Methodology for prediction of anticancer action of (2-oxo-2H-[1,2,4]triazino[2,3-c]-quinazolin-6-yl)thiones via QSAR and docking studies

Zaporizhzhia State Medical University

Key words: Quinazolines, Triazines, Casein Kinase II, Quantitative Structure-Activity Relationship, Molecular Docking Simulation.

Aimed to elaborate new group of protein kinase inhibitors we conducted receptor-based screening (docking, QSAR modeling) and biochemical testing for derivatives of (2-oxo-2H-[1,2,4]triazino[2,3-c]-quinazolin-6-yl)thiones.

Methods and results. This study allowed identifying of new potential anticancer compounds among (2-oxo-2H-[1,2,4]triazino[2,3-c]-quinazolin-6-yl)thiones’ derivatives.

Conclusion. Obtained data may be used for the development of more active and selective inhibitors of protein CK2 kinase. Besides that QSAR-models which were created may be used for planning of chemical modification of structure aimed to creation of new anticancer agents.
**QSAR and statistical analysis.** First of all, all molecules were built by MarvinSketch 6.3.0 [12]. Then they were preliminary optimized by program HyperChem8.0.8 using molecular mechanical MM+ algorithm combined with semi-empirical PM3 molecular modeling method with a maximum number of cycles and Polak-Ribiere (Conjugate Gradient) algorithm. Molecular mechanics has been used to produce more realistic geometry values for the majority of organic molecules owing to the fact of being highly parameterized. The next step was a re-optimization of the MM+ optimized structures by applying semi-empirical PM3 molecular modeling method with a maximum number of cycles. Obtained files were further used for calculations.

Descriptors were calculated using Dragon (> 1600 descriptors). The definition of all used molecular descriptors and the calculation procedures were summarized elsewhere [16,17]. Optimized structures were also used for calculation of additional important quantum-chemical parameters (final heat of formation, total energy, electronic energy, core-core repulsion, ionization potential, homo, lumo), that were also used as descriptors. MOPAC2012 was used to do mentioned computations [15]. Besides, scoring functions obtained by Autodock4 to CK2 kinase was added as a separate descriptor. It is a crucial parameter as it estimates the free energy of ligand binding to the receptor.

The correlation coefficients for all pair of descriptor variables used in the models were evaluated to identify highly correlated descriptors in order to detect redundancy in the data set. Hence, descriptors with constant variables and near-constant variables were excluded from the further consideration (r≥0.95).

The genetic algorithm (GA) and multiple linear regression analysis (MLRA) were used to select the descriptors and to generate the correlation models that relate the structural features to the cell growth percent of different cancer cell lines. The combination of the GA-MLRA technique was applied to obtain the best descriptors among 1671 calculated overall (DRAGON, MOPAC2012, Autodock4), and to construct QSAR models using the QSARINS 2.2.1 [8].

Calculation of QSAR-models was conducted separately for each line of non-small lung cancer (A549) ATCC, EKVX, HOP-62, HOP-92, NCI-H226, NCI-H23, NCI-H322M, NCI-H460, NCI-H522). Growth percent according to the NCI protocol wasn’t converted to any other value, it was used in original version to built models. Some cell lines were given the value of -999, which means, that they were not tested.

Preliminary calculation was made to find the cancer line, which according to the statistical parameters correlated with the calculated descriptors most accurately. Thus, the amount of generation algorithm setup was set until 5 descriptors, and generation per size was established to the value of 500, and the division into training and test sets was performed automatically at a ratio of 80 to 20 percent relatively. Models, which showed statistical significance according to the parameters at a higher level (r≥0.5), were selected for a thorough rendering. For these lines the following options were given: the amount of generation algorithm setup was set until 7 descriptors, and generation per size was established to the value of 10000. Seventy-six derivatives of (2-oxo-2H-[1,2,4]triazino[2,3-c]quinazolin-6-yl) thiiones were spitted into training and test sets and the division, was made such, as to establish equal distribution of substances of high and moderate percentage of inhibition of cell growth.

**Docking.** Receptor-oriented flexible docking was performed by software package Autodock 4.2.6 [13]. Ligands and macromolecules were prepared by software packages Vega ZZ (command line) [14] and MGL Tools 1.5.6 [13]. Autodock works with ligands and receptor molecules of PDBQT format, containing the coordinates of atoms and partial charges. Mol2 format was converted to PDBQT by means of Vega program, hydrogen atoms from non-polar atoms were removed and force field AUTODOCK was added. Changing of the receptor format from PDB to PDBQT and formation of the cards for docking was carried out in programs MGL Tools and AutoGrid.

The catalytic subunit of protein kinase CK2 was chosen as the target for the docking, namely, CK2 kinase, that was crystalized with inhibitor CX-494 (PDB code 3NSZ) [7]. Water molecules, ions and ligands were deleted from original PDB file.

The following parameters were set for the docking: step of forward movement equal 2 Å, quaternion angle – 50°, the torsion angle – 50°. The degree and coefficient of torsion freedom were 2 and 0.274 respectively. Cluster tolerance – 2 Å. External energy of the grid – 1000, the maximum initial energy – 0, the maximum number of attempts – 10000. The number of structures in the population – 300, the maximum number of stages assessing energy – 1000000, the maximum number of generations – 27000, the number of structures that move to the next generation – 1, the level of genetic mutations – 0.02, crossover rate – 0.8, way of crossover – arithmetic. α-Parameter of Gaussian distribution was equal to 0, β-parameter of Gaussian distribution – 1. The number of iterations of Lamarck genetic algorithm search is 10 for each ligand.

**Visual analysis of compounds** interaction with amino acid residues of ATP-binding pocket of protein kinase CK2 was performed in the program Discovery Studio Visualizer 4.0.

© I. S. Nosulenko, O. Yu. Voskoboynik, O. M. Antypenko, G. G. Berest, S. I. Kovalenko, 2015
Inhibition of protein kinase. Expressed in insect cells S21 (Upstate-Millipore) human CK2 kinase domain was used for in vitro test. Compounds' inhibitory activity to protein kinase CK2 was determined by inclusion of radioactive phosphorus in the peptide substrate during its kinase phosphorylation in the presence of γ-32P-ATP [9].

The total volume of the reaction mixture was 30 μL. First to 3 μL of reaction buffer (200 mM of Tris-HCl (pH 7.5), 500 mM KCl, 100 mM MgCl₂) was added 0.5 μL of peptide substrate solution (RRRDDDSDDD (New England Biolabs), 135 μM), 15.5 μL of water and 0.05 μL of protein solution (0.01 protein kinase relative activity). Then 1 microliter of inhibitor was added and after 3 minutes the reaction was initiated by adding to 20 μL of reaction mixture volume 10 μL 150 μM ATP solution, which also contained 1 microcurie of γ-32P-ATP. The final concentration of ATP in the reaction mixture was 50 μM. The reaction mixture was incubated for 30 min at 30 °С. Reaction was stopped by adding 8 μl of 5% phosphoric acid. The entire volume of sample was carried over onto a P-cellulose filter «Whatman P81», which were washed three times for 5 min with 0.75% phosphoric acid. Filters were dried, and their radioactivity was measured on a scintillation counter PerkinElmer Tri-Carb 2800-TR. As a negative control we used a sample of 1 μL DMSO (final concentration was 3.8%) instead of the inhibitor. The degree of inhibition of protein kinase was determined by the ratio of 32P in samples with inhibitor and in his absence.

Results and Discussion
According to the GA-MLRA we have obtained two good predictive models of non-small lung cancer (cell line EKVX and NCI-H522). The obtained equations consist of 6 descriptors. Most of the descriptors, used in models are among 3D ones (RDF, 3D-MoRSE, WHIM and GETAWAY descriptors). Such, it is clear, that not only presence of pharmacophore is important for biological activity, but also its spatial arrangement.

Fig. 2. Correlation of predicted versus experimental GP for model of non-small cell lung cancer, cell line NCI-H522 (Eqn.1)

GP = 192.6738(±98.2228)×SIC2+28.0662(±20.2474)×EEig08r-8.1859(±2.2395)×RDF130u-145.1481(±57.7604)×E3p-45.6237(±12.5782)×nThiazoles+5.2009(±22.3854)×B07[N-O]-132.9902(±102.6973) (Eqn.1)

Statistical data: training set (n=49; r²=0.7583; RMSE tr=13.3049; s=14.3709; F=21.9629; Q²LOO=0.6945); prediction set (n=12; r²=0.6951; RMSE ext=61.7888), where GP – growth percent, n – number of studied compounds, r² – squared regression coefficient, RMSE – root mean square error, F – variance ratio, Fisher coefficient, Q²LOO – weighted correlation coefficient by leave-one out method, and s – standard error.

According to the equation, higher value of SIC2, EEig08r and B07[N-O] is responsible for higher growth percent and responsively for lower anticancer activity. While higher value of RDF130u, E3p and nThiazoles decreases growth percent.

Fig. 3. Correlation of predicted versus experimental GP for model of non-small cell lung cancer, cell line EKVX (Eqn.2)

GP = 3.9897(±2.2446)×RDF145u-4.6468(±0.808)×RDF080e-27.9718(±17.9189)×Mor16v+47.6055(±21.3013)×Mor19v-767.041(±377.4656)×G2m-30.5949(±13.5572)×H-048+267.3181(±69.2583) (Eqn.2)

Statistical data: training set (n=48; r²=0.7878; RMSE tr=10.0414; s=10.8460; F=25.9905; Q²LOO=0.7098); prediction set (n=13; r²=0.7177; RMSE ext=22.5403).

Significance of descriptor contribution can be seen in table 1. The QSAR model containing only one descriptor has value of r²=0.3981. It consists of RDF080e descriptor, that corresponds to radial distribution function – 8.0/weighted by atomic Sanderson electronegativities. It is among the RDF descriptors, obtained by radial basis functions centered on different interatomic distances (from 0.5A to 15.5 A).

Ranking ligand binding was performed by the energy of the kinase domain. It uses a scoring function program of Autodock4.

© I. S. Nosulenko, O. Yu. Voskoboynik, O. M. Antypenko, G. G. Berest, S. I. Kovalenko, 2015
Table 1

| Desc. amount | Descriptors | Training set |
|--------------|-------------|--------------|
| 1            | RDF080e     |              |
| 2            | RDF080e R6p+|              |
| 3            | FINAL HEAT OF FORMATION RDF080e HATS2p |    |
| 4            | MATH2p RDF080u RDF010v Hypertens-80 |    |
| 5            | 6BELp2 RDF100u RDF080e E1s R1v |    |
| 6            | 12RDF145u RDF080e Mor16v Mor19v G2m H-048 |    |

Note: RDF080e – Radial Distribution Function – 8.0/weighted by atomic Sandorin electronnegativities; R6p+ – R maximal autocorrelation of lag 6/weighted by atomic polarizabilities; HATS2p – leverage-weighted autocorrelation of lag 2/weighted by atomic polarizabilities; MATH2p – Moran autocorrelation – lag 2/weighted by atomic polarizabilities; RDF080u – Radial Distribution Function – 8.0/unweighted; RDF010v – Radial Distribution Function – 1.0/weighted by atomic van der Waals volumes; Hypertens-80 - Ghose Viswanadhan-Wendoloski antihypertensive-like index at 80%; 6BELp2 – Lowest eigenvalue n. 2 of Burden matrix/weighted by atomic polarizabilities; RDF100u – Radial Distribution Function – 10.0/unweighted; E1s – 1st component accessibility directional WHIM index/weighted by atomic electrtopolotopical stes; R1v – R autocorrelation of lag 1/weighted by atomic van der Waals volumes; RDF145u – corresponds to: Radial Distribution Function – 14.5/unweighted; Mor16v, Mor19v – 3D-MoRSE – signal 16/19/weighted by atomic van der Waals volumes; G2m – 2st component symmetry directional WHIM index/weighted by atomic masses; H-048 – H attached to C2(sp3)/C1(sp2)/C0(sp).

Scoring function Autodock4 evaluates the free energy of ligand binding to the receptor in kcal/mol, smaller values correspond to more potent inhibitors.

In the table 2 ten compounds with the best affinity are present. We also show hydrogen bonds that were observed in the docking study with the residues of CK2 kinase.

For in vivo test on CK2 kinase we have selected two compounds. Namely, N-(2-fluorobenzyl)-2-((3-methyl-2-oxo-2H-[1,2,4]triazino[2,3-c]quinazolin-6-yl)thio)acetamide (MTB-67) with mean value of scoring function and 2-((2-oxo-3-phenyl-2H-[1,2,4]triazino[2,3-c]quinazolin-6-yl)thio)acetamide (MTB-13) with moderate scoring function. Second one turned to be quite active. Such, the percentage of rest of kinase activity, in concentration 33 μM is 80 and 4 respectively.

Cell growth percent of EKVX cell line according to the NCI protocol of N-(2-fluorobenzyl)-2-((3-methyl-2-oxo-2H-[1,2,4]triazino[2,3-c]quinazolin-6-yl)thio)acetamide is 89.22 and predicted by equation is 75.51. Such figures are comparable to measurements with CK2 kinase.

Table 2

| № Comp. (№ NCI) | Chemical formula | Scoring function | Hydrogen bond | cell line EKVX | Exp. | Pred. | cell line NCI-H522 | Exp. | Pred. |
|-----------------|------------------|------------------|---------------|---------------|------|-------|-------------------|------|-------|
| MTB-97 754975   |                  | -10,96           | LYS68, VAL116, ASN118 | 101,15 | 91,28 | 85,87 | 92,29 |
| MTB-100 754976  |                  | -10,53           | VAL116        | 77,80 | 73,11 | 41,24 | 56,87 |
| MTB-36 752628   |                  | -10,27           | VAL116       | 25,61 | 16,93 | -49,92 | 74,97 |
| MTB-62 753035   |                  | -10,22           | HIS160, ASP175, GLU114 | 97,36 | 104,32 | 104,33 | 76,90 |

© I. S. Nosulenko, O. Yu. Voskoboynik, O. M. Antypenko, G. G. Berest, S. I. Kovalenko, 2015
| № Comp. (№ NCI) | Chemical formula | Scoring function | Hydrogen bond | cell line EKVX | cell line NCI-H522 |
|-----------------|------------------|-----------------|---------------|----------------|-------------------|
| MTB-70 753040   | ![Chemical formula](image1) | -10,05          | VAL116        | 103,47         | 98,28             |
|                 |                  |                 |               | 83,69          | 89,31             |
| MTB-52 753027   | ![Chemical formula](image2) | -9,95           | GLU114        | 67,95          | 92,36             |
|                 |                  |                 |               | 91,22          | 73,05             |
| MTB-60 753033   | ![Chemical formula](image3) | -9,88           | HIS160, ASP175| 31,44          | 22,80             |
|                 |                  |                 |               | 62,5           | 67,99             |
| MTB-59 753032   | ![Chemical formula](image4) | -9,86           | ASP175, GLU114| 111,06         | 92,99             |
|                 |                  |                 |               | 92,27          | 69,22             |
| MTB-66 752624   | ![Chemical formula](image5) | -9,86           | GLU114        | 73,85          | 82,89             |
|                 |                  |                 |               | -              | 77,43             |
| MTB-67 753037   | ![Chemical formula](image6) | -9,77           | VAL116, ASN118, GLU114 | 89,22 | 75,51 |
|                 |                  |                 |               | 113,34         | 95,68             |
| MTB-128 754998  | ![Chemical formula](image7) | -9,30           | LYS68, VAL116, HIS160 | 95,89 | 99,58 |
|                 |                  |                 |               | 107,94         | 90,87             |
| MTB-121 (-)     | ![Chemical formula](image8) | -9,01           | LYS68, VAL116 | -              | 110,98            |
|                 |                  |                 |               | -              | 3,73              |
| MTB-123 754989  | ![Chemical formula](image9) | -8,63           | LYS68, VAL116 | 91,32          | 95,34             |
|                 |                  |                 |               | 77,43          | 86,13             |
| MTB-13 (-)      | ![Chemical formula](image10) | -8,13           | LYS68, ARG47  | -              | 80,48             |
|                 |                  |                 |               | -              | 118,55            |

© I. S. Nosulenko, O. Yu. Voskoboynik, O. M. Antypenko, G. G. Berest, S. I. Kovalenko, 2015
References
1. Barbosa, M. L., Lima, L. M., Tesch, R., Sant’Anna, C. M., Totzke, F., Kubbhat, M. H., et al. (2014). Novel 2-chloro-4-anilinoquinazoline derivatives as EGFR and VEGFR-2 dual inhibitors. *Eur. J. Med. Chem.*, 71, 1–14. doi: 10.1016/j.ejmech.2013.10.058.

2. Berest, G. G., Voskoboynik, O. Y., Kovalenko, S. I., Antypenko, O. M., Nosulenko, I. S., Katsen, A. M., & Standovskaya, O. S. (2011). Synthesis and biological activity of novel N-cycloalkyl-(cycloalkylkylaryl)-2-[(3-R-2-oxo-2-

3. Boyd, M. R. (1997). The NCI in vitro anticancer drug discovery screen. Concept, implementation and operation. *Humana Press*, 2, 23–43.

4. Cao, C., Albert, J. M., Geng, L., Ivy, P. S., Sandler, A., Johnson, D. H., & Lu, B. (2006). Vascular endothelial growth factor tyrosine kinase inhibitor azd2171 and fractionated radiotherapy in mouse models of lung cancer. *Cancer Res.*, 66(23), 11409–11415.

5. Carломагно, F., Vitagliano, D., Guida, T., Ciardiello, F., Tortora, G., Vecchio, G., et al. (2002). ZD6474, an orally available inhibitor of KDR tyrosine kinase activity, efficiently blocks oncogenic RET kinases. *Cancer Res.*, 62(24), 7284–7290.

6. Denny, W. A. (2001). The 4-anilinoquinazoline class of inhibitors of theerbB family of receptors. *Farmaco*, 56, 51–56.

7. Ferguson, A. D., Sheth, P. R., Basso, A. D., Paliwal, S., Gray, K., Fischmann, T. O., & Lee, H. V. (2011). Structural basis of CX-4945 binding to human protein kinase CK2. *FEBS Letters*, 585(1), 104–110. doi: 10.1016/j.febslet.2010.11.019.

8. Gramatica, P., Chirico, N., Papa, E., Cassani, S., & Kovarich, S. (2013). QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *Journal of Computational Chemistry, Software news and updates*, 34, 2121–2132. doi: 10.1002/jcc.23361.

9. Haste, C. J., McLaughlan, H. J., & Cohen, P. (2006). Assay of protein kinases using radiolabeled ATP: a protocol. *Nature Protocols*, 1(2), 968–971. doi: 10.1038/nprot.2006.149.

10. Kovalenko, S. I., Nosulenko, I. S., Voskoboynik, A. Y., Berest, G. G., Antypenko, L. N., Antypenko, A. N., & Katsev, A. M. (2012). Substituted 2-[(2-oxo-2H-[1,2,4]triazino[2,3-c][quinazolin-6-yl])thio]acetamides with thiazole and thiazolide fragments: synthesis, physicochemical properties, cytotoxicity, and anticancer activity. *Sci. Pharm.*, 80(4), 837–65. doi: 10.3797/scipharm.1208-07.

11. Kovalenko, S. I., Nosulenko, I. S., Voskoboynik, A. Y., Berest, G. G., Antypenko, L. N., Antypenko, A. N., & Katsev, A. M. (2013). Novel N-aryl(alkaryl)-2-[(3-R-2-oxo-2H-[1,2,4]triazino[2,3-c][quinazolin-6-yl])thio]acetamides: synthesis, cytotoxicity, anticancer activity, COMPARE analysis and docking. *Med. Chem. Res.*, 22, 2610–2632. doi: 10.1007/s00044-012-0257-x.

12. Marvin Sketch version: 6.3.0, (2014), Chem Axon. Retrieved from http://www.chemaxon.com.

13. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. & Olson, A. J. (2009). Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Computational Chemistry*, 16, 2785–2791.

14. Pedretti, A., Villa, L., & Vistoli, G. (2014). VEGA – An open platform to develop chemo-bio-informatics applications, using plug-in architecture and script” programming. *J.C.A.M.D.*, 18, 167–173.

15. Stewart, J. J. P. (2012). MOPAC2012. S.C. Chemistry. Retrieved from http://openmopac.net/MOPAC2012.html.

16. Todeschini, R., & Consonni, V. (2000). *Handbook of Molecular Descriptors*. Mannhold R., Kubinyi H., Timmerman H. (Eds.). Weinheim, New York.

17. Todeschini, R., & Consonni, V. (2003). DRAGON software for the Calculation of Molecular Descriptors, web version 3.0 for Windows.

18. Vangrevelinghe, E., Zimmermann, K., Schoepfer J., Portmann, R., Fabbro, D., & Furet, P. (2003). Discovery of a potent and selective protein kinase CK2 inhibitor by high-throughput docking. *J. Med. Chem.*, 46(13), 2656–2662. doi: 10.1021/jm030827c.