

RESEARCH ARTICLE

QSAR analysis of immune recognition for triazine herbicides based on immunoassay data for polyclonal and monoclonal antibodies

Andrey A. Buglak^{1,2*}, Anatoly V. Zherdev¹, Hong-Tao Lei³, Boris B. Dzantiev¹

1 A. N. Bach Institute of Biochemistry, Research Center of Biotechnology of the Russian Academy of Sciences, Moscow, Russia, **2** St. Petersburg State University, St. Petersburg, Russia, **3** Guangdong Provincial Key Laboratory of Food Quality and Safety, South China Agricultural University, Guangzhou, China

* andreybuglak@gmail.com

OPEN ACCESS

Citation: Buglak AA, Zherdev AV, Lei H-T, Dzantiev BB (2019) QSAR analysis of immune recognition for triazine herbicides based on immunoassay data for polyclonal and monoclonal antibodies. PLoS ONE 14(4): e0214879. <https://doi.org/10.1371/journal.pone.0214879>

Editor: Yoshihiro Uesawa, Meiji Pharmaceutical University, JAPAN

Received: December 18, 2018

Accepted: March 21, 2019

Published: April 3, 2019

Copyright: © 2019 Buglak et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The work was financially supported by the Russian Science Foundation (Project No. 14-16-00149).

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

Abstract

A common task in the immunodetection of structurally close compounds is to analyze the selectivity of immune recognition; it is required to understand the regularities of immune recognition and to elucidate the basic structural elements which provide it. Triazines are compounds of particular interest for such research due to their high variability and the necessity of their monitoring to provide safety for agricultural products and foodstuffs. We evaluated the binding of 20 triazines with polyclonal (pAb) and monoclonal (mAb) antibodies obtained using atrazine as the immunogenic hapten. A total of over 3000 descriptors were used in the quantitative structure-activity relationship (QSAR) analysis of binding activities (pIC_{50}). A comparison of the two enzyme immunoassay systems showed that the system with pAb is much easier to describe using 2D QSAR methodology, while the system with mAb can be described using the 3D QSAR CoMFA. Thus, for the 3D QSAR model of the polyclonal antibodies, the main statistical parameter q^2 ('leave-many-out') is equal to 0.498, and for monoclonal antibodies, q^2 is equal to 0.566. Obviously, in the case of pAb, we deal with several targets, while in the case of mAb the target is one, and therefore it is easier to describe it using specific fields of molecular interactions distributed in space.

Introduction

Triazines are herbicides which are widely used in agriculture and may accumulate in soil, as well as in food products [1]. Triazines are bound in soil to solids as well as to dissolved fractions of humic and fulvic acids, which leads not only to the accumulation of triazines in soil but also to the contamination of surface and ground waters [2] since triazines are water soluble.

Triazines may undergo chemical transformations of both a biotic and abiotic nature. The parent compound atrazine is subjected to oxidation of the alkyl substituents, oxidative dealkylation, hydroxylation, and also to ring cleavage [3–5]. The derivatives of atrazine may be toxic

to a greater or lesser extent. Triazines may induce different physiological disorders in humans and animals such as immunity suppression [6] and birth defects [7]. Thus, contamination of the environment by triazine herbicides is a significant risk factor. For this reason, it is necessary to monitor the triazine levels in soil, water, agricultural products and foods.

The methods of triazine detection, such as NMR, HPLC, mass-spectrometry and others, are time- and labour-consuming. The immunoassay determination of triazines is much easier, less expensive, highly sensitive and rapid. Due to its selectivity and simplicity, the immunoassay has been used to detect a wide variety of environmental pollutants including triazines [8–10]. However, the significant variety of triazines and their derivatives makes the detailed analysis of the regulations of their immune recognition necessary.

The quantitative structure-activity relationship (QSAR) is widely used to study the immune recognition of different classes of toxic food contaminants and veterinary drugs, including pesticides, etc. [11–14]. The work by Yuan and co-authors reported an immunoassay analysis of triazines; however, on only a small set of 11 compounds [9]. QSAR is used for the analysis of immunoassays on quinolones and fluoroquinolones [15–19], organophosphorus pesticides [20–21], phenylurea herbicides [22], and sulfonamides [23]. In this study, we used molecular modeling to study triazine recognition by monoclonal and polyclonal antibodies. Twenty triazines along with 2D and 3D-QSAR methodology were used to study the relationship between the antigen and antibody. The experimental data of microplate immunoenzyme assays with broad specificity for triazines were used from a classical work by A. Dankwardt and co-authors [24]. The comparative immunoassay of polyclonal and monoclonal antibodies used 4-arylamino-6-amino-1,3,5-triazines (Fig 1). Antibodies were grown using atrazine as an immunizing hapten.

The antibody-binding activity of triazines from Table 1, presented in the logarithmic form ($pIC_{50} = -\log IC_{50}$), was used as the dependent variable y during QSAR calculations. In Table 1, S2 stands for polyclonal sheep antibodies, while K4E7 stands for monoclonal antibodies.

Materials and methods

Conformational analysis and geometry optimization

The preparation of molecular geometries for 2D and 3D QSAR analysis was carried out in the Spartan v.16 program. A series of conformers for each of the 20 compounds was obtained by the systematic search method using the MMFF force field [25]. The optimization of the geometry for each conformer was carried out using the semi-empirical AM1 method [26]. In order

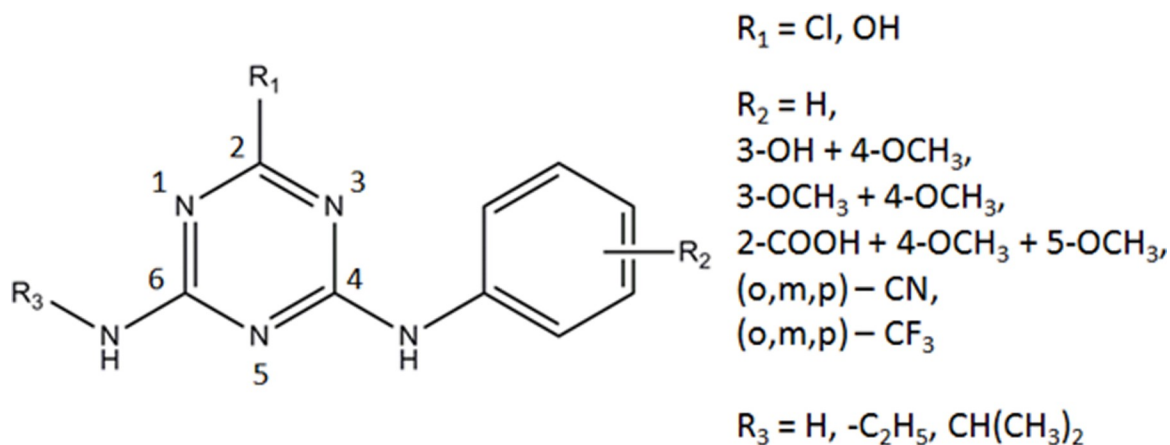


Fig 1. Molecular structure of arylamino-s-triazines (based on Dankwardt et al., 1996).

<https://doi.org/10.1371/journal.pone.0214879.g001>

Table 1. Cross-reactivity (CR) and concentration of 50% inhibition for the interaction of triazines with polyclonal (S2) and monoclonal antibodies (K4E7), according to Dankwardt et al., 1996.

№	Compound	CR, %		IC ₅₀ , nmol l ⁻¹	
		S2	K4E7	S2	K4E7
1	2-chloro-4-isopropylamino-6-ethylamino-1,3,5-triazine (atrazine)	100	100	0.93	0.37
2	2-chloro-4-isopropylamino-6-amino-1,3,5-triazine (deethylatrazine)	15	18	6.0	2.1
3	2-chloro-4-ethylamino-6-amino-1,3,5-triazine (deisopropylatrazine)	0.4	1.8	232	20.6
4	2-hydroxy-4-isopropylamino-6-ethylamino-1,3,5-triazine (hydroxyatrazine)	6	0.1	15.3	371
5	2-chloro-4-anilino-6-isopropylamino-1,3,5-triazine	106	140	0.88	0.27
6	2-chloro-4-(3',4'-dimethoxyanilino)-6-isopropylamino-1,3,5-triazine	126	140	0.74	0.27
7	2-chloro-4-(3',4'-dimethoxybenzylamino)-6-isopropylamino-1,3,5-triazine	109	148	0.84	0.25
8	2-chloro-4-(3'-hydroxy-4'-methoxyanilino)-6-isopropylamino-1,3,5-triazine	74	79	1.25	0.47
9	2-chloro-4-(2'-carboxy-4',5'-dimethoxyanilino)-6-isopropylamino-1,3,5-triazine	9	<0.1	10.3	>500
10	2-chloro-4-(2'-nitrilanilino)-6-isopropylamino-1,3,5-triazine	134	70	0.70	0.52
11	2-chloro-4-(3'-nitrilanilino)-6-isopropylamino-1,3,5-triazine	134	108	0.70	0.34
12	2-chloro-4-(4'-nitrilanilino)-6-isopropylamino-1,3,5-triazine	101	107	0.93	0.35
13	2-chloro-4-(2'-triflourmethylanilino)-6-isopropylamino-1,3,5-triazine	91	81	1.02	0.46
14	2-chloro-4-(3'-triflourmethylanilino)-6-isopropylamino-1,3,5-triazine	133	75	0.70	0.46
15	2-chloro-4-(4'-triflourmethylanilino)-6-isopropylamino-1,3,5-triazine	92	129	1.02	0.49
16	2-chloro-4-(3',4'-dimethoxyanilino)-6-ethylamino-1,3,5-triazine	5	10	18.6	3.68
17	2-hydroxy-4-anilino-6-isopropylamino-1,3,5-triazine	0.4	0.2	232	186
18	2-hydroxy-4-(3',4'-dimethoxyanilino)-6-isopropylamino-1,3,5-triazine	0.4	0.2	232	186
19	2-chloro-4-anilino-6-amino-1,3,5-triazine	0.1	<0.1	500	>500
20	2-chloro-4-(3',4'-dimethoxyanilino)-6-amino-1,3,5-triazine	<0.1	<0.1	>500	>500

<https://doi.org/10.1371/journal.pone.0214879.t001>

to confirm the correspondence between the obtained geometry and the minimum surface of the potential energy for each conformer, the Hessian was calculated. The conformers with the lowest total energy were used for 2D QSAR analysis.

2D QSAR

Using a random number generator, triazines were divided into two samples: the training set (80%) and test set (20%). The requirements for the maximum and minimum values in the test sample were as follows: 1) the maximum value of pIC₅₀ should be less than or equal to the maximum value of pIC₅₀ in the training sample; and 2) the minimum value of pIC₅₀ must be greater than or equal to the minimum value of pIC₅₀ in the training sample. The actual limit IC₅₀ values given in Table 1 are equal to 500 nmol l⁻¹, which were experimentally tested values in the original work by Dankwardt et al., 1996 [24]. When the activity IC₅₀ was equal to or greater than 500 nmol l⁻¹ the pIC₅₀ value of 6.30103 was used.

The required linear regression equation should have the following form:

$$y = a_1 * x_1 + \dots + a_n * x_n + c$$

where y is the dependent variable (pIC₅₀); a₁ and a_n are regression coefficients; x₁ and x_n are independent variables (descriptors); and c is a regression constant.

We estimated the contribution of a descriptor to the model using the following equation:

$$\alpha(x_1) = \frac{R^2(x_1, x_2, x_3) - R^2(x_2, x_3)}{3 \times R^2(x_1, x_2, x_3) - R^2(x_1, x_2) - R^2(x_1, x_3) - R^2(x_2, x_3)} \times 100\%$$

where α(x₁) is the relative contribution of the descriptor x₁ to the model with three descriptors;

$R^2(x_1, x_2, x_3)$ is the determination coefficient of the model with all three descriptors; and $R^2(x_2, x_3)$ is the determination coefficient of the model with two descriptors: x_2 and x_3 .

During the 2D QSAR analysis, different types of descriptors were used: constitutional (number of hydrogen bond donors, number of hydrogen bond acceptors, number of rings, number of chains, number of CH₃ groups, number of OH groups, etc.), electrostatic descriptors (maximum positive charge, maximum negative charge, the number of positively charged atoms, the number of negatively charged atoms, etc.), topological descriptors (kappa-indices describing the shape of the molecule, the indices of molecular bonds of Kier and Hall, etc.), 3D descriptors (volume, surface area, length and area of the projection on the coordinate axis, etc.), physico-chemical (lipophilicity (LogP), molecular refraction, polarizability, solubility in water, etc.) and quantum-chemical descriptors (energy of frontier molecular orbitals, electrostatic charges of atoms, the population of atoms according to Mulliken, etc.), amounting to a total of more than 3000 descriptors. The values of the descriptors were calculated using the program packages Spartan v. 16 and E-Dragon 1.0.

For the QSAR obtained model, the cross-validation was performed. Also, the models were tested for their predictive power using an external test set of compounds. The following statistical indicators were used:

- r^2 —coefficient of determination for the training sample [27];
- The determination coefficient R^2_{adj} , corrected for the number of descriptors involved in the model;
- q^2-r^2 from the results of the internal cross-validation of the training sample using the leave-one-out (LOO) method [28,29];
- LOF—error of Friedman approximation (Friedman Lack of Fit);
- RMSE—root-mean-square error;
- Max Error—the maximum prediction error for all the compounds in the test and training set;
- $pred_r^2-r^2$, which estimates the predictive ability of the model relative to the test set.

3D QSAR

Spatial alignment of molecules. The alignment of molecules plays a key role in 3D-QSAR. As a rule, the most energetically favorable conformation of the most active compound or the geometry corresponding directly to the interaction of the ligand and target (antibody), obtained on the basis of X-ray diffraction analysis or docking data, is chosen as the template for the alignment [30]. In our case, the most active conformation is unknown, and a number of the most active compounds (**10**, **14**) have a large number of conformers, which makes it difficult to analyze them correctly using quantum-chemical methods of required accuracy. Compounds **5**, **11** and **12** are also highly active, but with a small number of conformers compared to other triazines in the sample. All conformers of compounds **5** (12 conformers), **11** (23 conformers) and **12** (12 conformers) were optimized using the Hartree–Fock method and 6-31G(d) basis set (Fig 2). It was found that the most energetically favorable conformers of these compounds are identical, and therefore it can be assumed that this particular conformation is the most active and corresponds to the interaction with the antibody.

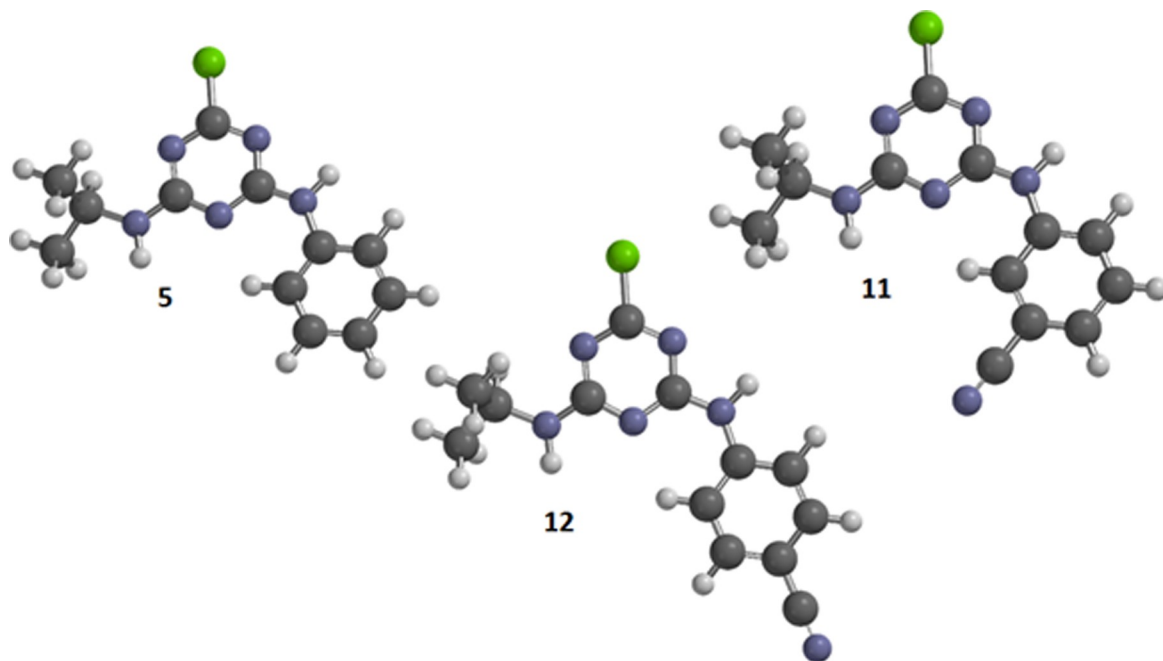


Fig 2. The most energetically favorable conformations of compounds 5, 11 and 12 according to the calculation by HF/6-31G(d) method.

<https://doi.org/10.1371/journal.pone.0214879.g002>

Triazine molecules in various conformations, optimized with the AM1 method, were superimposed on the template (compounds 5 and/or 11) using the algorithm presented in the open3DALIGN program [31]. Conformations corresponding to the largest overlap with the atoms of compounds 5 and 11 were selected (Fig 3).

The aligned geometries (S1 Fig) were used to create and analyze the fields of molecular interactions using the open3DQSAR program [32]. Visualization of 3D structures and contour maps of the force fields was carried out in the PyMol program.

Building of molecular interaction fields (MIF). Molecular interaction fields (MIF) were calculated to analyze the pIC_{50} values of triazines and to create CoMFA models for the S2 and K4E7 systems.

Two potentials were used to create the fields of molecular interaction: the steric potential in the form of the Lennard–Jones function 6–12 between the atoms of the molecule and the sp³ carbon atom:

$$E_{vdw} = \sum_{i=1}^n \left[\frac{A_i}{r_i^{12}} - \frac{B_i}{r_i^6} \right];$$

The electrostatic field was calculated by summing the Coulomb interactions between the test atom with charge +1 and triazine molecules:

$$E_{ele} = k \sum_{i=1}^n \left[\frac{q_i}{r_i^m} \right]$$

MIF preprocessing. The force-field data was preliminarily processed to remove non-informative x variables:

- 1) The values of the variables less than $0.05 \text{ kcal mol}^{-1}$ were equated to zero;

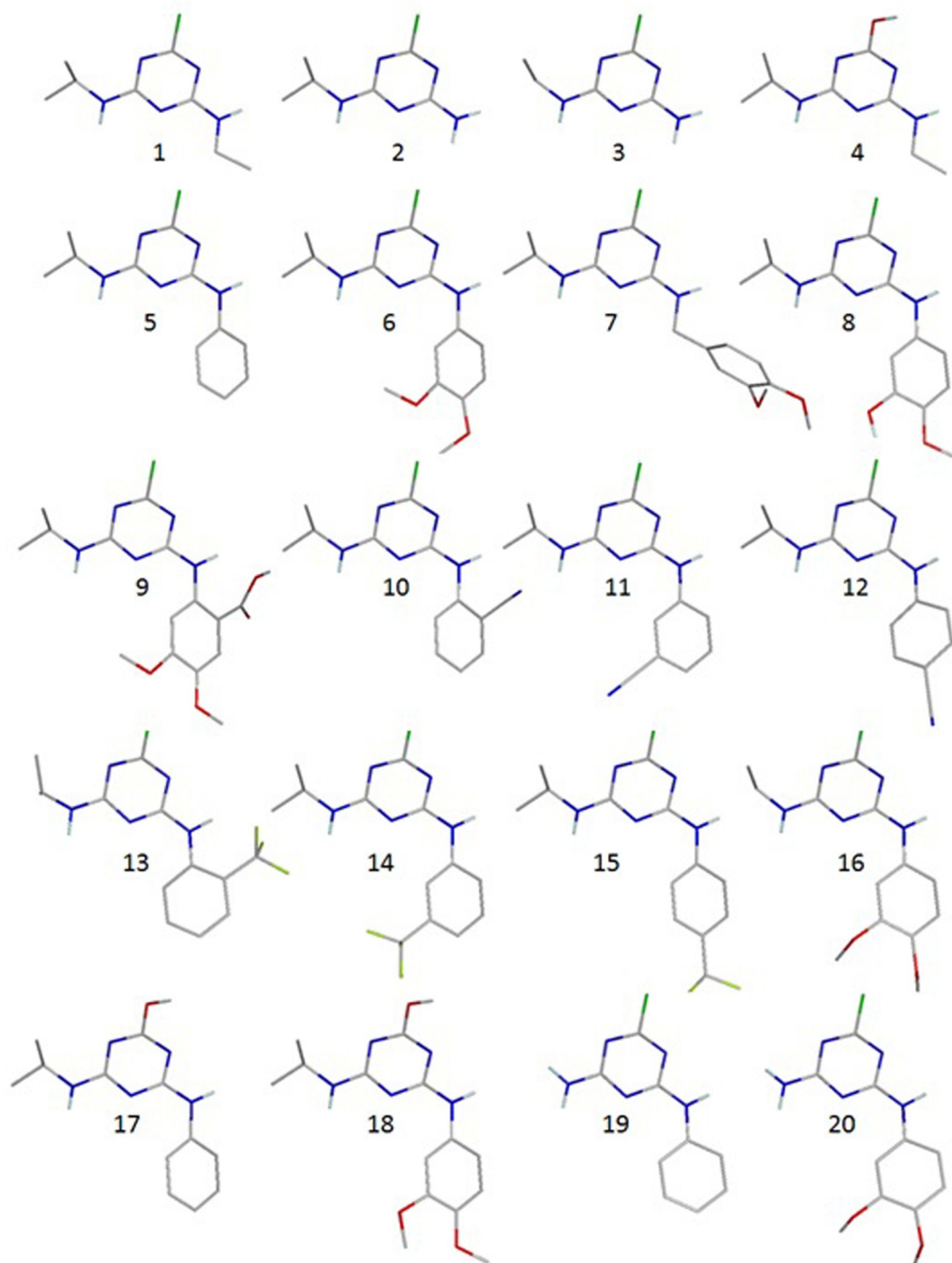


Fig 3. Triazine geometries optimized with the AM1 method (hydrogen atoms are represented only in -OH groups).

<https://doi.org/10.1371/journal.pone.0214879.g003>

- 2) Cut-off values were above 30 kcal mol⁻¹ and below -30 kcal mol⁻¹;
- 3) x variables with a standard deviation value below 0.1 were deleted;
- 4) Variables that had four or fewer nonzero values within the sample were removed.

Building of regression models. After all the preliminary operations were completed, we proceeded to construct the models using the partial least squares (PLS) method [33], based on the NIPALS algorithm (non-linear iterative partial least squares) [34]. The predictive power of the models was evaluated using the “leave-one-out” (LOO) cross-validation method:

$$q^2 = 1 - \frac{\sum (y_{obs} - y_{pred})^2}{\sum (y_{obs} - y_{mean})^2};$$

Also, we used the “leave-many-out” cross-validation method: a sample of 20 compounds was randomly broken into the training (75%) and test (25%) parts. The q^2 parameter was calculated for the test part. The partition procedure was repeated 50 times, and the average of the 50 partitioning variants was given the value q^2 (LMO).

To estimate the internal stability and predictive power of the 3D models, the standard prediction error (SDEP) parameter was calculated:

$$SDEP = \sqrt{\sum \frac{(y_{obs} - y_{pred})^2}{N}};$$

where y_{obs} is the experimental value; y_{pred} is the predicted value; y_{mean} is the mean value; and N is the number of molecules in the sample.

Procedures for selecting variables and methods for creating clusters of variables with common characteristics were used in order to increase the predictive ability of models: the SRD (smart region definition) method [35] groups variables based on their localization in three-dimensional space. This procedure reduces the number of descriptors bearing the same information; and fractional factorial design (FFD) [36,37] allows the selection of variables that have the greatest impact on the predictive ability of models. The FFD selection was conducted based on the leave-many-out cross analysis, using the SRD of the variable group.

Results

2D QSAR

S2 system analysis (polyclonal antibodies). For the S2 system, a series of multiple linear regression models was obtained:

Model 1

$$pIC50 = 23.68 + 0.985 * ES_Count_sssCH - 64.66 * Mulliken_Charge_CA + 0.056 * Solvation_E$$

Model 2

$$pIC50 = 12.43 + 4.393 * ES_Sum_sssCH - 107.3 * Jurs_FPSA_3$$

Model 3

$$pIC50 = 11.92 - 0.910 * CHI_2 + 1.932 * CHI_V_2 - 100.3 * Jurs_FPSA_3$$

Model 4

$$pIC50 = 8.807 + 5.366 * ES_Sum_sssCH + 0.828 * IC2 - 104.0 * Jurs_FPSA_3$$

Model 5

$$pIC50 = 6.302 + 6.182 * ES_Sum_sssCH + 0.946 * IC2 + 0.062 * Solvation_E$$

It is believed that QSAR is predictive if the following conditions are satisfied: $R^2 > 0.6$, $q^2 > 0.5$ and $pred_R^2 > 0.5$. Table 2 shows that all the presented models are statistically reliable and possess a predictive ability. Model 5 has the highest statistical parameters among the five models.

Let us consider the descriptors involved in Model 5. *Solvation_E* (relative contribution, $\alpha = 66.8\%$) is the solvation energy calculated with the AM1 method. Compound 18 has the highest solvation energy value ($-20.27 \text{ kJ mol}^{-1}$), while compound 13 has the lowest ($-71.22 \text{ kJ mol}^{-1}$) (S1 Table). The lower the solvation energy, the higher the solubility. However, the energy of solvation has a negative value, and therefore compounds that are less soluble in water have higher *pIC*₅₀ values.

ES_Sum_sssCH (26.1%) is the electrotopological index of methantriyl groups [38]. The *ES_Sum_sssCH* value is directly proportional to the activity of the molecule (1, 2 and similar compounds). *ES_Sum_sssCH* is equal to zero for compounds without >CH-groups. Obviously, the presence of the *ES_Sum_sssCH* descriptor in Model 5 indicates the participation of >CH-groups of triazine compounds in van der Waals interactions with antibodies.

IC2 (7.1%) is the information content index (neighborhood symmetry of 2-order). This descriptor makes it possible to estimate the degree of heterogeneity of the molecular structure. Compound 1 has the lowest value of *IC2* (3.431), and compound 9 has, on the contrary, the highest value of this descriptor (4.426). *IC2* can be considered as a measure of the complexity of the triazine topology. *IC2* makes a minor contribution to the model ($\alpha = 7.1\%$).

Experimental and predicted *pIC*₅₀ values for triazine interactions with polyclonal antibodies (S2) based on Model 5 are presented on Fig 4.

K4E7 system analysis (monoclonal antibodies). For the K4E7 system, a series of multiple linear regression models was obtained:

Model 6

$$pIC50 = 7.142 - 9.584 * Mor21v - 0.561 * NPlusO_Count + 33.83 * R6m+$$

Model 7

$$pIC50 = 5.859 - 0.018 * Jurs_TPSA - 7.118 * Mor21v + 27.20 * R6m+$$

Model 8

$$pIC50 = -3.941 + 1.291 * ES_Count_sssCH + 5.271 * IC2 - 0.216 * Jurs_DPSA_3$$

Model 9

$$pIC50 = 7.173 - 8.061 * Mor21v - 0.0385 * TPSA_NO + 34.56 * R6m+$$

Table 2. Statistical parameters of the models for the S2 system.

No	Statistical parameter	Model 1	Model 2	Model 3	Model 4	Model 5
1	R^2	0.941	0.890	0.937	0.937	0.945
2	R^2_{adj}	0.926	0.873	0.921	0.921	0.932
3	q^2	0.893	0.843	0.898	0.873	0.900
4	$pred_R^2$	0.913	0.924	0.870	0.920	0.942
5	Max error	0.567	0.904	0.746	0.549	0.454
6	RMS error	0.295	0.387	0.305	0.306	0.284
7	LOF	0.383	0.329	0.411	0.412	0.355

<https://doi.org/10.1371/journal.pone.0214879.t002>

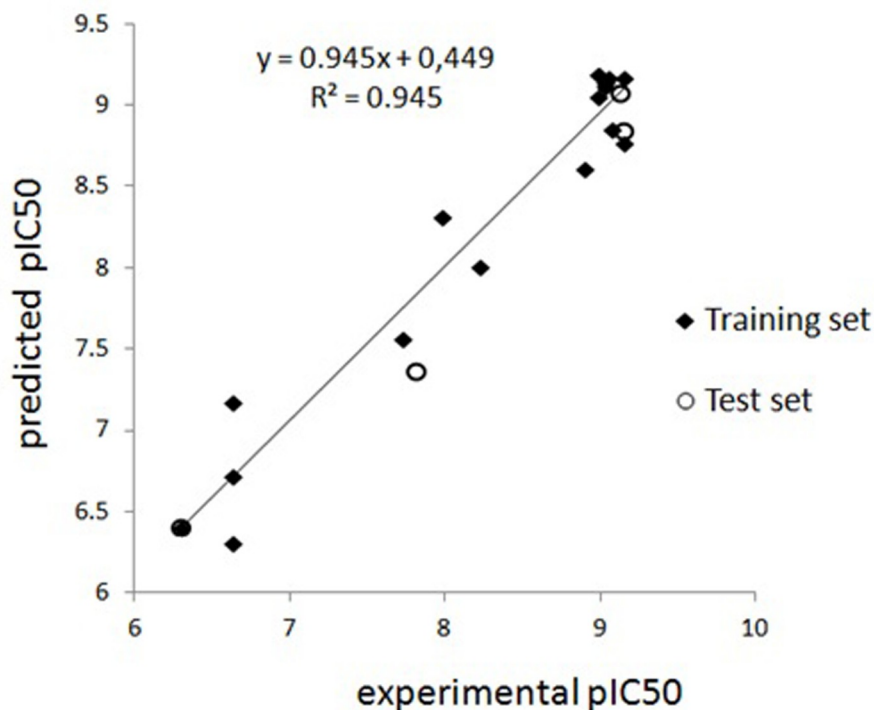


Fig 4. Comparison of experimental pIC_{50} values for triazine interactions with polyclonal antibodies (S2) and predicted activity based on Model 5; compounds 4, 6, 14 and 19 were used as a test set.

<https://doi.org/10.1371/journal.pone.0214879.g004>

Model 10

$$pIC_{50} = 7.369 + 1.637 * ES_Count_sssCH + 0.287 * ES_Sum_sCl - 0.791 * Num_H_Donors_Lipinski$$

Model 10 has a high predictive ability, and its statistical parameters are higher than for models 6–9 (Table 3).

Num_H_Donors_Lipinski ($\alpha = -37.9\%$) is the number of hydrogen bond donors (S2 Table). The value of the descriptor is inversely correlated with the activity: the greater the number of hydrogen bond donors, the lower the activity. Apparently, the presence of amine groups (-NH- and -NH₂) negatively affects the activity.

ES_Sum_sCl (34.6%) is the electrotopological index of chlorine atoms.³⁵ The *ES_Sum_sCl* value is directly proportional to the activity of the molecule. Apparently, the chlorine atom, as an electron-acceptor substituent, participates in electrostatic interactions.

Table 3. Statistical parameters of 2D QSAR models for the K4E7 system.

№	Statistical parameters	Model 6	Model 7	Model 8	Model 9	Model 10
1	R^2	0.921	0.915	0.917	0.925	0.966
2	R^2_{adj}	0.902	0.894	0.896	0.906	0.957
3	q^2	0.819	0.778	0.862	0.832	0.939
4	$pred_R^2$	0.767	0.788	0.898	0.704	0.969
5	Max error	0.947	1.110	1.065	0.913	0.469
6	RMS error	0.431	0.449	0.444	0.422	0.285
7	LOF	0.822	0.890	0.869	0.787	0.359

<https://doi.org/10.1371/journal.pone.0214879.t003>

ES_Count_sssCH (27.4%) is the number of methantriyl groups. Obviously, the presence of the *ES_Count_sssCH* descriptor in Model 10 indicates the participation of bulky substituents of triazine compounds in steric interactions with antibodies.

The experimental and predicted pIC_{50} values for triazine interactions with monoclonal antibodies (K4E7) based on Model 10 are presented in Fig 5.

3D QSAR

S2 system study. A 3D QSAR analysis was carried out for all 20 compounds. The interaction energy of the "probe" (carbon atom with charge +1) with the target molecule was calculated at each point of the regular 3D lattice. The dimensions of the 3D cubic grid were $26 \times 18 \times 22 \text{ \AA}$. The lattice spacing was 2 \AA .

One can consider Table 4 to compare the accuracy of the activity prediction for individual molecules within the model.

Van der Waals interactions make the main contribution to the interactions between the triazine and antibody in the S2 system in all three models (S3 Table).

To estimate the predictive ability of the 3D QSAR models, the following statistical parameters were used: the coefficient of determination (R^2), the mean-square error of the model (SDEC), the F-statistic value, the correlation coefficient of the leave-one-out (LOO) method ($q^2_{(LOO)}$), the correlation coefficient of the LMO method ($q^2_{(LMO)}$), and the mean-square error

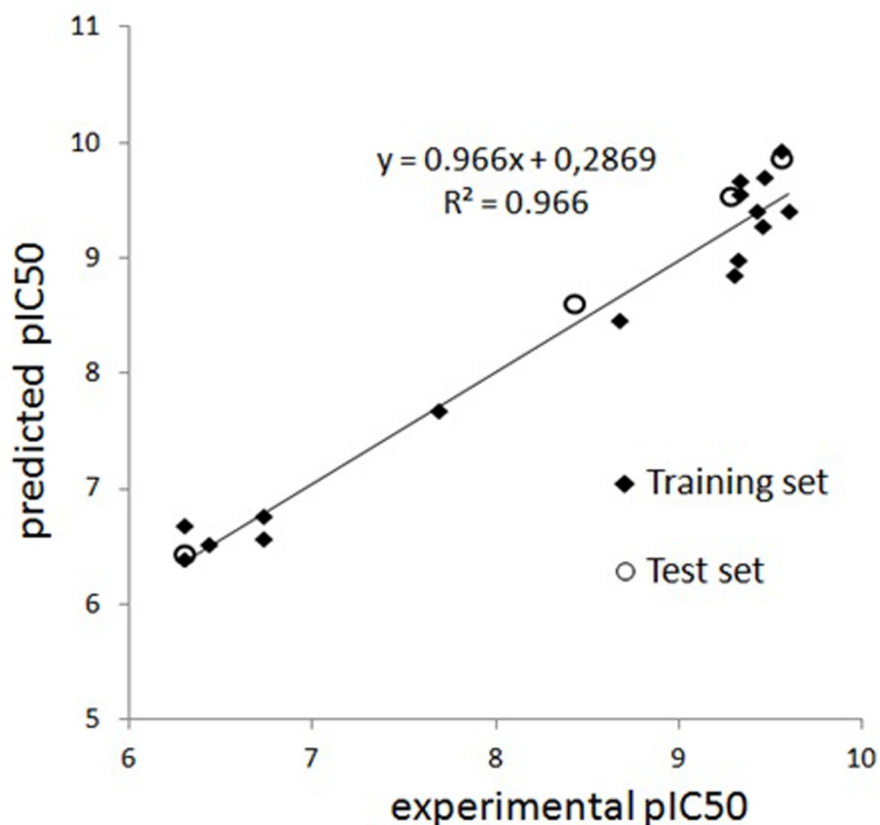


Fig 5. Comparison of experimental pIC_{50} values for triazine interactions with monoclonal antibodies (K4E7) and predicted activity based on Model 10; compounds 6, 10, 16 and 20 were used as a test set.

<https://doi.org/10.1371/journal.pone.0214879.g005>

Table 4. Experimental and predicted values of the triazine cross-reactivity logarithm (pIC₅₀) in the S2 system.

N ^o	Experiment	Prediction
1	9.032	9.056
2	8.222	8.405
3	6.635	6.846
4	7.815	7.297
5	9.056	8.641
6	9.131	8.763
7	9.076	9.393
8	8.903	8.713
9	7.987	7.878
10	9.155	9.492
11	9.155	9.276
12	9.032	8.811
13	8.991	9.147
14	9.155	9.108
15	8.991	9.016
16	7.731	7.406
17	6.635	7.039
18	6.635	6.950
19	6.301	6.529
20	6.301	6.170

<https://doi.org/10.1371/journal.pone.0214879.t004>

of prediction (SDEP). The S2 system 3D QSAE model showed satisfactory statistical results (Table 5).

As can be seen from Table 5, the model has high predictive ability. High coefficients of determination ($R^2 = 0.938$) and LOO cross-validation ($q^2 = 0.68$) indicate the statistical significance of the obtained model. The high correlation between predicted and experimental values of the model can be seen in Fig 6.

Molecular interaction field (MIF) contour maps were obtained to visualize the information about the 3D-QSAR models. Contour maps are presented in Fig 7. The steric fields presented by green contours are favorable for bulk substituents and have a positive effect on cross-reactivity (Fig 7A), while the yellow contours represent the areas where the presence of bulk substituents is unfavorable for high activity (Fig 7B). The electrostatic fields are represented by blue and red contours. The blue contour reflects the areas where positively charged atoms have a positive effect on activity (Fig 7C), while the red contours reflect the regions in which the presence of negatively charged atoms is favorable for high activity (Fig 7D).

K4E7 system study. A 3D QSAR analysis was carried out for all 20 compounds without splitting the sample into the training and test parts. The same alignment was used as for the S2 system.

The predicted values of triazine activity are shown in Table 6. The calculated values of the model have the best match with the experimental data. The level of correlation between the activities predicted by the CoMFA model and the experimental values is also shown in Fig 8.

Table 5. Statistical parameters of the CoMFA model for the S2 system. SDEP: mean-square error of prediction. LOO: leave-one-out.

N ^o	R ²	SDEC	F-test	q ² _(LOO)	SDEP _(LOO)	q ² _(LMO)	SDEP _(LMO)
3	0.938	0.269	80.39	0.680	0.610	0.498	0.754

<https://doi.org/10.1371/journal.pone.0214879.t005>

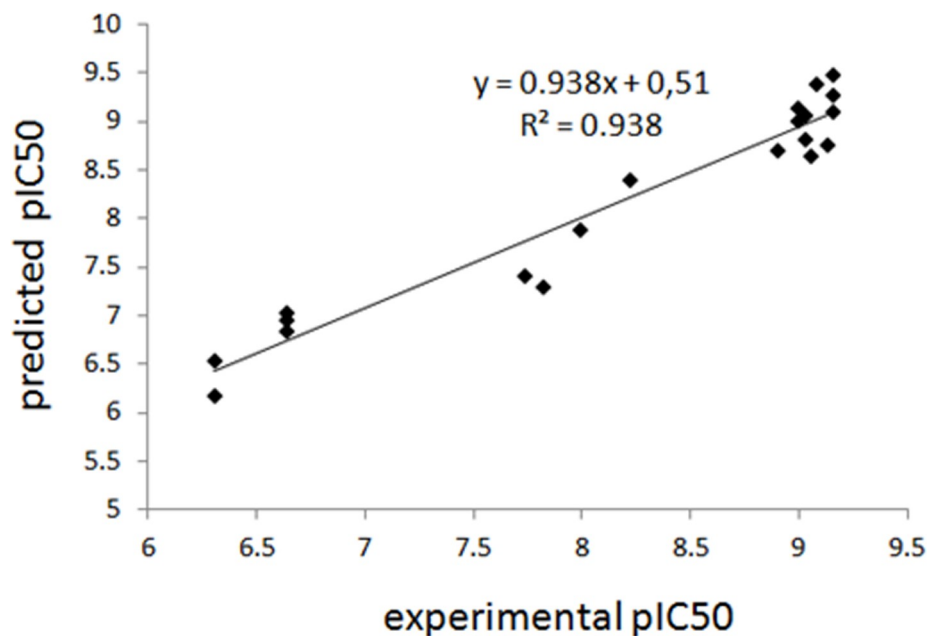


Fig 6. Experimental vs. predicted values of triazine cross-reactivity in the S2 system according to the CoMFA method (3D QSAR).

<https://doi.org/10.1371/journal.pone.0214879.g006>

Table 7 shows that the model has high values of the determination coefficient ($R^2 = 0.952$) and LOO cross-validation parameter ($q^2 = 0.637$), which indicates the statistical significance of the model obtained. The contour maps of the molecular interaction for this model are presented in Fig 9.

The contributions of the van der Waals and electrostatic interactions to the model are approximately equal at 51% and 49%, respectively (S4 Table).

Discussion

It seems reasonable to compare the most similar compounds with each other and find out which descriptors play a key role in their recognition.

Atrazine in the S2 system (polyclonal antibodies) exhibits significantly higher activity (CR 100%) than its metabolites: compounds 2, 3 and 4 (CR 15%, 0.4% and 6%, respectively). Atrazine has the highest value of the *Solvation_E* descriptor (see S1 Table). The value of the *Solvation_E* descriptor is directly proportional to pIC_{50} . The lower the *Solvation_E*, the higher the polarity of the molecule. For compounds 1–4, *Solvation_E* is equal to $-33.49 \text{ kJ mol}^{-1}$, $-49.34 \text{ kJ mol}^{-1}$, $-56.07 \text{ kJ mol}^{-1}$ and -58 kJ mol^{-1} , respectively. In general, this is consistent with our observation that the polarity of triazines is unfavorable for the interaction with the antibodies. Also, compounds with a CR higher than 1% (compounds 1, 2 and 4) possess *ES_Sum_sssCH* with a non-zero value, while the compound with the lowest *ES_Sum_sssCH* value (compound 3) has the lowest CR value— 0.4%.

For the K4E7 system (monoclonal antibodies), the highest pIC_{50} value of compound 1 among compounds 1–4 is determined by non-zero values of its descriptors *ES_Count_sssCH* and *ES_Sum_sCl*. Also, compound 1 has the lowest value of the descriptor *Num_H_Donors_Lipinski*. The descriptor *Num_H_Donors_Lipinski* is inversely correlated with pIC_{50} .

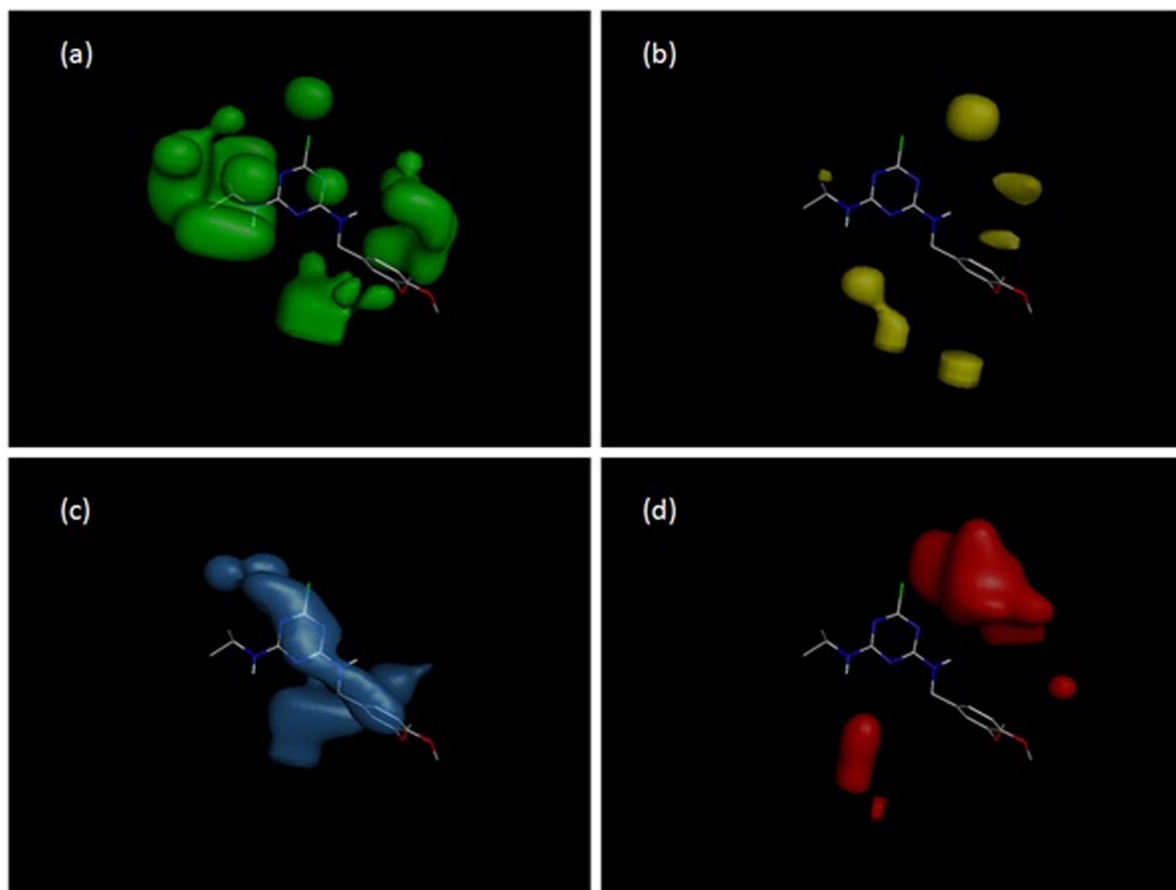


Fig 7. The molecular interaction field (MIF) contour maps for system S2 and compound 7: (a) favorable steric interactions; (b) unfavorable steric interactions; (c) sites favorable for positively charged groups; (d) sites favorable for negatively charged groups.

<https://doi.org/10.1371/journal.pone.0214879.g007>

For system S2, the highest CR value of compound **6** among compounds **6–9** (CR 126%, 109%, 74% and 9%, respectively) is determined by its lowest value of $E_{\text{Solvation}}$ ($-37.49 \text{ kJ mol}^{-1}$), while $E_{\text{Solvation}}$ is inversely correlated with CR.

In the case of the K4E7 system and compounds **6–9**, we should regard the CoMFA contour maps. The substituent at the C6 position does not differ in these compounds, and the variation of the substituent in the C4 position obviously plays the main role. The geometry of compound **7** is the most advantageous for the interaction with the antibody, and the geometry of **9** is the least advantageous (Table 1). The contour map of electrostatic charges (Fig 9D) indicates that the presence of a negatively charged carboxyl group at position 2' (compound **9**) is unfavorable for the binding with the antibody. Fig 9C, on the contrary, shows how the presence of negatively charged oxygen atoms of the methoxy groups of compound **7** has a positive effect on the activity. The contour map on Fig 9A shows that the aniline residue at the C4 position (namely, the benzene ring) generally has a positive effect on the activity of atrazines. Fig 9B shows that the bulky substituents localized at the meta position of the aniline have a negative effect on the activity; the methoxy substituents of compound **9** are located in these areas.

Compounds **10**, **11** and **12** have ortho-, meta- and para-positions of the nitrile of the aniline substituent, respectively (Fig 1). The contour maps of electrostatic interactions (Fig 7C and 7D) may explain why the activity of these compounds in system S2 is different. Fig 7D shows

Table 6. Experimental and predicted values of triazine cross-reactivity in the K4E7 system.

No	Experiment	Prediction
1	9.432	9.120
2	8.678	8.786
3	7.686	7.746
4	6.431	6.429
5	9.569	9.031
6	9.569	9.361
7	9.602	9.397
8	9.328	8.894
9	6.301	6.442
10	9.284	9.455
11	9.469	9.462
12	9.456	9.871
13	9.337	9.212
14	9.337	9.776
15	9.310	9.527
16	8.434	8.497
17	6.731	6.721
18	6.731	6.362
19	6.301	6.931
20	6.301	6.266

<https://doi.org/10.1371/journal.pone.0214879.t006>

that the contour map is favorable for the interaction with the nitrile located at the 2' and 3' positions of the aniline ring.

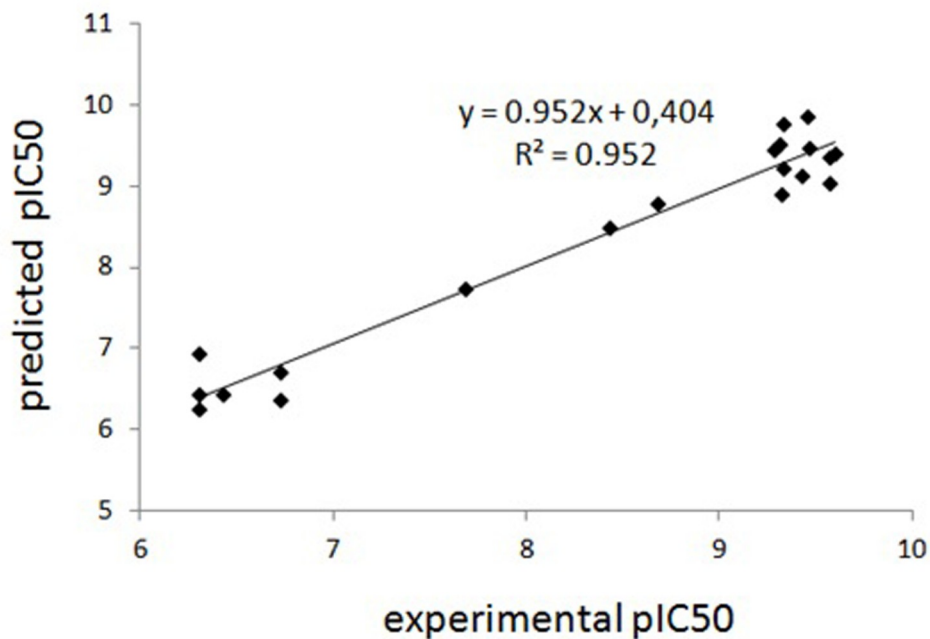


Fig 8. K4E7 system's experimental and predicted pIC_{50} triazine values based on the CoMFA method (3D QSAR).

<https://doi.org/10.1371/journal.pone.0214879.g008>

Table 7. Statistical parameters of the 3D QSAR model for the K4E7 system.

No	R ²	SDEC	F-test	q ² _(LOO)	SDEP _(LOO)	q ² _(LMO)	SDEP _(LMO)
3	0.952	0.291	105.0	0.637	0.796	0.566	0.861

<https://doi.org/10.1371/journal.pone.0214879.t007>

Let us regard compounds 13, 14 and 15, which possess trifluoromethyl substituent (-CF₃) in the ortho-, para- and meta- positions, respectively. In the S2 system, molecule 14 is more active than 13 and 15 (Table 1). In order to answer the question why these compounds exhibit different activities, it is required to consider the *Solvation_E* descriptor involved in 2D QSAR model 5. This descriptor is inversely correlated with pIC₅₀. For compounds 13, 14 and 15, *Solvation_E* is equal to -20.27 kJ mol⁻¹, -30.06 kJ mol⁻¹ and -27.35 kJ mol⁻¹, respectively. Obviously, the lower the solvation energy, the lower the activity. Also, the pIC₅₀ of compounds 13, 14 and 15 is correlated with IC₂ values of 4.15, 4.21 and 4.09, respectively. The higher value of IC₂ descriptor for compound 14 means a higher heterogeneity of 14 among compounds 13–15.

Obviously, the low activity of compounds 16, 19 and 20 in both systems (S2 and K4E7) is due to the absence of the >CH- group at the C6 position. This is confirmed both by the zero values of 2D descriptors *ES_Count_sssCH*, *ES_Sum_sssCH*, and by the contour maps of van der Waals interactions (Figs 7A and 9A). Also, it is quite obvious that compounds 17, 18 have low pIC₅₀ values due to the presence of the -OH group in the C2 position instead of chlorine.

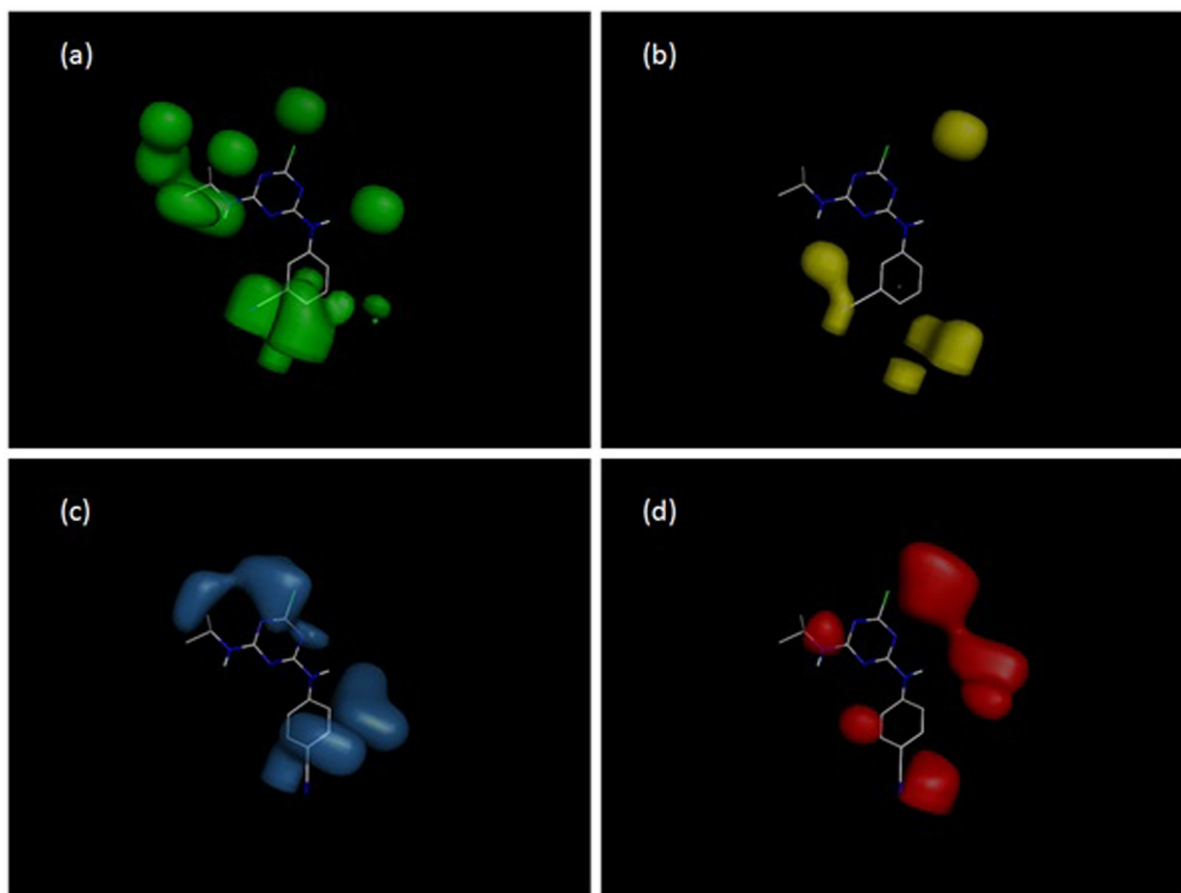


Fig 9. MIF contour maps for the K4E7 system: (a) favorable steric interactions and (b) unfavorable steric interactions for compound 11; (c) sites favorable for positively charged groups; (d) sites favorable for negatively charged groups for compound 12.

<https://doi.org/10.1371/journal.pone.0214879.g009>

If we compare the 2D QSAR models for monoclonal antibodies (mAb) and polyclonal antibodies (pAb), one can see that both models (Model 4 for mAb and Model 7 for pAb) contain the polar surface area parameters *Jurs_FPSA_3* and *Jurs_TPSA*, respectively. The polar surface area is inversely correlated with the activity, which means that the polarity of atrazines is unfavorable for the interaction with both monoclonal and polyclonal antibodies. Both mAb and pAb are affected by the amount of methantriyl groups (*ES_Sum_sssCH* and *ES_Count_sssCH* parameters, respectively), which means that, in both cases, bulk substituents are favorable for higher activity. The latter statement is supported by 3D molecular interaction field contour maps (Figs 7A and 9A). We may see that both mAb and pAb systems have a great deal in common.

Conclusions

In this study, we evaluated the efficiency of the interaction of 20 triazines with polyclonal and monoclonal antibodies grown using atrazine as the immunizing hapten.

The comparison of the two immunoassay systems showed that the system with the polyclonal antibodies (S2) is much easier to describe using 2D QSAR methodology, and the system with monoclonal antibodies can be described using the 3D QSAR CoMFA method. Thus, for the 3D QSAR model of the S2 system, the main statistical parameter q^2 (LMO) is equal to 0.498, and for K4E7, q^2 (LMO) = 0.566. Apparently, this is explained by the fact that in the case of polyclonal antibodies, we deal with several targets, while in interaction with monoclonal antibodies, the target is one, and therefore it is easier to describe it using specific fields of molecular interactions distributed in space.

There is an interesting fact: while for the S2 system the main contribution (72%) is made by steric interactions (S3 Table), for K4E7, the fraction of van der Waals and electrostatic interactions is equal (S4 Table).

Based on our analysis, we conclude that the effectiveness of the interaction with polyclonal antibodies significantly depends on the presence of the isopropylamino group in the C6 position of triazine, the presence of chlorine at the C2 position, and also the aniline substituent in the C4 position, whose substituents, in turn, participate in electrostatic interactions and hydrogen bonding.

The obtained data as well as the earlier-published results of QSAR technique application for immune recognition demonstrate the efficiency of this technique in the identification of key structures responsible for the distinguishing of structurally close antigens by antibodies. This information allows the theoretical formulation of criteria for molecules that could be recognized by the same antibodies (thus, extending the selectivity of one immunoassay) or requirements in different antibodies for detection (thus, realizing assays with mixtures or arrays of immunoreactants).

Supporting information

S1 Fig. Geometries of 20 triazine molecules aligned using compound 11 as a template.
(PDF)

S1 Table. The values of molecular descriptors from Model 5.
(PDF)

S2 Table. Values of molecular descriptors from Model 10.
(PDF)

S3 Table. Relative contribution (in %) of van der Waals forces and electrostatic interactions in 3D QSAR models for system S2.
(PDF)

S4 Table. Relative contribution (in %) of van der Waals forces and electrostatic interactions in the K4E7 model.

(PDF)

Author Contributions**Supervision:** Boris B. Dzantiev.**Writing – original draft:** Andrey A. Buglak.**Writing – review & editing:** Anatoly V. Zherdev, Hong-Tao Lei.**References**

1. Li Y, Sun Y, Beier RC, Lei H, Gee S, Hammock BD, et al. Immunochemical techniques for multianalyte analysis of chemical residues in food and the environment: A review. *Trends Anal Chem.* 2017; 88:25–40.
2. Bottoni P, Grenni P, Lucentini L, Barra Caracciolo A. Terbutylazine and other triazines in Italian water resources. *Microchem J.* 2013; 107:136–142.
3. Cheney MA, Shin JY, Crowley DE, Alvey S, Malengreau N, Sposito G. Atrazine dealkylation on a manganese oxide surface. *Colloids Surf A.* 1998; 137:267–273.
4. Shin JY, Cheney MA. Abiotic dealkylation and hydrolysis of atrazine by birnessite. *Environ Toxicol Chem.* 2005; 24:1353–1360. PMID: [16117110](https://pubmed.ncbi.nlm.nih.gov/16117110/)
5. Tian Y, Shen W, Jia F, Ai Z, Zhang L. Sulfite promoted photochemical cleavage of s-triazine ring: The case study of atrazine. *Chem Eng J.* 2017; 330:1075–1081.
6. Rowe AM, Brundage KM, Barnett JB. Developmental Immunotoxicity of Atrazine in Rodents. *Basic Clin Pharmacol Toxicol.* 2008; 102:139–145. <https://doi.org/10.1111/j.1742-7843.2007.00175.x> PMID: [18226067](https://pubmed.ncbi.nlm.nih.gov/18226067/)
7. Agopian AJ, Langlois PH, Cai Y, Canfield MA, Lupo PJ. Maternal residential atrazine exposure and gastroschisis by maternal age. *Matern Child Health J.* 2013; 17:1768–1775. <https://doi.org/10.1007/s10995-012-1196-3> PMID: [23184502](https://pubmed.ncbi.nlm.nih.gov/23184502/)
8. Xu ZL, Shen YD, Zheng WX, Beier RC, Xie GM, Dong JX, et al. Broad-Specificity Immunoassay for O, O-Diethyl Organophosphorus Pesticides: Application of Molecular Modeling to Improve Assay Sensitivity and Study Antibody Recognition. *Anal Chem.* 2010; 82:9314–9321. <https://doi.org/10.1021/ac1018414> PMID: [20958019](https://pubmed.ncbi.nlm.nih.gov/20958019/)
9. Yuan M, Na Y, Li L, Liu B, Sheng W, Lu X, et al. Computer-Aided Molecular Modeling Study on Antibody Recognition of Small Molecules: An Immunoassay for Triazine Herbicides. *J Agric Food Chem.* 2012; 60:10486–10493. <https://doi.org/10.1021/jf303256r> PMID: [23043348](https://pubmed.ncbi.nlm.nih.gov/23043348/)
10. Liu C, Dou X, Zhang L, Kong W, Wu L, Duan Y, et al. Development of a broad-specificity antibody-based immunoassay for triazines in ginger and the quantitative structure-activity relationship study of cross-reactive molecules by molecular modeling. *Anal Chim Acta.* 2018; 1012:90–99. <https://doi.org/10.1016/j.aca.2018.01.042> PMID: [29475478](https://pubmed.ncbi.nlm.nih.gov/29475478/)
11. Bitencourt M, Freitas MP. MIA-QSAR evaluation of a series of sulfonylurea herbicides. *Pest Manag Sci.* 2008; 64:800–807. <https://doi.org/10.1002/ps.1565> PMID: [18338340](https://pubmed.ncbi.nlm.nih.gov/18338340/)
12. Zuo Y, Wu Q, Su SW, Niu CW, Xi Z, Yang GF. Synthesis, Herbicidal Activity, and QSAR of Novel N-Benzothiazolyl-pyrimidine-2,4-diones as Protoporphyrinogen Oxidase Inhibitors. *J Agric Food Chem.* 2016; 64:552–562. <https://doi.org/10.1021/acs.jafc.5b05378> PMID: [26728549](https://pubmed.ncbi.nlm.nih.gov/26728549/)
13. Funar-Timofei S, Borota A, Crisan L. Combined molecular docking and QSAR study of fused heterocyclic herbicide inhibitors of D1 protein in photosystem II of plants. *Mol Divers.* 2017; 21:437–454. <https://doi.org/10.1007/s11030-017-9735-x> PMID: [28299584](https://pubmed.ncbi.nlm.nih.gov/28299584/)
14. Xie Y, Peng W, Ding F, Liu SJ, Ma HJ, Liu CL. Quantitative structure-activity relationship (QSAR) directed the discovery of 3-(pyridin-2-yl)benzenesulfonamide derivatives as novel herbicidal agents. *Pest Manag Sci.* 2018; 74:189–199. <https://doi.org/10.1002/ps.4693> PMID: [28762622](https://pubmed.ncbi.nlm.nih.gov/28762622/)
15. Wang Z, Zhu Y, Ding S, He F, Beier RC, Li J, et al. Development of a Monoclonal Antibody-Based Broad-Specificity ELISA for Fluoroquinolone Antibiotics in Foods and Molecular Modeling Studies of Cross-Reactive Compounds. *Anal Chem.* 2007; 79:4471–4483. <https://doi.org/10.1021/ac070064t> PMID: [17511422](https://pubmed.ncbi.nlm.nih.gov/17511422/)

16. Mu H, Wang B, Xu Z, Sun Y, Huang X, Shen Y, et al. Stereospecific recognition and quantitative structure–activity relationship between antibodies and enantiomers: ofloxacin as a model hapten. *Analyst*. 2015; 140:1037–1045. <https://doi.org/10.1039/c4an02155j> PMID: 25553733
17. Chen J, Lu N, Shen X, Tang Q, Zhang C, Xu J, et al. Investigation of an Immunoassay with Broad Specificity to Quinolone Drugs by Genetic Algorithm with Linear Assignment of Hypermolecular Alignment of Data Sets and Advanced Quantitative Structure–Activity Relationship Analysis. *J Agric Food Chem*. 2016; 64:2772–2779. <https://doi.org/10.1021/acs.jafc.6b00039> PMID: 26982746
18. Chen J, Wang L, Lu L, Shen X, Huang XA, Liu Y, et al. Four Specific Hapten Conformations Dominating Antibody Specificity: Quantitative Structure–Activity Relationship Analysis for Quinolone Immunoassay. *Anal Chem*. 2017; 89:6740–6748. <https://doi.org/10.1021/acs.analchem.7b00997> PMID: 28513147
19. Mu H, Xu Z, Liu Y, Sun Y, Wang B, Sun X, et al. Probing the stereoselective interaction of ofloxacin enantiomers with corresponding monoclonal antibodies by multiple spectrometry. *Spectrochim Acta A Mol Biomol Spectrosc*. 2018; 194:83–91. <https://doi.org/10.1016/j.saa.2018.01.008> PMID: 29328954
20. Xu ZL, Xie GM, Li YX, Wang BF, Beier RC, Lei HT, Wang et al. Production and characterization of a broad-specificity polyclonal antibody for O,O-diethyl organophosphorus pesticides and a quantitative structure–activity relationship study of antibody recognition. *Anal Chim Acta*. 2009; 647:90–96. <https://doi.org/10.1016/j.aca.2009.05.025> PMID: 19576391
21. Xu ZL, Shen YD, Zheng WX, Beier RC, Xie GM, Dong JX, et al. Broad-Specificity Immunoassay for O, O-Diethyl Organophosphorus Pesticides: Application of Molecular Modeling to Improve Assay Sensitivity and Study Antibody Recognition. *Anal Chem*. 2010; 82:9314–9321. <https://doi.org/10.1021/ac1018414> PMID: 20958019
22. Yuan M, Liu B, Liu E, Sheng W, Zhang Y, Crossan A, et al. Immunoassay for Phenylurea Herbicides: Application of Molecular Modeling and Quantitative Structure–Activity Relationship Analysis on an Antigen–Antibody Interaction Study. *Anal Chem*. 2011; 83:4767–4774. <https://doi.org/10.1021/ac200227v> PMID: 21539295
23. Wang Z, Kai Z, Beier RC, Shen J, Yang X. Investigation of Antigen–Antibody Interactions of Sulfonamides with a Monoclonal Antibody in a Fluorescence Polarization Immunoassay Using 3D-QSAR Models. *Int J Mol Sci*. 2012; 13:6334–6351. <https://doi.org/10.3390/ijms13056334> PMID: 22754368
24. Dankwardt A, Hock B, Simon R, Freitag D, Kettrup A. Determination of Non-extractable Triazine Residues by Enzyme Immunoassay: Investigation of Model Compounds and Soil Fulvic and Humic Acids. *Environ Sci Technol*. 1996; 30:3493–3500.
25. Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comp Chem*. 1996; 17: 490–519.
26. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J Am Chem Soc*. 1985; 107:3902–3909.
27. Alexander DL, Tropsha A, Winkler DA. Beware of R(2): Simple, Unambiguous Assessment of the Prediction Accuracy of QSAR and QSPR Models. *J Chem Inf Model*. 2015; 55:1316–1322. <https://doi.org/10.1021/acs.jcim.5b00206> PMID: 26099013
28. Golbraikh A, Tropsha A. Beware of q2! *J Mol Graph Model*. 2002; 20:269–276. PMID: 11858635
29. Tropsha A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol Inform*. 2010; 29:476–488. <https://doi.org/10.1002/minf.201000061> PMID: 27463326
30. Sliwoski G, Kothiwale S, Meiler J, Lowe EW Jr. Computational methods in drug discovery. *Pharmacol Rev*. 2013; 66:334–395. <https://doi.org/10.1124/pr.112.007336> PMID: 24381236
31. Tosco P, Balle T, Shiri FJ. Open3DALIGN: an open-source software aimed at unsupervised ligand alignment. *Comput Aided Mol Des*. 2011; 25:777–783.
32. Tosco P, Balle T. An open-source software aimed at highthroughput chemometric analysis of molecular interaction fields. *J Mol Model*. 2011; 17:201–208. <https://doi.org/10.1007/s00894-010-0684-x> PMID: 20383726
33. Stahle L, Wold S. Partial least squares analysis with cross-validation for the two-class problem: A Monte Carlo study. *J Chemom*. 1987; 1:185–196.
34. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemometrics Intell Lab Syst*. 2001; 58:109–130.
35. Pastor M, Cruciani G, Clementi S. Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure–activity relationships. *J Med Chem*. 1997; 40:1455–1464. <https://doi.org/10.1021/jm9608016> PMID: 9154968
36. Baroni M, Clementi S, Cruciani G, Costantino G, Riganelli DJ. Predictive ability of regression models. Part II: Selection of the best predictive PLS model. *Chemometr*. 1992; 6:347–356.

37. Baroni M, Costantino G, Cruciani G, Riganelli D, Valigi R, Clementi S. Generating optimal linear PLS estimations (GOLPE): An advanced chemometric tool for handling 3D-QSAR problems. *Quant Struct-Act Relat.* 1993; 12:9–20.
38. Hall LH, Mohny B, Kier LB. The electrotopological state: structure information at the atomic level for molecular graphs. *J Chem Inf Comput Sci.* 1991; 31:76–82.