



Support Vector Machine Classification of Streptavidin-Binding Aptamers

Xinliang Yu^{1,2*}, Yixiong Yu³, Qun Zeng⁴

1 College of Chemistry and Chemical Engineering, Hunan Institute of Engineering, Xiangtan, Hunan, China, **2** State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha, Hunan, China, **3** G1302, Lushan International Experimental School, Changsha, Hunan, China, **4** Department of Neurological Surgery, Xiangtan Central Hospital, Xiangtan, Hunan, China

Abstract

Background: Synthesizing and characterizing aptamers with high affinity and specificity have been extensively carried out for analytical and biomedical applications. Few publications can be found that describe structure–activity relationships (SARs) of candidate aptamer sequences.

Methodology: This paper reports pattern recognition with support vector machine (SVM) classification techniques for the identification of streptavidin-binding aptamers as “low” or “high” affinity aptamers. The SVM parameters C and γ were optimized using genetic algorithms. Four descriptors, the topological descriptor $PW4$ (path/walk 4 - Randic shape index), the connectivity index $X3A$ (average connectivity index chi-3), the topological charge index $JG12$ (mean topological charge index of order 2), and the free energy E of the secondary structure, were used to describe the structures of candidate aptamer sequences from SELEX selection (Schütze *et al.* (2011) PLoS ONE (12):e29604).

Conclusions: The predicted fractions of winning streptavidin-binding aptamers for ten rounds of SELEX conform to the aptamer evolutionary principles of SELEX-based screening. The feasibility of applying pattern recognition based on SVM and genetic algorithms for streptavidin-binding aptamers has been demonstrated.

Citation: Yu X, Yu Y, Zeng Q (2014) Support Vector Machine Classification of Streptavidin-Binding Aptamers. PLoS ONE 9(6): e99964. doi:10.1371/journal.pone.0099964

Editor: Mark Isalan, Imperial College London, United Kingdom

Received: February 28, 2014; **Accepted:** May 21, 2014; **Published:** June 13, 2014

Copyright: © 2014 Yu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Hunan Provincial Natural Science Foundation of China (Contract No. 12JJ6011). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: yxliang5602@sina.com.cn

Introduction

Aptamers are structured single-stranded oligonucleotides that show an affinity toward a variety of targets, including proteins, viruses, and whole cells [1,2]. Compared to antibodies, aptamers are economical and easy to synthesize and modify, possess long-term stability, and display low immunogenicity, fast blood clearance, rapid tissue and tumor penetration. Furthermore, unlike antibodies developed *in vivo*, aptamers can be developed *in vitro* [3–5]. Aptamers, a promising class of compounds, both for target recognition and therapy, can be derived from a process termed Systematic Evolution of Ligands by EXponential enrichment (SELEX). This is a reiterative process of partitioning of aptamer candidates from non-binding sequences by an affinity method, followed with amplification of the bound sequences by polymerase chain reaction (PCR) [5].

The conventional SELEX method usually takes 10–15 cycles of selection and amplification, which are labor-intensive and time-consuming [6]. Furthermore, the consumption of samples/reagents is relatively high. By applying the relationship between the chemical structure of a molecule and its biological activity or properties, i.e., the structure–activity relationship (SAR) model, pattern recognition can be used to screen a series of candidate molecules, including those not yet synthesized, on the computer in order to select the structures having the desired set of predicted

activities/properties [7]. It is then possible to select the most promising candidates for synthesis, laboratory testing, and optimization. Thus, the pattern recognition approach based on such SARs as classification models can conserve resources and accelerates the process of selecting candidates for any purpose.

Streptavidin is widely used as a detection tool in biology research because of the strongest non-covalent interaction known in nature between streptavidin and biotin [8,9]. In addition, biotin-streptavidin system can be explored for detection of infection and tumor in clinical medicine [10]. Support vector machines (SVMs) are a popular technique for classification. The aim of this paper is to develop a pattern recognition model for aptamers against streptavidin, using SVM as the classification technique. A genetic algorithm is employed to find suitable parameters for SVM model that has relatively optimal prediction performance.

Materials and Methods

Data set

Table S1 shows the candidate aptamer sequences, which were taken from Reference [9]. These candidate aptamers were obtained through ten rounds of SELEX selection and the 100 most frequent clones of each round were listed. Each of the clones

contains 40 bases with a tolerance of two bases. The candidate aptamers were split into two sets: a training set and a test set. The training set includes the sequences from the 1st and 10th rounds of SELEX selection. The sequence R10#86 in the 10th round recurs in the 1st selection round, i.e., R1#1. In addition, the sequences R1#11 and R1#60 have no pair structure. Therefore, the three sequence, R10#86, R1#11 and R1#60 are removed from the data set of the 1st round selection. Totally, 197 different sequences were obtained to generate the training set. The test set contains the sequences from the rest of SELEX rounds, i.e., the 2nd, 3rd, 4th, 5th, 6th, 7th, 8th, and 9th rounds. Similarly, the sequences R2#55, R2#63 and R2#94 without pair structure are deleted from the test set.

In general, the binding affinity of aptamers with targets increases exponentially with increasing selection rounds. Such exponential increase in binding affinity is obvious during the first few rounds. In the end, the affinity is close to a saturation point after a certain number of rounds are carried out, and subsequently the binding affinity does not increase obviously [11–13]. Thus, the class labels (or target values) of sequences from the 1st round of SELEX were set as 1, denoting the low affinity and specificity aptamer candidates. The class labels of sequences in the 10th round were defined as 2, denoting the high affinity and specificity aptamer candidates. The training set was used to train and optimize algorithm parameters of SVM models. The test set was used later to evaluate predictive performance of the developed model.

Molecular descriptor calculation

The RNAstructure package (version 5.3) was used for prediction of secondary structures of candidate aptamers by minimizing the free energy (E) [14]. The loop structures were adopted to calculate molecular descriptors for corresponding candidate aptamers. Since the size of the loop is important for binding, the priority was given to the loops with 5–7 nucleotides, which were selected to calculate descriptors. Besides the free energy descriptor (E), three groups of molecular descriptors were calculated with Dragon software [15], which are 119 topological descriptors, 21 topological charge indices, and 33 connectivity indices. Totally, 174 molecular descriptors were calculated for each sequence. To calculate molecular descriptors, the loop of each aptamer candidate was sketched using ChemBioDraw Ultra 11.0 [16], and optimized using molecular mechanics (MM2) in ChemBio3D Ultra 11.0 until the ms of gradient value became smaller than 0.1 kcal/mol Å. The energy minimized molecules were then used as the inputs for Dragon software [15].

Topological descriptors based on a graph representation of a molecule can describe one or more such chemically interesting features as size, shape, symmetry, branching and cyclicality. They can also reflect chemical information like atom type and bonding environments [15]. Connectivity indices are calculated from the non-hydrogen part of a molecule where each vertex (non-hydrogen atom) is weighted by the vertex degree, i.e., the number of connected non-hydrogen atoms [17]. Topological charge descriptors are derived from an unsymmetrical matrix CT , whose single element CT_{ij} is equal to the vertex degree δ_i of the i th atom under the condition of $i = j$. Otherwise, CT_{ij} equals to the difference of m_{ij} and m_{ji} . Here m_{ij} and m_{ji} are elements of the matrix obtained by multiplying the adjacency matrix by the reciprocal square distance matrix [15,18]. For each path of length k , a topological charge index GGk is defined as the half-sum of all charge terms CT_{ij} (absolute values) corresponding to pair of vertices with topological distance equal to k [15,18].

Principle of support vector machine classification

SVMs are known as maximum-margin classifiers, since they find the optimal hyperplane that is equidistant from two classes and defined by a number of support vectors. In general, the larger the margin or distance between these parallel hyperplanes, the smaller the generalization error of the classifier will be [19–22].

Let $(\mathbf{x}_i; y_i)$ be a set of training examples, where $i = 1, 2, \dots, n$, $\mathbf{x}_i \in \mathbb{R}^d$ is an input vector, $y_i \in \{-1, 1\}$ is its corresponding desired output, i.e., a constant denoting the class to which that point \mathbf{x}_i belongs, n is the number of training data, and d denotes the number of dimensions of input data.

The SVM requires the solution of the following optimization problem

$$\min_{\mathbf{w}, b, \xi} J(\mathbf{w}, \xi, b) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_i \xi_i \quad (1)$$

subject to:

$$y_i(\mathbf{w}^T \phi(x_i) + b) \geq 1 - \xi_i \quad (2)$$

$$\xi_i \geq 0 \quad (3)$$

Here \mathbf{w} is the weight vector, b is the bias, ξ is a non-negative slack variable for the data points, and C is a penalty factor that controls the tradeoff between the complexity of the decision function and the number of training examples misclassified. SVM maps input vectors \mathbf{x}_i into a higher (may be infinite) dimensional space, where a margin hyperplane with the maximal margin is constructed. Under constraints $\sum_i \alpha_i y_i = 0$ and $0 \leq \alpha_i \leq C$, the optimization problem becomes

$$\min W(\alpha) = - \sum_i \alpha_i + \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j (x_i, x_j) \quad (4)$$

Quadratic programming method can be adopted to solve the above extreme problem. The points \mathbf{x}_i with $\alpha_i > 0$ are called support vectors. Patterns with $0 < \alpha_i < C$ are called unbounded support vectors, while those with $\alpha_i = C$ are called bounded support vectors [22]. SVM can be easily generalized to non-linear decision surfaces by replacing the inner product $(\mathbf{x}_i, \mathbf{x}_j)$ with a kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$. The Gaussian radial basis function kernel (RBF) is a popular kernel function used in SVM classification and can be expressed with

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (5)$$

Here γ is a kernel parameter. For the SVM classification models based on the RBF kernel, two parameters, C and γ , should be carefully tuned to the problem at hand. If the factor C is too large, a large penalty is assigned to non-separable points, which leads to store too many support vectors and thus over fit. On the other hand, if C is too small, an under fitting can occur. The γ parameter specifies the radius of the RBF, also exerting a strong impact on model performance [23].

In this paper, we used the genetic algorithm to optimize the SVM parameters, C and γ . Genetic algorithm belongs to the

family of evolutionary algorithms, which generate solutions to optimization problems using techniques inspired by the evolution of species emphasizing the law of survival of the strongest [24]. It uses random mutation, crossover and selection procedures to breed better models or solutions from an originally random starting population or sample [25]

The parameter values used in our experiments were set as following: the population size of genetic algorithm being 20, evolutionary generation being 200, and both the SVM parameter, C and γ , being selected in the ranges [1, 1000]. The 5-fold cross validation procedure was carried out for the training set during the optimization of SVM parameters, C and γ . LibSVM package [26] was used to develop the SVM models.

Results and Discussion

The selection of appropriate theoretical descriptors is crucial to obtain good classification performance. With the number of selection rounds being the dependant variable and 174 molecular descriptors being the independent variables, the correlation between independent and dependent variables was analyzed with stepwise multiple linear regression (MLR) in SPSS 11.5, to select an optimal subset of variables used as the inputs of the SVM model. Four descriptors, the topological charge index $JGI2$ (mean topological charge index of order 2), the topological descriptor $PW4$ (path/walk 4 - Randic shape index), the connectivity index $X3A$ (average connectivity index chi-3), and the free energy (E) of the secondary structure were obtained to describe the structure features of each candidate aptamer, which are listed in Table S2.

Molecular connectivity indices are used widely in various areas of physical, chemistry, biology, pharmacology, polymer, and environmental science. One of the most important reasons about their successful applications is that these indices are based on sound chemical, structural (topologic and geometrical), and mathematical grounds. The descriptor $X3A$ (a connectivity index; average connectivity index chi-3) belongs to average connectivity indices XkA , which are obtained by dividing each connectivity index by the number of paths involved in its calculation [15]

$$XkA = \sum_{k=1}^k \left(\prod_{a=1}^n \delta_a \right)^{-1/2} \quad (6)$$

Where δ_a represents the corresponding vertex degree, k is the total number of m th order subgraphs, and n is the number of vertices in the subgraph. The descriptor $X3A$ ($k=3$ for XkA) relates the characteristic dimension of the molecule to the atomic parameters (quantum number, bond indexes, etc.) [27]. $X3A$ can denote the molecular size and the electronic distribution of a loop, which is related to the induced fitting behavior and molecular recognition of an aptamer.

Mean topological charge indices $JGIk$ (from order 1 to order 10) are obtained by dividing the corresponding topological charge index ($GGIk$) by the total number of summation terms in $GGIk$. The mean topological charge index $JGI2$ ($k=2$ for $JGIk$) is related to the topological valence of the atoms and the net charge transferred from the atom j to the atom i [15].

Path/walk Randic shape indices (PWk) are calculated by summing the ratios of the atomic path count over the atomic walk count of the same order k and then dividing by the number of non-H atoms [15,28]. DRAGON calculates path/walk shape indices from order 2 up to 5. The index of first order is not provided as the counts of the paths and walks of length one are

equal and, therefore, the corresponding molecular index always equals one for all molecules [15]. Since path/walk count ratio is independent of molecular size, the topological descriptor $PW4$ (path/walk 4 - Randic shape index) can be considered as a shape descriptor.

The free energy E of an aptamer secondary structure is approximated as the sum of individual contributions from loops, stacked base pairs, and other secondary structure elements. Aptamer molecules fold by intramolecular base pairs and are stabilized by hydrogen bonds that form between the base pairs along the DNA or RNA molecule. In addition, base pair stacking in a helix also stabilizes the molecule and decreases the free energy of the folded aptamer. Thus the free energy (E) of the secondary structure reflects the conformational stability of an aptamer. Generally, an aptamer with a high free energy E (absolute value) does not mean a stronger binding with its targets. Because the binding is an aptamer-target interaction provided by different intermolecular interactions such as electrostatic interactions between charged groups, stacking of aromatic structures contained in organic compounds and the nucleobases, hydrogen bonds, and the complementary in three-dimensional shape [29], while the free energy here is the property of an aptamer only. For streptavidin-binding aptamers, a longer stem of an aptamer secondary structure leads to a larger descriptor E (absolute value). On the other hand, a longer stem can more effectively maintain the loop and bulge structures for binding [8]. Therefore, for a streptavidin DNA aptamer, its free energy E is correlated with its binding ability to streptavidin. Figure 1 shows the correlation of the mean free energy (E) of the secondary structure and the number of rounds of SELEX [9]. As can be seen, the descriptor E decreases with the increasing number of performed rounds.

Selecting appropriate values for parameters, C and γ , is also important for SVM performance. The optimization results show that the relatively optimal SVM model possesses parameters C of 705.933 and γ of 749.802. The model based on $C=705.933$ and $\gamma=749.802$ has classification accuracy of 97.98% for the training set. To evaluate the model, we calculated prediction values for the test set, which are listed in Table S2. In Reference [9], at identical concentration (1 mM), the binding affinity of eight sequences of R10#1, R10#2, R10#4, R10#6, R10#10, R10#17, R10#62, and R10#86, were studied. Only the sequence R10#62 has low

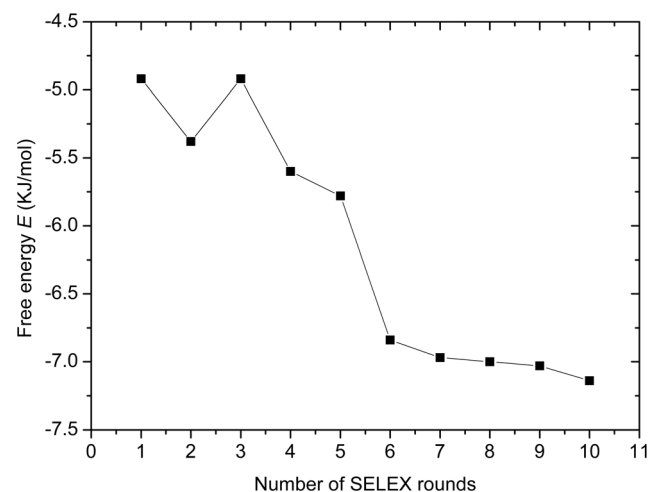


Figure 1. Correlation of the mean free energy and the number of performed rounds.

doi:10.1371/journal.pone.0099964.g001

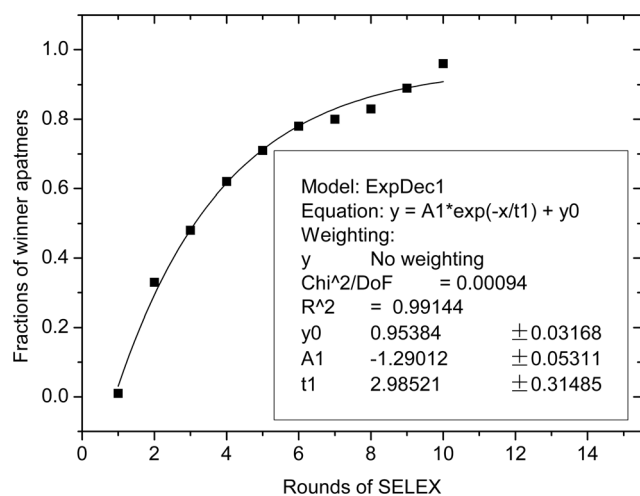


Figure 2. Evolutionary curve of streptavidin-binding aptamers.
doi:10.1371/journal.pone.0099964.g002

prediction class labels being 1, which is acceptable since R10#62 really has a weak binding ability to streptavidin [9].

For *in vitro* DNA aptamer selection procedures based on SELEX technology, starting point of each SELEX process is a synthetic random DNA oligonucleotide library consisting of a multitude of single-stranded DNA fragments (approximately 10^{15}) with different sequences. After the library and the target molecules are incubated for binding, unbound oligonucleotides are removed by several stringent washing steps of the binding complexes. The target-bound oligonucleotides being eluted and subsequently amplified by PCR are then used to generate a new enriched pool for the next selection round. During the SELEX process, the average binding affinity of the selected sequence against specific targets can increase exponentially with the number of selection rounds, just as the word “exponential” suggests in the term of Systematic Evolution of Ligands by EXponential enrichment (SELEX) [11–13]. The fractions of winning aptamers (i.e. their prediction class labels being 2) from the first round to 10th rounds of experiments are 0.01, 0.33, 0.48, 0.62, 0.71, 0.78, 0.80, 0.83, 0.89, and 0.96, respectively. We can obtain the following fitting curve and the corresponding exponential equation in Figure 2,

References

- Wilson DS, Szostak JW (1999) In vitro selection of functional nucleic acids. *Annu Rev Biochem* 68: 611–647. doi: 10.1146/annurev.biochem.68.1.611
- Bunka DH, Stockley PG (2006) Aptamers come of age – at last. *Nat Rev Microbiol* 4: 588–596. doi: 10.1038/nrmicro1458
- Hwang DW, Ko HY, Lee JH, Kang H, Ryu SH, et al. (2010) A nucleolin-targeted multimodal nanoparticle imaging probe for tracking cancer cells using an aptamer. *J Nucl Med* 51: 98–105. doi: 10.2967/jnumed.109.069880
- Shangguan D, Meng L, Cao ZC, Xiao Z, Fang X, et al. (2008) Identification of liver cancer-specific aptamers using whole live cells. *Anal Chem* 80: 721–728. doi: 10.1021/ac701962v
- Tan WH, Donovan MJ, Jiang JH (2013) Aptamers from cell-based selection for bioanalytical applications. *Chem Rev* 113: 2842–2862. doi: 10.1021/cr300468w
- Levine HA, Nilsen-Hamilton M (2007) A mathematical analysis of SELEX. *Comput Biol Chem* 31: 11–35. doi: 10.1016/j.compbiolchem.2006.10.002
- Karelson M, Lobanov VS, Katritzky AR (1996) Quantum–chemical descriptors in QSAR/QSPR studies. 96: 1027–1043. doi: 10.1021/cr950202r
- Bing T, Yang XJ, Mei HC, Cao ZH, Shangguan D (2010) Conservative secondary structure motif of streptavidin-binding aptamers generated by different laboratories. *Bioorg Med Chem* 18: 1798–1805. doi: 10.1016/j.bmc.2010.01.054
- Schütze T, Wilhelm B, Greiner N, Braun H, Peter F, et al. (2011) Probing the SELEX Process with Next-Generation Sequencing. *PLoS ONE* 6: e29604. doi: 10.1371/journal.pone.0029604
- Yomogida K, Chou Y, Pang J, Baravati B, Maniack BJ, et al. (2012) Streptavidin suppresses T cell activation and inhibits IL-2 production and CD25 expression. *Cytokine* 58: 431–436. doi: 10.1016/j.cyto.2012.02.007
- Djordjevic M (2007) SELEX experiments: new prospects, applications and data analysis in inferring regulatory pathways. *Biomol Eng* 24: 179–189. doi: 10.1016/j.bioeng.2007.03.001
- Vant-Hull B, Payano-Baez A, Davis RH, Gold L (1998) The mathematics of SELEX against complex targets. *J Mol Biol* 278: 579–597. doi: 10.1006/jmbi.1998.1727
- Djordjevic M, Sengupta AM (2006) Quantitative modeling and data analysis of SELEX experiments. *Phys Biol* 3: 13–28. doi: 10.1088/1478-3975/3/1/002
- Reuter JS, Mathews DH (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* 11: 129–129. doi: 10.1186/1471-2105-11-129
- Talet SRL (2006) DRAGON for Widows (Software for the Calculation of Molecular Descriptors), Version 5.4. Milan, Italy.
- Cambridge Soft Inc (2008) ChemBioOffice Ultra Version 11.0. Cambridge, USA.
- Mandloi M, Sikarwar A, Sapre NS, Karmarkar S, Khadikar PV (2000) A comparative QSAR study using Wiener, Szeged, and molecular connectivity indices. *J Chem Inf Comput Sci* 40: 57–62. doi: 10.1021/ci980139h
- Gálvez J, Garcia-Domenech R, Salabert-Salvador MT, Soler R (1994) Charge indexes. New topological descriptors. *J Chem Inf Comput Sci* 34: 520–525. doi: 10.1021/ci00019a008

which shows the average binding affinity of the candidate sequences towards streptavidin increases exponentially during the first six rounds of SELEX (the experimental results also show that the binding signal increased strongly until round six [9]). After that, the affinity is close to the saturation point, the fractions of winning aptamers being 08, although the binding affinity increases gently in subsequent rounds. Obviously, the prediction result consists with the aptamer evolutionary principles of SELEX based screening, as stated above [11–13].

Conclusions

The four descriptors, *PW4*, *X3A*, *JGI2*, and *E*, reflecting the structures of candidate aptamer sequences, were used as the input variables of SVM model. Genetic algorithms were chosen to optimize the SVM parameters, *C* and γ . The relatively optimal SVM model with parameters *C* of 705.933 and γ of 749.802 has classification accuracy of 97.98% for the training set. Furthermore, the prediction fractions of winning aptamers from the 1st round to 10th round are 0.01, 0.33, 0.48, 0.62, 0.71, 0.78, 0.80, 0.83, 0.89, and 0.96, respectively. The prediction result consists with the aptamer evolutionary principles of SELEX based screening, which shows that pattern recognition for streptavidin-binding aptamers is successful. The investigation may encourage the application of pattern recognition methods to the designs of candidate aptamers.

Supporting Information

Table S1 List of 995 candidate aptamer sequences for streptavidin containing 40 bases with a tolerance of two bases.

(DOC)

Table S2 List of the topological charge index *JGI2*, topological descriptor *PW4*, connectivity index *X3A*, the free energy *E*, and predicted class labels for 995 candidate aptamer sequences.

(DOC)

Author Contributions

Conceived and designed the experiments: XY YY. Performed the experiments: XY YY. Analyzed the data: XY YY. Contributed reagents/materials/analysis tools: XY QZ. Wrote the paper: XY QZ.

19. Birzele F, Kramer S (2006) A new representation for protein secondary structure prediction based on frequent patterns. *Bioinformatics* 22: 2628–2634. doi: 10.1093/bioinformatics/btl453
20. Wang B, Chen JW, Li XH, Wang YN, Chen L, et al. (2009) Estimation of Soil organic carbon normalized sorption coefficient (K_{oc}) using least squares-support vector machine. *QSAR Comb Sci* 28:561–567. doi: 10.1002/qsar.200860065
21. Li S, Fedorowicz A, Andrew ME (2007) A new descriptor selection scheme for SVM in unbalanced class problem: a case study using skin sensitisation dataset. *SAR QSAR Environ Res* 18: 423–441. doi: 10.1080/10629360701428474
22. Afantiis A, Melagraki G, Sarimveis H, Koutentis PA, Igglessi-Markopoulou O, et al. (2010) A combined LS-SVM & MLR QSAR workflow for predicting the inhibition of CXCR3 receptor by quinazolinone analogs. *Mol Divers* 14: 225–235. doi: 10.1007/s11030-009-9163-7
23. Plewczynski D, Tkacz A, Wyrwicz LS, Godzik A, Kloczkowski A, et al. (2006) Support-vector-machine classification of linear functional motifs in proteins. *J Mol Model* 12: 453–461. doi: 10.1007/s00894-005-0070-2
24. Yu X, Yu R (2013) Setschenow Constant Prediction Based on the IEF-PCM Calculations. *Ind Eng Chem Res* 52:11182–11188. doi: 10.1021/ie400001u
25. Turabekova MA, Rasulev BF (2004) A QSAR Toxicity Study of a Series of Alkaloids with the Lycotoxine Skeleton. *Molecules* 9: 1194–1207. doi: 10.3390/91201194
26. Chang CC, Lin CJ (2011) LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol* 2: 27:1–27: 27. doi: 10.1145/1961189.1961199
27. Petrova T, Rasulev BF, Leszczynski J, Toropov AA, Leszczynska D (2011) Improved model for fullerene C 60 solubility in organic solvents based on quantum-chemical and topological descriptors. *J Nanopart Res* 13: 3235–3247. doi: 10.1007/s11051-011-0238-x
28. Randić M (2001) Novel shape descriptors for molecular graphs. *J Chem Inf Comput Sci* 41: 607–613. doi: 10.1021/ci0001031
29. Nikolaus N, Strehlitz B (2014) DNA-Aptamers Binding Aminoglycoside Antibiotics. *Sensors*, 14(2): 3737–3755. doi:10.3390/s140203737