Identification of a putative anti-rheumatoid arthritis molecule by virtual screening

Shazi Shakil1-3*, Adel M Abuzenadah1-3, Suzan M Attar4, Omar Fathaldin4, Rajaa Al-Raddadi5, Mansour I Sulaiman6

1King Fahd Medical Research Center, 2Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, 3Center of Excellence in Genomic Medicine Research, 4Rheumatology Department, 5Department of Community Medicine, Faculty of Medicine, 6Department of Pharmacology, College of Medicine, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

*For correspondence: Email: shazibiotech@gmail.com, sfaruqi@kau.edu.sa

Sent for review: 3 February 2020

Abstract

Purpose: To propose an improved chemical skeleton whose scaffolds could be used for the design of future thymidylate synthase (TS)-inhibitors against rheumatoid arthritis.

Methods: The drug discovery platform, ‘MCULE’, was employed for inhibitor-screening. The ‘methotrexate-interaction site’ in the crystal (PDB ID 5X66) was used as a target. One ‘RO5 violation’ was permitted. A maximum of ‘10 rotatable bonds’ and ‘100 diverse molecules’ were also allowed in the protocol. The ‘threshold similarity cut off’ was 0.7. The input values describing the remaining parameters were kept as ‘default’. The ‘Open Babel Linear Fingerprint’ was used for the analyses of molecular descriptors, followed by ADME-check.

Results: 4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine corresponding to the MCULE ID-7590816301-0-93 exhibited the overall best binding with TS. The free energy of binding was -8.6 kcal/mol. A total of 17 amino acid residues were significant for the binding interactions. Importantly, 9 residues were common to methotrexate binding. It satisfied pertinent ADME conditions.

Conclusion: 4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine may emerge as a potent seed molecule for TS-inhibitor design in the context of rheumatoid arthritis. It has satisfied pertinent ADME features. However, there is need for further wet laboratory validation.

Keywords: Anti-rheumatoid arthritis, Inhibitor design, Methotrexate, Seed molecule, Thymidylate synthase, Virtual screening

INTRODUCTION

Target-based computational screening focuses on finding new molecules that might act as inhibitors (or putative drugs) for a chosen protein. This is achieved by exhaustive scanning of huge databases of three-dimensional ligand-structures. In this way, molecules displaying the ‘best-fit’ to the binding-site present on the target are identified [1]. A promising lead-molecule is supposed to have acceptable pharmacokinetic properties.
Methotrexate (a traditional anti-rheumatoid arthritis drug) has been reported to possess several limitations due to toxicity concerns [2]. The actual mechanism by which methotrexate exerts its anti-rheumatoid arthritis action still remains unclear [3]. Nevertheless, methotrexate is used in the treatment of rheumatoid arthritis [4]. On the other hand, methotrexate inhibits thymidylate synthase (TS) enzyme. Inhibition of TS enzyme is an important strategy for anticancer drug design [5,6]. In this way, design of novel TS-inhibitors is important for cancer as well as rheumatoid arthritis [7,8].

Rheumatoid arthritis is a common ailment in the world population. Yet, no permanent cure appears to be available for this autoimmune disease till date. The objective of the research work was to propose an improved chemical skeleton whose scaffolds could be used for design of future TS-inhibitors in the context of rheumatoid arthritis.

METHODS

Binding site examination

The three-dimensional structure of the methotrexate interaction site present on the TS-protein was thoroughly explored using ‘CASTp3.0’ as suggested by Tian et al [9]. This web-based server is free for academic use and indeed gives pretty significant and thorough information about all major and minor grooves relevant to the possible binding interactions for a protein-ligand pair. This server employs the α-shape method to identify pertinent protein features, to measure the volume and area and to compute imprint [9]. The reality crystal structure having the PDB ID as 5X66 available in the Protein Data Bank was used in this study.

Structure-based virtual screening

MCULE online drug discovery platform was employed to perform structure-based virtual screening of five million molecules against the ‘methotrexate interaction site’ of human TS enzyme [10]. The aim was to identify a putative anti-rheumatoid arthritis seed molecule that would occupy the same binding groove as methotrexate. All the required parameters were entered into the MCULE workflow builder. The value for ‘allowed RO5 violation(s)’ was entered as 1 with the intention to keep the early filters somewhat flexible so that a broad range of pharmacophores could be included. The other input parameters for the ‘Basic Property Filter’ tab within the MCULE workflow builder were as follows. A maximum of ‘10 rotatable bonds’ and ‘100 diverse molecules’ were allowed in the protocol. The value entered for ‘sampler size’ was 1000. The ‘threshold similarity cut off’ was fixed at 0.7. The input values describing the remaining parameters were kept at their ‘default’ as given in the drug discovery platform. The value allocated for “the maximum number of compounds after sphere exclusion” was 3 million. The ‘Open Babel Linear Fingerprint’ was used for the analyses of molecular descriptors in the screening process.

Computational docking

All the ligands and water molecules were removed from the complex crystal i.e. PDB ID 5X66. The modified information was saved as a separate .pdb file. This dock-ready file was supplied to the MCULE screening platform. The docking experiments were performed by AutoDock Vina [11]. A grid of 60 Å ×60 Å ×60 Å³ was used to completely cover the methotrexate interaction spot located on the TS enzyme. The values for grid position coordinates which were required for a precise docking procedure i.e x, y and z grid coordinates, were extracted from the reality complex (PDB ID 5X66). These values were 154.725848, 150.510455 and 24.558182, respectively. Docking was performed as per the method of Trott and Olson [11].

VINA scores and ADME features

The test molecules were assigned ranks using VINA [11]. In this way, 45 molecules displaying upper VINA scores were identified. These molecules were further subjected to ADME analyses bya SWISS ADME server which incorporates major drug-likeness and medicinal chemistry filters routinely employed in drug design [12]. The molecules that were able to pass through a minimum of 4 filters were selected.

G criterion, TOX-CHECK and Lipinsky filters

Molecules exhibiting binding free energy (ΔG) greater than -8.5 kcal/mol were excluded. The remaining molecules were subjected to ‘toxicity-checking’ by the TOX-CHECKER of MCULE [10]. Finally, molecules that managed to pass the toxicity filter were examined for RO5 violations (Lipinsky filter).

Molecular contacts of the ‘best molecule’ and ‘reference molecule’ with TS

Re-docking of the reference ligand (i.e. methotrexate) to human TS enzyme was
performed after separating it from the complex crystal (PDB ID 5X66) with the aid of Discovery Studio visualizer. The position coordinates were 154.725848, 150.510455 and 24.558182 for x, y and z, respectively, located in 60 x 60 x 60 Å³ grid volume. Molecular interactions holding the ‘best molecule’ in the binding site of the TS protein were compared with those of the reference ligand with the aid of ‘Molecular Overlay tool’ of the visualizer.

RESULTS

Binding site examination

Molecular exploration of the 3-D structure (PDB ID 5X66) showed that 11 amino acid residues were significant for the binding of methotrexate with human TS enzyme. These residues were Arg 78, Phe 80, Ile 108, Asp 218, Leu 221, Gly 222, Phe 225, Asn 226, Tyr 258, Met 311, and Ala 312.

Structure-based virtual screening

The filtration cascade for in silico screening of 500,000 drug candidates against human thymidylate synthase generated a set of ninety ligands (out of 5 x 10⁶ test molecules) [Figure 1].

The upper layer consisting of 45 molecules chosen on the basis of VINA ranks was fetched to the next filters. These 45 candidate inhibitors were subjected to tests using Lipinsky (Pfizer), Ghose, Veber (GSK), Egan (Pharmacia), Muegge (Bayer), Brenk and PAINS filters. Table 1 shows the significant ‘SWISS ADME’-characteristics of the two best ligands as putative TS-inhibitors (Table 1).

Fifteen molecules that were able to pass through a minimum of 4 filters were selected. However, 10 molecules exhibiting binding free energy (ΔG) greater than -8.5 kcal/mol were excluded. In this way, we narrowed down to 2 molecules that passed the toxicity filter. The MCULE IDs of these two best ligands were MCULE-7590816301-0-93 and MCULE-2794455216-0-61. The corresponding IUPAC names generated from their SMILES with the aid of CHEMSPIDER were 4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine and (2E)-4-(4-Biphenylyl)-N,3-bis(4-methylphenyl)-1,3-thiazol-2(3H)-imine, respectively. The molecule that was a pyrimidine perfectly passed the tests using Lipinsky (Pfizer), Ghose, Veber (GSK), Egan (Pharmacia), Muegge (Bayer), Brenk and PAINS filters.

In contrast, MCULE-2794455216-0-61 did not pass the Ghose filter whereby two violations were observed (i.e. WLOGP>5.6, MR>130). Moreover, it failed the Egan (Pharmacia) filter test due to one violation (i.e. WLOGP>5.88). Its XLOGP3 value was found to be greater than 5 which led to its rejection by Muegge (Bayer) filter. It was found to have a low gastrointestinal absorption. However, it did pass the Brenk and PAINS filters (Table 1).

MCULE-7590816301-0-93, the pyrimidine exhibited an acceptable MLOGP value of 2.67. MCULE-2794455216-0-61 got rejected due to one ‘RO5 violation’. Despite the fact that the aforementioned ligand displayed a slightly higher (negative) binding energy, MCULE-7590816301-0-93 (IUPAC name: 4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine) was designated as the top screened out molecule in the present study owing to its overall best performance against all the filters.

Figure 1: Filtration cascade for structure-based virtual screening of 500,000 drug candidates against human thymidylate synthase

MCULE-7590816301-0-93, the pyrimidine exhibited an acceptable MLOGP value of 2.67. MCULE-2794455216-0-61 got rejected due to one ‘RO5 violation’. Despite the fact that the aforementioned ligand displayed a slightly higher (negative) binding energy, MCULE-7590816301-0-93 (IUPAC name: 4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine) was designated as the top screened out molecule in the present study owing to its overall best performance against all the filters.
**Table 1:** Significant 'SWISS ADME'-characteristics of the two best candidate ligands obtained by virtual screening of 500,000 molecules against human thymidylate synthase

<table>
<thead>
<tr>
<th>MCULE ID and 'SWISS ADME'-characteristics:</th>
<th>MCULE-7590816301-0-93</th>
<th>MCULE-2794455216-0-61</th>
</tr>
</thead>
<tbody>
<tr>
<td>IUPAC Name</td>
<td>4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine</td>
<td>(2E)-4-(4-Biphenylyl)-N,3-bis(4-methylphenyl)-1,3-thiazol-2(3H)-imine</td>
</tr>
<tr>
<td>Chemical Formula</td>
<td>C21H20N4O</td>
<td>C29H24N2S</td>
</tr>
<tr>
<td>Average mass (Da)</td>
<td>344.41</td>
<td>432.58</td>
</tr>
<tr>
<td>Log Po/w (MLOGP)</td>
<td>2.67</td>
<td>6.00</td>
</tr>
<tr>
<td>RO5 violation</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>H-bond acceptors</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>H-bond donors</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No. of rotatable bonds</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Topoligical PS Area (Å²)</td>
<td>45.40</td>
<td>45.53</td>
</tr>
<tr>
<td>Molar Refractivity</td>
<td>111.01</td>
<td>136.11</td>
</tr>
<tr>
<td>GI-Absorption</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Water Solubility, Log S (Ali)</td>
<td>-4.61</td>
<td>-8.63</td>
</tr>
<tr>
<td>Synthetic accessibility</td>
<td>3.33</td>
<td>3.85</td>
</tr>
</tbody>
</table>

**Drug-Likeness and Medicinal Chemistry Filters:**

- **Lipinsky (Pfizer):**
  - YES
  - YES; 1 violation: MLOGP>4.15

- **Ghose:**
  - YES
  - No; 2 violations: WLOGP>5.6, MR>130

- **Veber (GSK):**
  - YES
  - YES

- **Egan (Pharmacia):**
  - YES
  - No; 1 violation: WLOGP>5.88

- **Muegge (Bayer):**
  - YES
  - No; 1 violation: XLOGP3>5

- **Brenk:**
  - YES
  - YES

- **PAINS:**
  - YES
  - YES

**Molecular contacts of the ‘best molecule’ and ‘reference molecule’ with TS**

The free energy of re-docking of methotrexate with human TS was -8.6 kcal/mol. Amino acid residues important for binding of methotrexate molecule within the interacting groove for the re-docked-complex were essentially same as the corresponding crystal housed in the Protein Data Bank (PDB ID: 5X66). The ‘best molecule’ i.e. 4-(4-methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine also displayed a binding energy of -8.6 kcal/mol with the human TS protein. The chemical structures of the two best candidate ligands screened out in this study are shown (Figure 2).

‘Discovery Studio-2-D-Diagram’ of the ‘best candidate ligand’ complexed with the human thymidylate synthase is presented in Figure 3.

Ligand binding residues as well as pertinent molecular-contacts key to clasp the ‘best candidate ligand’ onto the interaction site of the human TS are labeled. Moreover, molecular-interactions of 4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine and those of the reference ligand (methotrexate) with human TS in the corresponding complexes were duly compared. The top screened out ligand was found to interact with human TS through 17 residues [Table 2].

Importantly, 9 of 17 binding residues of the ‘TS-top ligand-complex’ were found to be common to that of ‘TS-reference ligand-complex’. The common binding residues were Phe 80, Ile 108, Asp 218, Leu 221, Gly 222, Phe 225, Tyr 258, Met 311, and Ala 312.

**Figure 2:** Chemical structures of the two best candidate ligands obtained by virtual screening

**Figure 3:** Discovery Studio-2-D-Diagram of the ‘best candidate ligand’ complexed with the human thymidylate synthase.
**Table 2:** Molecular docking interaction energy ($\Delta G$) and important amino acid residues involved in the binding interactions of the two best candidate ligands with the human thymidylate synthase

<table>
<thead>
<tr>
<th>Ligand ID as per MCULE database</th>
<th>MCULE-7590816301-0-93</th>
<th>MCULE-2794455216-0-61</th>
<th>Reference ligand (Methotrexate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular docking Interaction energy ($\Delta G$)</td>
<td>-8.6 kcal/mol</td>
<td>-9.4 kcal/mol</td>
<td>-8.6 kcal/mol</td>
</tr>
</tbody>
</table>

**DISCUSSION**

Successful elaboration of molecular interactions in the past articles duly boosted the scientific morale to carry out the described work. Previous publications have described interactions of several important proteins, namely CTX-M-15, acetylcholinesterase, butyrylcholinesterase, sodium-glucose co-transporter 2, MMP-2 and MMP-9 [13-18]. Binding interactions involving different type of ligands have also been described. The ligands included Forxiga, Fawcettimine, certain plant-based compounds, and the anti-rheumatoid drug, Methotrexate [19-23].

As of now, *in silico* screening has acquired a giant and indispensable share in the area of drug research and development. It has become feasible to examine millions of putative drug structures that possess relevant pharmacophores. Selected molecules could be subjected to required lab tests *in vitro* as well as *in vivo*. SWISS ADME is a well cited yet free to use online facility employed to evaluate the pharmacokinetic properties, lead-likeness, drug-likeness and also medicinal-chemistry related friendliness of small drug candidates [12]. Absorption, distribution, metabolism and excretion are important parameters for drug discovery.

These parameters are abbreviated as ‘ADME’. ‘$\Delta G$’ that stands for the ‘free energy of binding’, is a well-known criterion used in docking studies. The ‘methotrexate binding residues’ of the ‘reference ligand’ and the ‘screened out ligand’ are shown in CPK and stick representations, respectively.
docked-complex’ were essentially the same as those of the corresponding crystal housed in the Protein Data Bank (PDB ID: 5X66). This further confirmed the accuracy of the docking experiments. A twofold strategy was used to impart flexibility to the initial part of the virtual screening. Molecules that could pass a minimum of 4 filters were selected. Moreover, 1 ‘RO5 violation’ was ignored [24]. Consequently, MLOGP value for 4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine was found to be 2.67 by SWISS ADME server. This is in harmony with the accepted norm that lead molecules should possess MLOGP < 4.15. The molecule was found to have a very good GI-absorption. This is a plus for orally administered drugs. It displayed an easy synthetic accessibility score of 3.33.

Presence of only 2 rotatable bonds makes it even more preferable for drug design experiments. This finding is supported by another study where the authors have reported pyrimidine-based TS-inhibitors [25]. In essence, 4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine passed all the major drug screen filters. Hence, it could be an ideal lead molecule for wet laboratory validation.

CONCLUSION

4-(4-Methyl-1-piperazinyl)-2-phenyl[1]benzofuro[3,2-d]pyrimidine displays a robust interaction with human TS, and satisfied pertinent ADME criteria. Therefore, it could serve as a potent seed molecule for TS-inhibitor design in the context of rheumatoid arthritis. However, there is need for further wet laboratory validation.

DECLARATIONS

Acknowledgement

This project was funded by the National Plan for Science, Technology and Innovation (MAARIFAH) – King Abdulaziz City for Science and Technology - the Kingdom of Saudi Arabia (award no. 11-MED1553-03). The authors also acknowledge with thanks Science and Technology Unit, King Abdulaziz University for technical support.

Conflict of interest

No conflict of interest is associated with this work.

We declare that this work was done by the authors named in this article and all liabilities pertaining to claims relating to the content of this article will be borne by the authors.

Open Access

This is an Open Access article that uses a funding model which does not charge readers or their institutions for access and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0) and the Budapest Open Access Initiative (http://www.budapestopenaccessinitiative.org/idea), which permit unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

REFERENCES