chem. 43 (2000) 2797–2804.
[6] C. Kunick, K. Lauenroth, K. Wieking, X. Xie, C. Schultz, R. Gussio, D. Zaharevitz, M. Leost, L. Meijer, A. Weber, F.S. Jørgensen, T. Lemcke, Evaluation and comparison of 3D-QSAR CoMSIA models for CDK1, CDK5, and GSK-3 inhibition by paullones, J. Med. Chem. 47 (2004) 22–36.
[7] M.T.D. Cronin, T.W. Schultz, Pitfalls in QSAR, J. Mol. Struct. 622 (2003) 39–51 (Theochem).
[8] M. Akamatsu, Current state and perspectives of 3D-QSAR, Curr. Top. Med. Chem. 12 (2002) 1381–1394.
[9] P.W. Sprague, K. Müller (Eds.), Perspectives in Drug Discovery and Design, Automated Chemical Hypothesis Generation and Database Searching with Catalyst, vol. 3, ESCOM Science, 1995, pp. 1–20.
[10] M.O. Taha, Y. Bustanji, M. Al-Ghussein, M. Mohammad, H. Zalloum, I.M. Al-Masri, N. Atallah, 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 (2008) 2062–2077.
[11] I.M. Al-masri, M.K. Mohammad, M.O. Taha, Discovery of DPP IV inhibitors by pharmacophore modeling and QSAR analysis followed by in silico screening, ChemMedChem 3 (2008) 1763–1779.
[12] M.O. Taha, A.M. Qandil, D.D. Zaki, M.A. AlDamen, Ligand-based assessment of factor Xa binding site flexibility via elaborate pharmacophore exploration and genetic algorithm-based QSAR modeling, Eur. J. Med. Chem. 40 (2005) 701–727.
[13] M.O. Taha, N. Atallah, A.G. Al-Bakri, C.P. -Bleau, H. Zalloum, K. Younis, R.C. Levesque, Discovery of new MurF inhibitors via pharmacophore modeling and QSAR analysis followed by in-silico screening, Bioorg. Med. Chem. 16 (2008) 1218–1235.
[14] M.O. Taha, Y. Bustanji, A.G. Al-Bakri, M. Yousef, W.A. Zalloum, I.M. Al-Masri, N. Atallah, Discovery of new potent human protein tyrosine phosphatase inhibitors via pharmacophore and QSAR analysis followed by in silico screening, J. Mol. Graph. Model 25 (2007) 870–884.
[15] A.M. Abu Hammad, M.O. Taha, 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 (2009) 978–996.
[16] R. Abu Khalaf, G. Abu Sheikha, Y. Bustanji, M.O. Taha, Discovery of new cholesteryl ester transfer protein inhibitors via ligand-based pharmacophore modeling and QSAR analysis followed by synthetic exploration, Eur. J. Med. Chem. 45 (2010) 1598–1617.
[17] A. Al-Nadaf, G. Abu Sheikha, M.O. Taha, Elaborate ligand-based pharmacophore exploration and QSAR analysis guide the synthesis of novel pyridinium-based potent β-secretase inhibitory leads, Bioorg. Med. Chem 9 (2010) 3088–3115.
[18] Catalyst User Guide. Accelrys Software Inc., San Diego, 2005.

[19] P. Ducrot, M. Legraverend, D.S. Grierson, 3D-QSAR CoMFA on cyclin-dependent kinase inhibitors, J. Med. Chem. 43 (2000) 4098–4108.
[20] P. Gund, Three-dimensional pharmacophoric pattern searching, Prog. Mol. Subcell. Biol. 11 (1977) 117–143.
[21] P.W. Sprague, R. Hoffmann Curr. Tools Med. Chem, in: H. Waterbeemd, V.d. Testa, B.G. Folkers (Eds.), CATALYST Pharmacophore Models and Their Utility as Queries for Searching 3D Databases, VHCA, Basel, 1997, pp. 223–240.
[22] D. Barnum, J. Greene, A. Smellie, P. Sprague, Identification of common functional configurations among molecules, J. Chem. Inf. Comput. Sci. 36 (1996) 563–571.
[23] A. Smellie, S. Teig, P.P. Towbin, Promoting conformational variation, J. Comput. Chem. 16 (1995) 171–187.
[24] H. Li, J. Sutter, R. Hoffmann, in: O.F. Güner (Ed.), Pharmacophore Perception, Development, and Use in Drug Design, InternationalUniversity Line, California, 2000, pp. 173–189.
[25] J. Sutter, O. Güner, R. Hoffmann, H. Li, M. Waldman, Effect of Variable Weights and Tolerances on Predictive Model Generation. Pharmacophore Perception, Development, and Use in Drug Design. International University Line, California, 2000, pp. 501–511.
[26] Y. Kurogi, O.F. Güner, Pharmacophore modeling and three-dimensional database searching for drug design using catalyst, Curr. Med. Chem. 8 (2001) 1035–1055.
[27] I.B. Bersuker, B.S. eci, J.E. . Boggs, in: O.F. Güner (Ed.), Pharmacophore Perception, Development, and Use in Drug Design, International University Line, California, 2000, pp. 457–473.
[28] K. Poptodorov, T. Luu, R. Hoffmann, in: T. Langer, R.D. Hoffmann (Eds.), Methods and Principles in Medicinal Chemistry, Pharmacophores and Pharmacophores Searches, vol. 2, Wiley-VCH, Weinheim, 2006, pp. 17–47.
[29] J. Singh, C.E. Chuaqui, P.A.B. -Sjodin, W.-C. Lee, T. Pontz, M.J. Corbley, H.-K. Cheung, R.M. Arduini, J.N. Mead, M.N. Newman, J.L. Papadatos, S. Bowes, S. Josiah, L.E. Ling, Successful shape-based virtual screening: the discovery of a potent inhibitor of the type I TGFbeta receptor kinase (TbetaRI), Bioorg. Med. Chem. Lett. 13 (2003) 4355.
[30] P.A. Keller, M. Bowman, K.H. Dang, J. Garner, S.P. Leach, R. Smith, A. McCluskey, Pharmacophore development for corticotropin-releasing hormone: new insights into inhibitor activity, J. Med. Chem. 42 (1999) 2351.
[31] R.G. Karki, V.M. Kulkarni, Computer-aided design and synthesis of *Candida albicans* MyristoylCoA: protein *N*-myristoyltransferase inhibitors, Eur. J. Med. Chem. 36 (2001) 147.
[32] M.O. Taha, A.G. Al-Bakri, W.A. Zalloum, Discovery of potent inhibitors of pseudomonal quorum sensing via pharmacophore modeling and in silico screening, Bioorg. Med. Chem. Lett. 16 (2006) 5902–5906.
[33] K. Moffat, V.J. Gillet, M. Whittle, G. Bravi, A.R. Leach, ChemInform abstract: a comparison of field-based similarity searching methods: CatShape, FBSS, and ROCS, J. Chem. Inf. Model 48 (2008) 719–729.
[34] S. Huang, R. Li, P.J. Connolly, S. Emanuel, S.A. Middleton, Synthesis of 2-amino-4-(7-azaindol-3-yl)pyrimidines as cyclin dependent kinase 1 (CDK1) inhibitors, Bioorg. Med. Chem. Lett. 16 (2006) 4818–4821.
[35] R.P. Sheridan, S.K. Kearsley, Why do we need so many chemical similarity search methods? Drug Discov. Today 7 (2002) 903–911.
[36] R. Fisher, The Principle of Experimentation Illustrated by a Psycho-physical ExpeHafner Publishing Co, eighth ed.. Hafner Publishing, NewYork, 1966, (Chapter II).
[37] A. Tropsha, P. Gramatica, V.K. Gombar, The importance of being earnest: validation is the absolute essential for successful application and inerpretation of QSPR Models, QSAR Comb. Sci. 22 (2003) 69–77.
[38] CERIUS2 QSAR Users' Manual. Accelrys Inc., San Diego, CA, 2005.
[39] L.F. Ramsey, W.D. Schafer, The Statistical Sleuth, first ed.. Wadesworth Publishing Company, USA, 1997.
[40] E.M. Krovat, T. Langer, Non-peptide angiotensin II receptor antagonists: chemical feature based pharmacophore identification, J. Med. Chem. 46 (2003) 716–726.
[41] D. Bednarczyk, S. Ekins, J.H. Wikel, S.H. Wright, Influence of molecular structure on substrate binding to the human organic cation transporter, hOCT1, Mol. Pharmacol. 63 (2003) 489–498.
[42] S. Deswal, N. Roy, Quantitative structure activity relationship of benzoxazinone derivatives as neuropeptide Y Y5 receptor antagonists, Eur. J. Med. Chem. 41 (2006) 552–557.
[43] D.H. Cho, S.K. Lee, B.T. Kim, K.T. No, Quantitative structure–activity relationship (QSAR) study of new fluorovinyloxyacetamides, Bull. Korean Chem. Soc. 22 (2001) 388–394.
[44] P. Buchwald, N. Bodor, Computer-aided drug design: the role of quantitative structure–property, structure–activity and structure–metabolism relationships (QSPR, QSAR, QSMR), Drugs Fut. 27 (2002) 577–588.
[45] M.O. .Taha, L.A. Dahabiyeh, Y. Bustanji, H. Zalloum, S. Saleh, 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 (2008) 6478–6494.
[46] O.O. Clement, A.T. Mehl, Pharmacophore perception, development, and use in drug design. in: O.F. Guner (Ed.), IUL Biotechnology Series. InternationalUniversity Line, La Jolla, CA, 2000, pp. 71–84.
[47] J.H. Van Drie, Pharmacophore discovery–lessons learned, Curr. Pharm. Des. 9 (2003) 1649–1664.
[48] A.P. Graves, R. Brenk, B.K. Shoichet, Decoys for docking, J. Med. Chem. 48 (11) (2005) 3714–3728.
[49] C.A. Lipinski, F. Lombardo, B.W. . Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug. Del. Rev. 46 (2001) 3.
[50] D.F. Veber, S.R. Johnson, H.Y. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates, J. Med. Chem. 45 (2002) 2615.
[51] M.L. Verdonk, L. Marcel, V. Berdini, M.J. Hartshorn, W.T.M. Mooij, C.W. Murray, R.D. Taylor, P. Watson, Virtual screening using protein-ligand docking: avoiding artificial enrichment, J. Chem. Inf. Comput. Sci. 44 (2004) 793–806.
[52] J. Kirchmair, P. Markt, S. Distinto, G. Wolber, T. Langer, 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. 22 (2008) 213–228.
[53] J.J. Irwin, B.K. Shoichet, ZINC – a free database of commercially available compounds for virtual screening, J. Chem. Inf. Comput. Sci. 45 (2005) 177–182.
[54] N. Triballeau, F. Acher, I. Brabet, J.-P. Pin, H.-O. Bertrand, 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 (2005) 2534–2547.
[55] M. Jacobsson, P. Liden, E. Stjernschantz, H. Bostroem, U. Norinder, Improving structure-based virtual screening by multivariate analysis of scoring data, J. Med. Chem. 46 (2003) 5781–5789.
[56] J.K. Holmes, M.J. Solomon, A predictive scale for evaluating cyclin-dependent kinase substrates, J. Biol.Chem. 271 (1996) 25240–25246.
[57] S.M. Rodems, B.D. Hamman, C. Lin, J. Zhao, S. Shah, D. Heidary, A FRET-based assay platform for ultra-high density drug screening of protein kinases and phosphatases, Assay Drug Dev. Technol. 1 (2002) 9–19.
[58] S. Zhang, J. Ma, Y. Bao, P. Yang, L. Zou, K. Li, X. Sun, Nitrogen-containing flavonoid analogues as CDK1/cyclin B inhibitors: synthesis, SAR analysis, and biological activity, Bioorg. Med. Chem. 16 (2008) 7128–7133.
[59] D.W. Lamson, M.S. Brignall, Antioxidants and cancer, part 3: quercetin, Altern. Med. Rev. 5 (2000) 196–208.
[60] C. Patrick, H. Nicole, O. David, R. Thomas, L. Mary, B. Amanda, A. George, C. W. Douglas, L. Brian, A novel class of cyclin-dependent kinase inhibitors identified by molecular docking act through a unique mechanism, J. Bio. Chem. 284 (2009) 29945–29955.
[61] L. Zhang, H. Zhu, Q. Wang, H. Fang, W. Xu, M. Li, Homology modeling, molecular dynamic simulation and docking studies of cyclin dependent kinase 1, J. Mol. Model., in press, doi:10.1007/s00894-010-0710-z.
[62] S. Vadivelan, B.N. Sinha, S.J. Irudayam, A.R. Sarma, P. Jagarlapudi, Virtual screening studies to design potent CDK2-cyclin A inhibitors, J. Chem. Inf. Model 47 (2007) 1526–1535.
[63] S.Y. Wu, I. McNae, G. Kontopidis, S.J. McClue, C. McInnes, K.J. Stewart, S. Wang, D.I. Zheleva, H. Marriage, D.P. Lane, P. Taylor, P.M. Fischer, M.D. Walkinshaw, Discovery of a novel family of CDK inhibitors with the program LIDAEUS: structural basis for ligand-induced disordering of the activation loop, Structure 11 (2003) 399–410.
[64] D. Lesuissea, G. Dutruc-Rosseta, G. Tiraboschic, M.K. Dreyerf, S. Maignand, A. Chevaliera, F. Halleya, P. Bertrand, M.-C. Burgevinb, D. Quarteronet, T. Rooney, Rational design of potent GSK3b inhibitors with selectivity for Cdk1 and Cdk2, Bioorg. Med. Chem. Lett. 20 (2010) 1985–1989.
[65] F. Sliman, M. Blairvacq, E. Durieu, L. Meijer, J. Rodrigo, D. Desmaële, Identification and structure–activity relationship of 8-hydroxy-quinoline-7-carboxylic acid derivatives as inhibitors of Pim-1 kinase, Bioorg. Med. Chem. Lett. 20 (2010) 2801–2805.
[66] D.A. Scott, K.J. Bell, C.T. Campbell, D.J. Cook, L.A. Dakin, D.J.D. Valle, L. Drew, T.W. Gero, M.M. Hattersley, C.A. Omer, B. Tyurin, X. Zheng, 3-Amido-4-anilinoquinolines as CSF-1R kinase inhibitors 2: optimization of the PK profile, Bioorg. Med. Chem. Lett. 19 (2009) 701–705.
[67] L. Chen, C. Shaoqing, M. Christophe, Preparation of quinolinyl thiazolinones as CDK1 inhibitors for treatment of cancers, PCT Int. Appl. (2006) 55 CODEN: PIXXD2 WO 2006029863 A1 20060323 CAN 144:331426 AN 2006:273697 CAPLUS.