Available online at www.sciencedirect.com
ScienceDirect
Mendeleev Commun., 2015, **25**, 214–215
Mendeleev Communications

Combining 3D-QSAR and molecular docking for the virtual screening of PARP inhibitors

Eugeny I. Prokhorov,[a] Aleksandra V. Bekker,[a] Alexander V. Perevoznikov,[a] Mikhail I. Kumskov[a] and Igor V. Svitanko*[b,c,d]

[a] *Department of Mechanics and Mathematics, M. V. Lomonosov Moscow State University, 119991 Moscow, Russian Federation*
[b] *Department of Chemistry, M. V. Lomonosov Moscow State University, 119991 Moscow, Russian Federation*
[c] *N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation. Fax: +7 499 135 5328; e-mail: svitanko@mail.ru*
[d] *Higher Chemical College, Russian Academy of Sciences, 125047 Moscow, Russian Federation*

DOI: 10.1016/j.mencom.2015.05.019

3D-QSAR and molecular docking were applied to predict the inhibitory activity of 196 compounds towards poly(ADP-riboso) polymerase-1 (PARP). A proportion of experimentally active ligands was higher among compounds with good rankings from both methods (57%) compared to compounds scored as inactive by at least one method (40% for docking-active, QSAR-inactive compounds).

Molecular docking[1] and 3D-QSAR[2,3] represent two common approaches to preliminary molecular modeling. We have demonstrated that preliminary molecular modeling could be efficiently used in the synthesis of novel inhibitors of poly(ADP-riboso) polymerase-1 (PARP),[4] which is localized in the cellular nucleus and catalyzes the poly-ADP-ribosylation of miscellaneous protein and DNA substrates.[5] The aim of this work was to compare the performance of these methods and to study whether they produce consolidated predictions. For this purpose, we have performed a search for potential PARP inhibitors in a test set of 196 compounds followed by the experimental validation of the best predicted ligands.

Molecular docking was performed using the Lead Finder 1.1.15 software (available at www.moltech.ru)[6] with default parameters. The full-atomic model of PARP was prepared using the Model Builder software supplied in the Lead Finder distribution package.[7] Ligand binding energies were estimated with the dG-scoring function of Lead Finder.[8]

A structural filtration methodology[9] was employed in order to increase the accuracy of a docking-based rational search for inhibitors. The analysis of the available X-ray structural data of PARP–inhibitor complexes revealed that active ligands form a pair of correlated hydrogen bonds with carbonyl and amide groups of G863 residue (Figure 1). Presumably, these hydrogen bonds are characteristic of the binding mode of active ligands. Therefore, only ligands that formed the above interactions with protein were selected from the docking results, while other ligands were considered inactive.

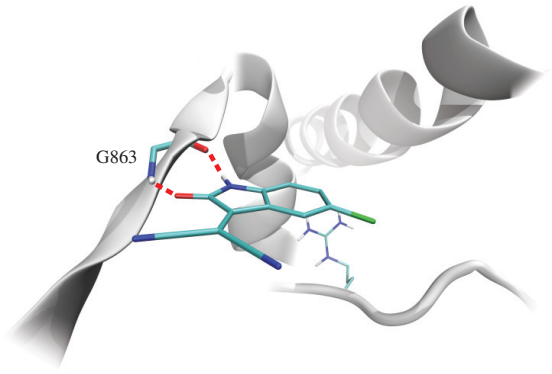

Figure 1 Hydrogen bonds between a PARP backbone residue and an inhibitor used as structural filtration criteria.

Alternatively, the same test set of 196 compounds[10] was analyzed with 3D-QSAR methods.[11] Here, we used a QSAR approach based on concerted activity prediction by several recognizing models, which were constructed by the following algorithms: an iterative algorithm based on a variety of different clusterizations[12] of the training set, a local classifier based on the *k* nearest neighbors (K-NN) method,[13] and a two-phase classifier based on the SVM method.[9] The training set comprised 86 active and 34 inactive PARP inhibitors.

To describe chemical structures, we have employed the allocation of linear fragments using the labeling of molecular graph vertices.[14] Fragment descriptors pertain to information-based descriptors, which tend to code information stored in molecular structures. An important advantage of above descriptors is related to the simplicity of their calculation, storage and interpretation. Descriptor values were defined as the number of repetitive fragments corresponding to molecular graph vertices, the pairs of vertices, triples of the vertices and quadruples of vertices. Adjustable parameters of description are the length of linear fragments (k = 2, 3, 4) and the labels involved in the description (the degree of molecular graph vertices and information on a chemical bond and its position in the ring structure). By varying these parameters, 24 varieties of description were constructed; 72 models were constructed by varying the description of molecular graphs and using different algorithms of recognition.

The molecular docking and 3D-QSAR results were generally in a good agreement (Table 1). A large part (165 of 196) of the test compounds was predicted to be inactive by both methods;

Table 1 Comparison of molecular docking and QSAR-predicted PARP inhibitory activity on the test set of 196 compounds.

Molecular docking	QSAR	
	Active	Non-active
Active	7	19
Non-active	5	165

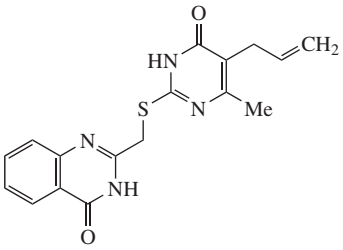

Figure 2 PARP inhibitor from the QSAR-active, docking-inactive set.

12 compounds were predicted to be active according to molecular docking and 26, according to 3D-QSAR. Three subsets were further selected for experimental characterization: compounds predicted to be active by both methods (7 structures), by molecular docking but not by 3D-QSAR (5 structures) and *vice versa* (19 structures).

Among seven active/active (both methods) compounds, four demonstrated a significant *in vitro* inhibitory activity in 10 μM resulting in a 57% screening success rate. Only two of five substances were active in the set of docking-active, QSAR-inactive compounds. One of the most QSAR-active, docking-inactive compounds (Figure 2) demonstrated activity (IC_{50} of 200 μM).†

The above data indicate the importance and noninterchangeability of both of the modeling methods applied. A principal difference between 3D-QSAR and molecular docking is that the former relies only on the ligand structure while the latter directly assesses full-atomic protein–ligand interaction. Thus, molecular docking (especially in combination with structural filtration) allows one to confidently reject ligands that do not form crucial contacts in the protein binding site and thus tend to be inactive, leading to low overall rate of false negative predictions. Nevertheless, the number of false positive molecular docking predictions (50% in this study) is still large and it can certainly be improved. Our results indicate that the concomitant use of QSAR allows one to significantly enrich docking-predicted set indicating that both of the methods can be successfully used in virtual screening.

† PARP inhibitory activity was measured at the Novosibirsk Institute of Chemical Biology and Fundamental Medicine.

This work was supported by the Ministry of Education and Science of the Russian Federation (programme ‘Scientific and Educational Research Personnel of Innovative Russia’, contract no. 8431) and the Russian Foundation for Basic Research (grant nos. 12-03-01036-a and 15-03-08344-a).

References

1 (*a*) F. L. Stahura and J. Bajorath, *Comb. Chem. High Throughput Screen*, 2004, **7**, 259; (*b*) S. F. Sousa, A. J. Ribeiro, J. T. S. Coimbra, R. P. P. Neves, S. A. Martins, H. N. S. Moorthy, P. A. Fernandes and M. J. Ramos, *Curr. Med. Chem.*, 2013, **20**, 2296.
2 N. S. Zefirov and V. A. Palyulin, *Mendeleev Commun.*, 2010, **20**, 243.
3 G. A. Gazieva, Yu. B. Vikharev, L. V. Anikina, T. B. Karpova, A. N. Kravchenko, E. A. Permyakov and I. V. Svitanko, *Mendeleev Commun*., 2013, **23**, 202.
4 L. V. Romashov, A. A. Zeifman, A. L. Zakharenko, F. N. Novikov, V. S. Stroylov, O. V. Stroganov, G. G. Chilov, S. N. Khodyreva, O. I. Lavrik, I. Yu. Titov and I. V. Svitan’ko, *Mendeleev Commun*., 2012, **22**, 15.
5 D. D’Amours, S. Desnoyers, I. D’Silva and G. G. Poirier, *Biochem. J*, 1999, **342** (part 2), 249.
6 O. V. Stroganov, F. N. Novikov, V. S. Stroylov, V. Kulkov and G. G. Chilov, *J. Chem. Inf. Model*., 2008, **48**, 2371.
7 O. V. Stroganov, F. N. Novikov, A. A. Zeifman, V. S. Stroylov and G. G. Chilov, *Proteins*, 2011, **79**, 2693.
8 F. N. Novikov, A. A. Zeifman, O. V. Stroganov, V. S. Stroylov, V. Kulkov and G. G. Chilov, *J. Chem. Inf. Model*., 2011, **51**, 2090.
9 F. N. Novikov, V. S. Stroylov, O. V. Stroganov and G. G. Chilov, *J. Mol. Model*., 2010, **16**, 1223.
10 F. N. Novikov, *PhD Thesis*, Moscow, 2011.
11 E. I. Prokhorov, *Moscow University Math. Bull*., 2013, **68**, 246 [*Vestn. Mosk. Univ., Mat. Mekh.*, 2013, **67** (5), 34].
12 A. V. Bekker, A. A. Suleymanov, G. N. Apryshko and M. I. Kumskov, *Pattern Recognition Image Analysis*, 2011, **21**, 454.
13 A. V. Perevoznikov, A. M. Shestov, E. A. Permyakov and M. I. Kumskov, *Pattern Recognition Image Analysis*, 2011, **21**, 545.
14 M. I. Kumskov, E. A. Smolensky, L. A. Ponomareva, D. F. Mityushev and N. S. Zefirov, *Dokl. Akad. Nauk*, 1994, **336**, 64 (in Russian).

Received: 27th October 2014; Com. 14/4494