CDK2 is considered as a potential target for cancer therapy. In order to find novel CDK2 inhibitors which have different scaffolds, a training set of 17 compounds and a test set of 39 compounds were collected from the published literatures. Pharmacophore models were built and validated by different methods. The hypotheses chosen were then used for chemical databases virtual screening, the selected compounds were further analyzed and refined using drug-like filters and ADMET analysis. Finally, forty five hits with different scaffolds were picked out for docking studies to find tophits. These hits were predicted to have high inhibitory activity and good ADMET properties; they may act as novel leads for CDK2 inhibitors designing.

Keywords: CDK2; inhibitor; pharmacophore; docking; virtual screening.

INTRODUCTION

Cyclins-dependent kinases (CDKs) are members of the CMGC family of the human kinome, they contain the group which mediate cell cycle progression (e.g. CDK1, CDK2, CDK4 & CDK6). The CDKs are regulated in the cell by their partner’s cyclins, specific CDK inhibitors (CDKIs), their state of phosphorylation and ubiquitin mediated deregulation. Specific CDK/cyclin complexes are responsible for each phase of the cell cycle, CDK2 has two partners’ cyclins, A & E; CDK2 activated during G1 phase (CDK2/E complexe) is the checkpoint of G1/S phase transition, while CDK2/A complex drives the cell cycle through the S phase.

Disruption of the normal regulation controls can be caused by either expensive production of CDKs or insufficient production of CDKIs. Therefore, discovery of novel compounds that target CDK2 has become important for treating cancers specially breast, lung and prostate cancers. Many published reports describe various chemical compounds that target ATP binding site within CDK2 kinase, its inactive form was found by Knapp et al to be type II inhibitors. This type binds not only to the ATP region (with Glu81 & Leu83) but also exploits hydrogen bonding and hydrophobic interactions within the allosteric site.

In the present paper, our goal is to generate a pharmacophore model that can be applied in virtual screening to discover of novel lead compounds. Pharmacophore modeling is a practical tool helping in the discovery of novel lead compounds while the choice of compounds used for the discovery (training set) can determine the quality of the pharmacophore model; two important criterias in choosing the training set were structural diversity and activity ranges of selected compounds. Our aim was to use Discovery Study program (DS) to generate an accurate and powerful 3D-pharmacophore model for CDK2. Validated pharmacophore model was further used in screening study of two large chemical databases; Asinex BioDesign and Elite Libraries. Selected molecules with different scaffolds were then refined using certain drug-like filters and ADMET analysis to choose forty five hits. Molecular docking study of hits was carried out to analyze their bind mode with the protein. Our study shows that one molecule (16) seems to act as novel lead for CDK2 inhibitors designing.

MATERIALS AND METHODS

Generation of pharmacophore model

The concept of pharmacophore (pharmacon-phoros) was identified by Errlich in 1909, it defines pharmacophore as a molecular framework that carries the essential features responsible for the biological activity of a drug. Structure-based pharmacophore modeling uses the accessible information of the 3D structure of a macromolecular target and its active site. The information extract, directly from the macromolecule and its interactions with a ligand (macromolecule ligand complex), can be employed to determine the key interaction points between ligands and macromolecule. Discovery Studio 2.5 program was used to achieve this study.

In our study, a crystal structure (PDB codes: 1W0X) of CDK2 complexed with Olomoucine as known inhibitor (IC50 = 0,007-7 μM) was employed to generate pharmacophore models. The crystal structure of the protein was prepared by using general purpose protocol, Define and Edite binding site was used to determine a sphere within 9.5 Å distance from the inhibition, water molecules were removed due to their high mobility and to prevent complication. The pharmacophore hypotheses can be further generated into two methods: common
feature pharmacophore model and 3D QSAR pharmacophore model. A set of 56 ligands with known activity were collected from published literatures and were used for our studies as in Figure 1 & 2. This set will be subdivided into two categories; a training set of 17 molecules to generate hypothesis and a test set of 39 molecules to validate these hypotheses [18-38].

3D structures of these molecules were built by ChemBioOffice version 2008; the energies of their 3D structures were minimized by Minimization protocol in DS. Generate Conformations protocol was applied to generate all the energetically reasonable conformations, the most stable conformation of each compound was used for the study.

![Figure 1: Compounds of training set](image-url)
Figure 2: Compounds of test set.

Generation of common feature pharmacophore model

HipHop algorithm, which uses information from the activity of ligands collected, was employed to generate common feature pharmacophores. A list of different necessary features including hydrogen-bond acceptor (HBA), hydrogen-bond Donor (HBD), hydrophobic (HY); AR features were effectively mapped in all the training set molecules. Common feature pharmacophore model of the DS protocol with general parameters was applied to generate pharmacophoric features corresponding to all the possible points of interaction at the active site.
Generation of 3D QSAR pharmacophore model

HypoGen algorithm of DS program was also used to generate the hypothesis; this step uses the activity of training set compounds 40. The activity values of 17 compounds were spanned over 5 orders of magnitude, a set of different chemical features was selected (maximum 5) to construct the hypotheses. 3D QSAR pharmacophore generation from protocol was applied to build and identify 9 scoring hypotheses models 40.

Pharmacophore validation

All the active molecules must have the same pharmacophore for CDK2 therefore, a test set of 39 chosen active compounds were used to assess and to validate the generated pharmacophore models. Ligand pharmacophore mapping with Calayst algorithm was employed to identify ligands that map to a pharmacophore, and aligns the ligands to the query; two omitted features and rigid method of fitting were selected parameters11-12. Fit value for each hypothesis was given to estimate the activities of the test set compounds; two hypotheses (Hypo8 & Hypo7) were chosen to virtual screening study.

Virtual screening

The purpose of virtual screening is to find potential leads with different scaffolds and high inhibitory activity to CDK2. We used the best 3D-QSAR pharmacophores (Hypo8 and then Hypo7) as 3D query for searching potent compounds from two chemical downloaded databases; Asinex libraries including Bio Design and Elite. In order to further validate the quality of the hypo 8, all the 56 active molecules were also introduced into the screening study 12. Fast conformation generation for each compound was done and fit values were chosen to measure how well the ligand fits the pharmacophore. The compounds which fitted the features of the best 3D-QSAR pharmacophore model and then the common feature pharmacophore model were extracted to undergo an additional analysis; Lipinski’s rule ‘rule of five’ and some of the ADMET analysis such as aqueous solubility levels, BBB penetration levels, CYP2D6 inhibition and hepatotoxicity were also carried out 41, 18. Only the molecules which in accordance with restriction were picket out.

Docking study and validation

In order to understand how these ligands bind to the enzyme and to minimize the number of the false positive, docking study was carried out for the hits molecules. Crystal structure 1W0X was prepared by general purpose and used to define the interaction; all the molecules used in the docking study were also minimized and prepared by general purpose to identify the best conformation. All the molecules were docked to the CDK2 binding site using the CDOCKER program and then they were sort by their score (-CDOCKER ENERGY) 11.

Before docking of the compounds selected by recent study, docking study was validated. The validation can be achieved by measuring the RMSD between the orientation of Olomoucine existing in the crystal complexe (1W0X) and his orientation found by our study. The RMSD value must be less than 2Å 42-43.

RESULTS AND DISCUSSION

Pharmacophore generation and validation

In our study, we have extracted 259 CDK2/Cycline complexes from the Protein Data Bank (PDB), most of them consist of the inactive monomer form of CDK2 with different inhibitors. 56 molecules with good structural diversity, different activity’s order were just selected to the pharmacophore study in using the crystal structure of CDK2/Olomoucine complexe (1W0X pdb code). All the molecules were prepared and minimized for the study. The pharmacophore models were generated by two methods; Common feature pharmacophore using HipHop algorithm and 3D QSAR pharmacophore by using the Hipogen algorithm. 4 features were demanded for the generation (HA, HD, HY & HY-ar) in the two methods. 17 active compounds were employed as a training set; they were classified into five activity scales based on their activity values: very high active; IC50 ≤ 1nM, high active; IC50 ≤ 10nM, moderate active; 1000nM < IC50 ≤ 5000nM, low active; 5000nM < IC50 ≤ 10000nM and very low active IC50 > 10000nM

Figure 3: Chemical features of 9 Hypotheses by 3D QSAR Pharmacophore models
The common feature pharmacophore method has led to 10 pharmacophore hypotheses while the 3D QSAR method has produced 9 hypotheses as clear from Figure 3; the hypotheses of 3D QSAR were chosen principally to continue the study. All the hypotheses were found to have good results, total cost values, correlation coefficient, RMS deviation as reported in Table 1.

### Table 1: Statistical parameters of the generated pharmacophore models

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Total cost</th>
<th>RMS</th>
<th>Correlation</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypo1</td>
<td>70.4378</td>
<td>0.448923</td>
<td>0.534673</td>
<td>HA, Ar</td>
</tr>
<tr>
<td>Hypo2</td>
<td>70.458</td>
<td>0.452405</td>
<td>0.510675</td>
<td>HA, Ar</td>
</tr>
<tr>
<td>Hypo3</td>
<td>70.5673</td>
<td>0.46195</td>
<td>0.531861</td>
<td>HA, HD</td>
</tr>
<tr>
<td>Hypo4</td>
<td>70.6523</td>
<td>0.471666</td>
<td>0.511198</td>
<td>HA, HD</td>
</tr>
<tr>
<td>Hypo5</td>
<td>71.0512</td>
<td>0.513776</td>
<td>0.416961</td>
<td>HA, HD</td>
</tr>
<tr>
<td>Hypo6</td>
<td>71.1722</td>
<td>0.525699</td>
<td>0.387559</td>
<td>HA, HD</td>
</tr>
<tr>
<td>Hypo7</td>
<td>73.0666</td>
<td>0.577204</td>
<td>0.551907</td>
<td>HA, HD, H</td>
</tr>
<tr>
<td>Hypo8</td>
<td>73.5297</td>
<td>0.548833</td>
<td>0.664549</td>
<td>HA, HD, H</td>
</tr>
<tr>
<td>Hypo9</td>
<td>73.6102</td>
<td>0.607342</td>
<td>0.490234</td>
<td>HA, HA, H</td>
</tr>
</tbody>
</table>


Two hypotheses (Hypo 7 & 8), which comprise three pharmacophoric features; one hydrogen bond acceptor, one hydrogen bond donor and one hydrophobic group were selected. There are about 6,4 Å between HA and both of HD and Hydrophobic groups while the distance appears to be far between HBD and Hydrophobic feature (Figure 4a). Hydrogen bond donor is located near the Glu83 while we can see hydrogen bond acceptor in front of Lys33.

![Figure 4: a) Chemical features of Hypo 7 & 8 and the distance between them, b) Chemical features of Phar8 with the nearby amino acids.](image)

In order to validate our hypotheses, 39 Molecules of test set were used to map with the hypo8 as clear in Figure 4b. All the test set appears to map with high fit values especially molecules which have good activity such as HDT & NS9 as reported in Table 2. For the hypo 8, all the pharmacophoric features are around the important active site of CDK2. Hydrogen bond Acceptor (HA) is near the amino group of Leu83. Hydrogen bond donor (HD) is near the carboxylic group of Glu81. Hydrophobic feature (HY) is close to hydrophobic amino acid Phe82 as shown in Figure 5. Compounds mapping on some of these identified features may have potential to inhibit CDK2 with high affinity.

**Virtual screening and filter**

The purpose of virtual screening is to find potential leads with different scaffolds and high inhibitory activity to CDK2. Asinex Databases including BioDesign and Elite (116,563 & 104,577 compounds, respectively) were downloaded to use for virtual screening in our study, all the 56 active molecules were also used in the study. Hypo8 was firstly used to screen databases and then Hypo7 was applied to screen the selected molecules. The virtual screening selected all the compounds containing the pharmacophore (Hypo8 & Hypo7), fast conformation generation and 2 maximum omitted features were chosen as parameters. All the compounds were sorted by their fit values score to select only the one which has fit values equal or more than 3. The molecular properties and ADMET Descriptors of selected compounds were
then calculated by Analyze Small Molecules from ‘Tools’. Drug-likeness properties are usually used for filtering compounds for vitro inhibition studies. Here, the restrictions of Lipinski’s rule of five (Lipinski, 1997) were as follows: logP ≤ 5, molecular weight ≤ 500D, hydrogen bond acceptors (HBA) ≤ 10, hydrogen bond donors (HBD) ≤ 5, and rotated bonds ≤10.

**Table 2: Best Fit values of the test set group with the Hypo8.**

<table>
<thead>
<tr>
<th>PDB Code</th>
<th>Reference IC50 µM</th>
<th>Best Fit value</th>
<th>PDB Code</th>
<th>Reference IC50 µM</th>
<th>Best Fit value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDT</td>
<td>0.003</td>
<td>3.39648</td>
<td>5BN</td>
<td>0.012</td>
<td>2.63545</td>
</tr>
<tr>
<td>UCN</td>
<td>0.03</td>
<td>3.29038</td>
<td>U32</td>
<td>0.02</td>
<td>2.62813</td>
</tr>
<tr>
<td>C95</td>
<td>47</td>
<td>3.21449</td>
<td>107</td>
<td>0.01</td>
<td>2.62796</td>
</tr>
<tr>
<td>N59</td>
<td>0.003</td>
<td>2.87522</td>
<td>740</td>
<td>0.03</td>
<td>2.61787</td>
</tr>
<tr>
<td>T7Z</td>
<td>0.123</td>
<td>2.71789</td>
<td>STU</td>
<td>0.007</td>
<td>2.61599</td>
</tr>
<tr>
<td>2BZ</td>
<td>0.035</td>
<td>2.7149</td>
<td>U55</td>
<td>2.61435</td>
<td></td>
</tr>
<tr>
<td>18K</td>
<td>0.7</td>
<td>2.71261</td>
<td>3TI</td>
<td>0.005</td>
<td>2.60979</td>
</tr>
<tr>
<td>C85</td>
<td>0.18</td>
<td>2.6998</td>
<td>CDK</td>
<td>0.00011</td>
<td>2.56831</td>
</tr>
<tr>
<td>CK8</td>
<td>0.22</td>
<td>2.69948</td>
<td>SCJ</td>
<td>0.1</td>
<td>2.54172</td>
</tr>
<tr>
<td>FRT</td>
<td>0.001</td>
<td>2.69871</td>
<td>BWP</td>
<td>0.3</td>
<td>2.52826</td>
</tr>
<tr>
<td>CK4</td>
<td>0.29</td>
<td>2.69454</td>
<td>B49</td>
<td>130</td>
<td>2.52318</td>
</tr>
<tr>
<td>EZR</td>
<td>10</td>
<td>2.68995</td>
<td>LS2</td>
<td>0.0057</td>
<td>2.52272</td>
</tr>
<tr>
<td>G62</td>
<td>27</td>
<td>2.68637</td>
<td>FAL</td>
<td>22</td>
<td>2.52130</td>
</tr>
<tr>
<td>L27</td>
<td>0.85</td>
<td>2.67481</td>
<td>EZV</td>
<td>1.04</td>
<td>2.52016</td>
</tr>
<tr>
<td>LIA</td>
<td>0.003</td>
<td>2.67466</td>
<td>1PU</td>
<td>0.44</td>
<td>2.46359</td>
</tr>
<tr>
<td>T3E</td>
<td>0.004</td>
<td>2.66728</td>
<td>2PU</td>
<td>10</td>
<td>2.35764</td>
</tr>
<tr>
<td>1N3</td>
<td>0.07</td>
<td>2.66376</td>
<td>SU9</td>
<td>0.110</td>
<td>2.28519</td>
</tr>
<tr>
<td>L29</td>
<td>0.003</td>
<td>2.65269</td>
<td>5BP</td>
<td>0.12</td>
<td>2.14347</td>
</tr>
<tr>
<td>L8</td>
<td>0.14</td>
<td>2.64961</td>
<td>SCE</td>
<td>20</td>
<td>1.93466</td>
</tr>
<tr>
<td>CT9</td>
<td>0.059</td>
<td>2.6465</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Chemical structures](image-url)
ADMET (absorption, distribution, metabolism, elimination, toxicity) analysis is important in drug design. Some properties including aqueous solubility levels, BBB penetration levels, CYP2D6 inhibition and hepatotoxicity were used in our study. Their Brain-Blood ratio must be less than 0.3/1, they must be unlikely to inhibit CYP2D6 enzyme, unlikely to cause dose-dependent liver injuries, aqueous solubility level log(Sw) must be between -4.0 to 0.0. Only these molecules which in accordance with Lipinski’s rule and have good predicted activity and good ADMET properties can be considered as hits; only 45 molecules have been meet the conditions as shown in Figure 6.

**Molecular docking studies**

In order to understand the binding mode into the active site of CDK2, docking study was carried out using the CDOCKER. Docking program can generate all the biologically active conformations of the ligands and identifying the binding modes to the CDK2. The CDK2

**Figure 6:** Compounds selected from screening study.
protein was prepared and the marcomelcule was
minimized and a sphere of 9.5 Å around the active site
was determined. All the molecules were also prepared
from ‘General Purpose’. CHARMM forcefield have been
used, 10 top conformers for each compound have been
demanded and the results were filtered by their CDOCKER
ENERGY score. The study was validated by measuring the
RMSD between the position of Olomoucine as reference
and the position of the same molecule by our study, all
the conformers have been found to have RMSD 0.45Å.
45 molecules (hits) retrieved from virtual screening were
docked to the crystal structure of CDK2 binding site and
the mode of binding was analyzed. Multi hits with good
CDOCKER score have been found, for example the
molecule 16 was found to have high CDOCKER score (40,
04). This compound can connect with Glu83 by hydrogen
bond while his aromatic ring can form π-interaction with
Phe82 as shown in Figure 7. It may act as potential
inhibitor to CDK2 to continue In vitro study and it is
considered as lead compound to prepare different
analogs with high infinity to the active site of CDK2.

CONCLUSION
In this study, two kinds of pharmacophores were built to
screen novel and potent CDK2 inhibitors. Hypo8, the
pharmacophore model, is composed of three chemical
features including one hydrogen bond acceptors, one
hydrogen bond donor, and one hydrophobic feature. Both
generated pharmacophore models were validated for its
quality to identify new reliable chemical compounds by
using a test set of 39 molecules. All the test set
compounds have mapped all the features of Hypo8 and
Hypo7. 221,140 compounds of two libraries (Asinex
BioDesign & Elite) have been downloaded to map to the
chemical features of Hypo8 & Hypo7. The selected
compounds were further analyzed and refined using
drug-like filters and ADMET analysis. At last, 45 hits were
selected out; molecular docking procedure was carried
out for these hits to understand the interaction with the
active site of the protein. The docking study allowed us to
discover molecule 16 to be a tophit and may act as good
leads against CDK2.

Acknowledgment: We would like to make honor to
Prof. Mohamed Abdel Hamid Ismail; Professor of
Pharmaceutical Chemistry in the Faculty of pharmacy in
the Ain Shams University, for providing the DS program
and we give our work as a tribute to his spirit.

REFERENCES
1. Matthews DJ. Gerritsen ME. kinase and cancer. targeting
protein kinases for cancer. John Wiley & Sons. INC. United
States, 2010, 32-33.
2. Sethi A, Yadav S. Pharmacophore mapping-An Approach for
designing cyclin dependent kinesis -2 cancer treatment. Int.
5. Sielecki TM, Boylan JF, Benfie PA, et al. cyclin dependant
kinase inhibitors: useful targets in cell cycle regulation. J.
6. Akiyama T, Ohuchi T, Sumida S, et al. Phosphorylation of the
89, 1992, 7900-7904.
7. Canavese M, Santo L, Raje N. Cyclin dependent kinases in
cancer Potential for therapeutic intervention. Cancer Biology
8. Alexander LT, Möbitz H, et al. Type II Inhibitors Targeting
Exploiting the Active Kinase Conformation. J. Med. Chem.
screening, and molecular docking studies for discovery of
11. Ece A, Sevin F. The discovery of potential cyclin A/CDK2
inhibitors: a combination of 3D QSAR pharmacophore
modeling, virtual screening, and molecular docking studies.


27. Kang YN, Stueky JA. Crystal structure of the cdk2 in complex with thiazolylpyrimidines. To be published.


42. Zavodszky MI, Kuhn LA. Improving Docking Validation. kuhnlab.bmb.msu.edu.


Source of Support: Nil, Conflict of Interest: None.