

Structures and biological activity of cinnamoyl derivatives of coumarins and dehydroacetic acid and their boron difluoride complexes*

K. V. Tambov,^a I. V. Voevodina,^a A. V. Manaev,^a Ya. A. Ivanenkov,^b N. Neamati,^c and V. F. Traven^{a*}

^aD. I. Mendeleyev University of Chemical Technology of Russia,

9 Miusskaya pl., 125047 Moscow, Russian Federation.

Fax: +7 (499) 978 8660. E-mail: traven@muctr.ru

^bChemical Diversity Research Institute,

Build. 1, 2-a ul. Rabochaya, 141400 Khimki, Moscow Region, Russian Federation.

Fax: +7 (495) 626 9780. E-mail: chemrar@chemrar.ru

^cCollege of Pharmacy, University of Southern California,

USA, 1985 Zonal Avenue, Los Angeles, CA 90089-9121.

Fax: (323) 442 1369. E-mail: neamati@usc.edu

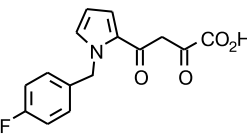
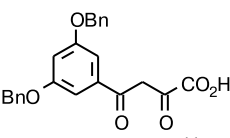
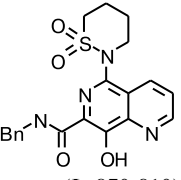
Computer-assisted screening for Kohonen self-organizing maps in terms of the quantitative structure–activity relationship (QSAR) model revealed the high potential activity of cinnamoyl derivatives of coumarin and dehydroacetic acid and their boron difluoride complexes against a number of biological targets, including HIV-1 integrase. The pronounced inhibitory properties of dehydroacetic acid derivatives and their boron difluoride complexes against HIV-1 integrase were experimentally confirmed by *in vitro* testing of their antiviral activity with respect to HIV-infected cells. The data obtained suggests a correlation between the structure of the compounds studied and their biological activity.

Key words: coumarin derivatives, cytotoxicity, biological activity, boron difluoride complexes, dehydroacetic acid derivatives.

Many 1,3-dicarbonyl, cinnamoyl, and polymethine derivatives of coumarin and its analogs are known for pronounced biological activity. For instance, some pyrone derivatives exhibit anticancer, antimicrobial, and anti-coagulant properties. Boron-containing derivatives have been used in boron neutron capture therapy of tumors and in inhibition of HIV protease.¹ Polymethine hetarene derivatives can act as fluorescent labels of various proteins.^{2,3} It should be emphasized that many dicarbonyl hetarene derivatives produce an inhibitory effect on HIV-1 integrase;^{4–10} some examples are given in Table 1.

In the study of condensation reactions of boron difluoride complexes of acyl(hydroxy)hetarenes, we have obtained a great number of chalcone analogs: cinnamoyl derivatives of coumarin, 2-quinolone, and dehydroacetic acid.^{13–16} Because novel compounds can have high biological activity, we studied them in a systematic manner. In the present work, we examined a correlation between the structures and biological activity of cinnamoyl derivatives coumarin, dehydroacetic acid, and their boron difluoride complexes by computer-assisted screening and

Table 1. Some HIV-1 integrase inhibitors

Inhibitor	Inhibitory activity, IC ₅₀ /μmol L ⁻¹	
	3'-Processing	Chain transfer
 (L-731,988) ¹¹	6	0.050
 (L-708,906) ¹¹	1	0.100
 (L-870,810) ¹²	0.085	0.008

* Dedicated to Academician of the Russian Academy of Sciences O. M. Nefedov on the occasion of his 80th birthday.

tested these compounds *in vitro* for antiviral activity in HIV-infected cells.

Results and Discussion

Prediction of the profile of the target-specific activity of coumarin derivatives and its analogs using the algorithm of Kohonen self-organizing maps. First, we tried to predict the profile of the target-specific activity of the compounds obtained using the algorithm of Kohonen self-organizing maps.^{17–20} For this purpose, we designed a computer-assisted classifying model for estimation of the biological activity of organic molecules. In terms of this model describing a quantitative structure–activity correlation (QSAR), we predicted the profile of the target-specific activity for all the derivatives of coumarin and its analogs. According to the prediction results, we grouped the compounds obtained together in separate focused libraries with the most probable mechanism of action. This approach allows one to predict, in the context of a single model, the binding efficiency for a number of therapeutically significant biological targets rather than for a specific target only. The theoretical and practical aspects of the approach have been discussed earlier.^{21–25}

Training sample. Drugs from the Integrity Prous database (Prous Science, URL: <http://www.prous.com/integrity>) were used as a training sample. The sample consisted of more than 17 000 compounds whose target-specific activity against various receptors and enzymes (*e.g.*, tyrosine kinases, G protein-coupled receptors, nuclear receptors, HIV integrase, DNA topoisomerase, caspase enzymes, chemokine receptors, *etc.*) were experimentally determined and confirmed. The total number of unique types of biological targets in the training sample was over 200. It was demonstrated that the key statistical parameters of selected model structures such as the number of unique heterocycles, the diversity coefficients, and the number of different substructural fragments is indicative of their wide structural diversity.²⁶ Therefore, the results of the computer-assisted modeling are not distorted by the predominance of some chemotypes but they reflect general objective trends.

Molecular descriptors. The choice and calculation of the most appropriate molecular descriptors for modeling is an essential step in the design of an effective computer-assisted model. For all the compounds from the training sample, we calculated a wide set of molecular descriptors with the SmartMining program package.^{27–30} The most significant features were selected with the *t*-test and principal component analysis (PCA).^{31,32} Using the *t*-test, we compared in pairs sets of compounds from different target-specific groups and revealed a number of descriptors making the greatest contributions to the discrimination between most of the sets. The variability of the input space of descriptors was analyzed by PCA. We selected the de-

scriptors that contribute most substantially to the first five components describing more than 90% of the system variability. The number of principal components was estimated according to the *Broken Stick Model* rule. As a result, the final set included eight descriptors that describe the input chemical space most accurately (given below).

Descriptor	Definition
$\log P$	Calculated logarithm of the partition coefficient of an organic compound in the octan-1-ol/water system
$S(-CH_2-)$	Electrotopological index of the structural fragment $-CH_2-$
$S(-N=)$	Electrotopological index of the structural fragment $-N=$
$S(-OH)$	Electrotopological index of the structural fragment $-OH$
$S(=O)$	Electrotopological index of the structural fragment $=O$
HB2	Calculated strength value of potential hydrogen bond acceptors
Zagreb 1 st	Gutman topological index
HBD	Number of potential hydrogen bond donors

Kohonen self-organizing maps. The next step of our computer-assisted screening was building of a general Kohonen map. Kohonen self-organizing maps are a modern computer-assisted tool for highly efficient data analysis and visualization as 2D or 3D images of multiparametric feature spaces.¹⁷ This method finds increasingly wide use in diverse areas of organic and medicinal chemistry, including prediction of various physiological and physico-chemical properties of organic compounds.

We used the above eight molecular descriptors to build two Kohonen maps (Fig. 1) with the SmartMining software. The positions of the Kohonen neurons correspond to the network nodes; in other words, the point in the map with $x = 1$ and $y = 1$ corresponds to the "first" node of the Kohonen lattice, the point with $x = 2$ and $y = 1$, to the "second" node, *etc.*

To train the algorithm, we employed the following parameters and settings: (1) the number of nodes (neurons) of 2D Kohonen maps is 196 (the nodes are ordered in a standard rectangular topology); (2) the number of epochs is 2000; (3) the WTA (winner takes all) strategy for selection of a winning neuron; (4) the initial learning radius is 8; (5) the initial learning rate is 0.3; (6) the standard Gaussian activation function; (7) the Euclidean metrics; (8) random initialization of the weighting factors (the values of the descriptors are normalized).

Such modifications as "two training cycles", "neural gas", "convex combination", "Duane DeSieno method", and "noise technique" were not used.

The constructed maps are 2D images of the eight-dimensional space of molecular descriptors; note that near neighbors in the input feature space are also such on Kohonen maps. Such a map dimensionality was chosen

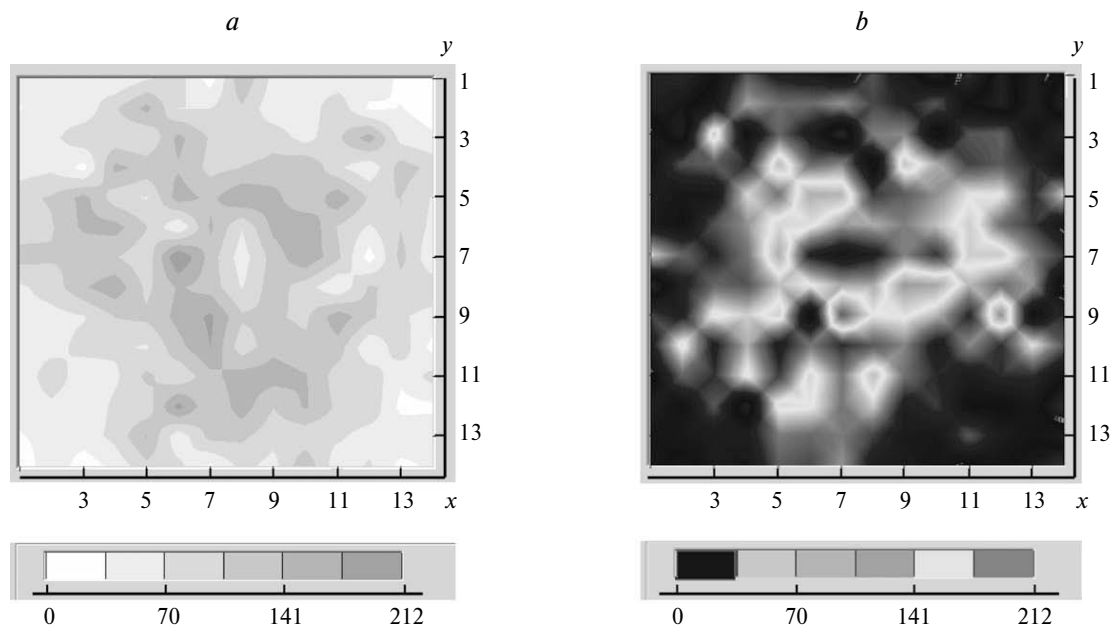


Fig. 1. Kohonen maps (14×14) for the whole input training sample of compounds for the first (a) and second randomizations (b).

for optimum resolution of the image, as well as for avoiding unnecessary fragmentation giving rise to sparsely populated areas with undefined classification.

Then, in either Kohonen map we revealed areas of domination for compounds from different target-specific groups (data for first randomization is shown in Fig. 2*). It can be seen in Fig. 2 that the compounds from the training sample are broadly scattered in the maps, forming irregularly shaped areas of different sizes. The map shows clear-cut areas in which the groups of the compounds under study dominate or, in contrast, are scarce. The populations of such areas can differ by more than two or three orders of magnitude.

Each point in the maps was classified as a cell belonging to a certain group of target-specific activity if the corresponding molecules dominate in it (in percentage). Such a division of the map makes its analysis and testing easier, though this may lead to incorrect assignments of points in which the molecules of one category only slightly dominate over others. A cell was assigned to a certain class (*e.g.*, tyrosine kinase inhibitors) if the percentage of the compounds of this group in the cell was higher the percentage of the compounds of other groups by a factor >1.2. When this assignment was completed for all the areas, it turned out that the constructed Kohonen maps have high potential predicting ability. The mean degree of discrimination between the model groups of compounds was 77 (first randomization) and 75% (second randomization). With consideration of the specificity of our task

(the design of an accurate model for classification of compounds with respect to a number of biological targets), these values are quite acceptable, which in turn suggests a high discriminating ability of the designed model. It is significant that the incorrectly classified compounds are predominantly localized in borderline regions between different areas of target-specific activity. No wonder that the binary method we used for categorization of compounds and areas in the maps gives some erroneously predicted compounds. Obviously, addition of intermediate categories (*e.g.*, for compounds showing a non-selective profile of target-specific activity) can increase the fraction of correctly classified compounds in key groups.

The designed model (the best randomization) was tested with internal validation. Following this method, we randomly excluded from the general training sample 10% of the compounds in each target-specific group. The selected structures were combined into a tested sample. This procedure was repeated three times by randomly excluding 10% of the compounds. We found that the prediction accuracy is comparable with the classification accuracy with training examples (on average, 73%). The model was not subjected to independent testing with biological screening data. However, such tests have been carried out earlier for other similar models also designed by us and the results obtained have been published.

The most important result of these experiments is that the eight-dimensional feature space under examination (defined with eight selected molecular descriptors) contains compact areas corresponding to groups of compounds with different target-specific activity profiles. As a result,

* In statistics theory, randomization is taken to mean an outcome of random data mixing within a separate sample.³³

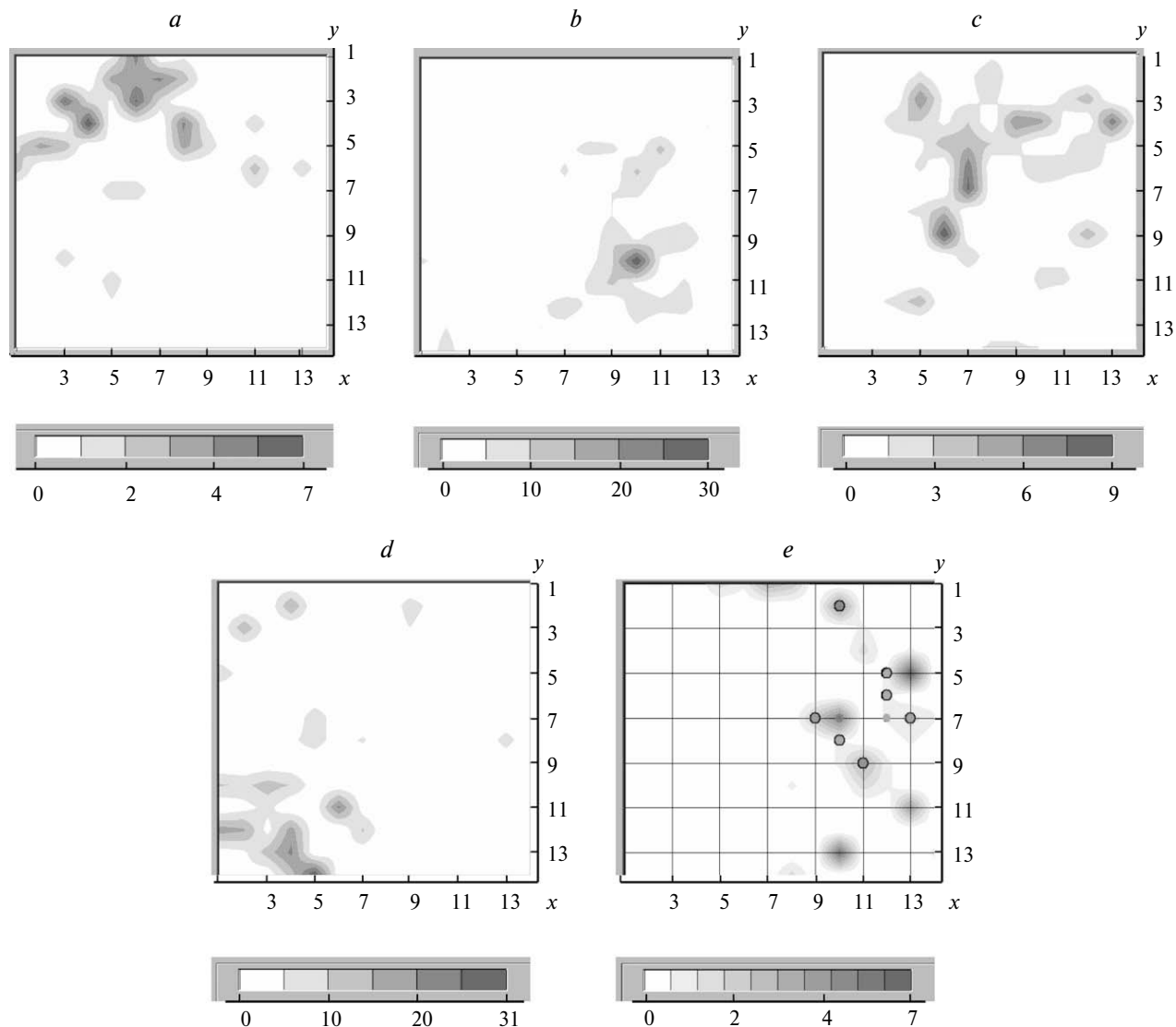


Fig. 2. Distribution of the representative groups of compounds with different target-specific activities in the Kohonen maps (first randomization): P38MAPK inhibitors (*a*), β - (*b*) and α -adrenoceptor antagonists (*c*), thrombin inhibitors (*d*), and thromboxane synthase inhibitors (*e*).

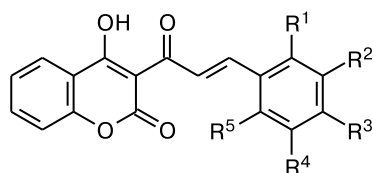
the constructed Kohonen maps, which are 2D images of the above eight-dimensional space, can distinguish between the locations of these groups.

Target-specific activity of derivatives of coumarin and its analogs. The revealed trends in the location of various groups of drugs on Kohonen maps were used to predict the profile of the target-specific activity of the compounds from the test set of derivatives of coumarin and its analogs. In particular, the tested sample included compounds **1–6** distributed among groups I (**1** and **3**) and II (**2** and **4–6**).

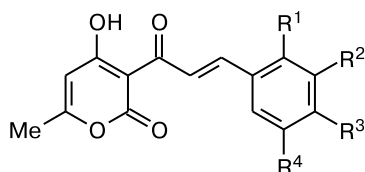
The profiling methodology consisted in calculation of eight descriptors (analogous to those used in the design of the model) for each derivative of coumarin and its analogs (the tested set) followed by location of each structure in

the aforesaid Kohonen maps in the testing mode. The tested compound was classified as belonging to a unique class(es) of physiologically active drugs if it was located in the area of domination of molecules with a particular target-specific activity from the training sample. The corresponding representative Kohonen maps reflecting the distribution of the compounds from the tested sample by target-specific activity are shown in Figs 3 and 4.

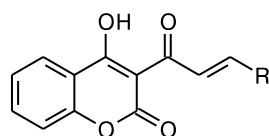
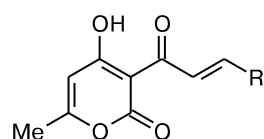
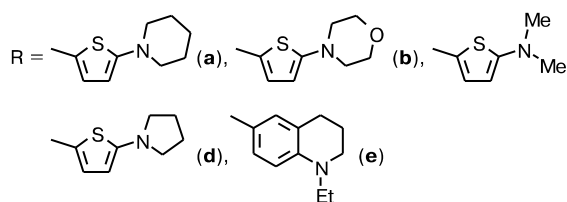
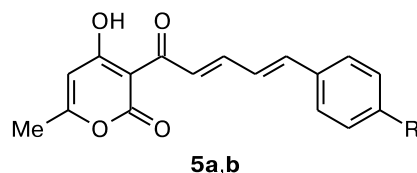
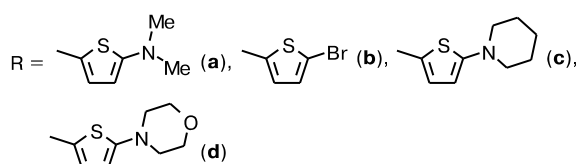
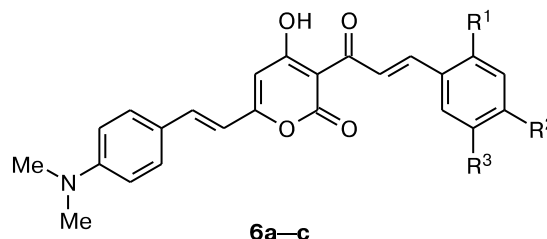
It is evident from Figs 3 and 4 that the derivatives of coumarin and its nearest analogs can be highly active against various biological targets. Moreover, our model of computer-assisted screening makes it possible to estimate a relationship between the biological activity of these derivatives and their isomeric and tautomeric transforma-

**1a–e**

	R ¹	R ²	R ³	R ⁴	R ⁵
a	H	H	OH	OEt	H
b	OH	OMe	H	H	H
c	OH	H	OMe	H	H
d	H	OMe	OH	OMe	H
e	H	H	OMe	OH	H
f	OH	OEt	H	H	H
g	OMe	OH	H	H	H
h	OH	H	NEt ₂	H	H
i	H	H	C ₈ H ₁₆ OH	H	H
j	H	H	OMe	H	H
k	OMe	H	OMe	H	H
l	H	H	NMe ₂	H	H

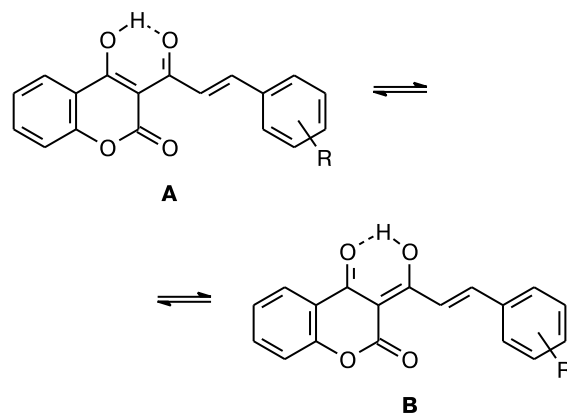
**2a–f**

	R ¹	R ²	R ³	R ⁴
a	H	H	OMe	H
b	H	H	NMe ₂	H
c	H	OMe	OMe	OMe
d	H	H	F	H
e	OMe	H	H	Br
f	H	H	Br	H

**3a–e****4a–d****5a,b**R = OMe (**a**), NMe₂ (**b**)**6a–c**

R¹ = H (**a**, **b**), OMe (**c**)
 R² = NMe₂ (**a**), OMe (**b**), H (**c**)
 R³ = H (**a**, **b**), Br (**c**)

tions. In particular, tautomerizable structures from group I were tested in two tautomeric forms: *endo*-enol (**A**) and *exo*-enol (**B**) (Scheme 1).

Scheme 1

It can be seen in Fig. 3 that the areas of the dioxo and hydroxy tautomers from tested sample I do not coincide in most cases. The distributions of the compounds from group I in the Kohonen maps show a moderate degree of discrimination between *endo*- and *exo*-enols. Within a single group, we demonstrated that *endo*- and *exo*-enols can exhibit fundamentally different target-specific activities. It should be emphasized that no difference in target-specific activity was observed for the tautomers of the compounds from group II.

All the observations and conclusions are based only on the trends emerging from the constructed Kohonen maps, with eight selected molecular descriptors used. Two randomizations showed that this model has the property of reproducing prediction results. A more detailed cross analy-

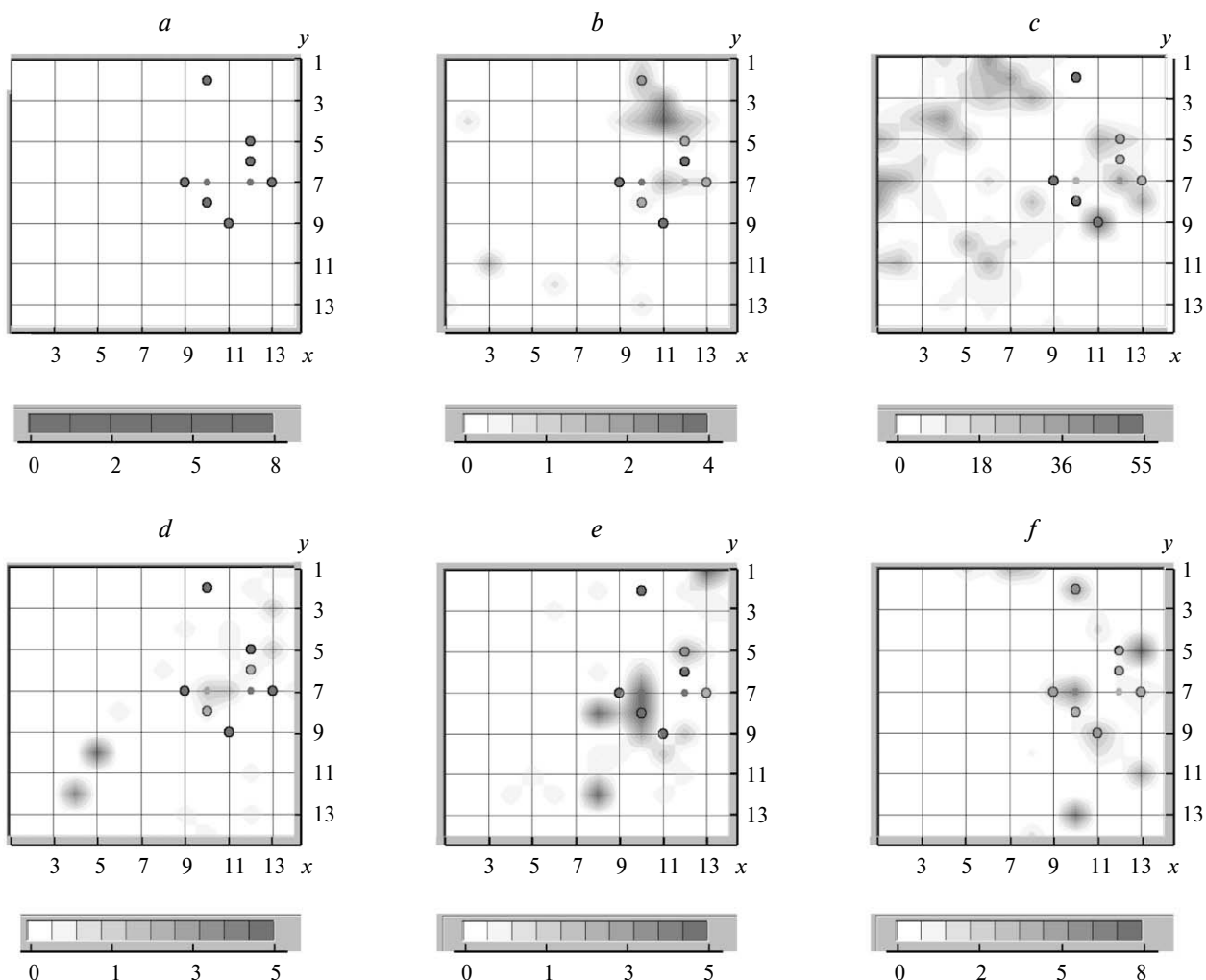


Fig. 3. Localization of compounds from the tested sample I (here and in Fig. 4, black and gray dots refer to the *endo*- and *exo*-enol forms, respectively) against the background of representative target-specific areas: compounds from group II (a), lactamase inhibitors (b), squalene synthase inhibitors (c), HIV integrase inhibitors (d), DNA gyrase inhibitors (e), aldose reductase inhibitors (f), and potassium channel inhibitors (g).

sis of the positions of the tested compounds in both Kohonen maps revealed that the same compounds from the tested sample are classified in the majority of cases (95%) with an equal probability as the same class of target-specific activity. Insignificant fluctuations (5%) in the prediction can be explained by the nonuniformity of the initial random distribution of the model structures before the training procedure.

As a conclusion, one can state that the designed model provides high-quality classification of compounds from an independent experimental sample. It is worth noting that the training and tested samples contain different chemical types of compounds. Therefore, the designed model reveals some fundamental trends that are not associated with the direct structural similarity of molecules from one of the classes of target-specific activity under consideration.

The literature data suggest that the coumarin derivatives and their analogs studied can have some effect on HIV-1 integrase.^{13–16} Our analysis of the target-specific activity of these compounds through the use of the algorithm of Kohonen self-organizing maps confirmed that both cinnamoylcoumarins (compounds of group I) and cinnamoylpyrones (compounds of group II) can act as HIV-1 integrase inhibitors. At the same time, the compounds of group II are located much closer to the area of domination of the molecules with specific activity against HIV-1 integrase from the training sample (see Figs 3 and 4).

Experimental study of 3-acyl-4-hydroxypyrones and their boron difluoride complexes as HIV-1 integrase inhibitors. The promising use of cinnamoylpyrones as HIV-1 integrase inhibitors was convincingly confirmed by experimental *in vitro* tests of the compounds from groups I and

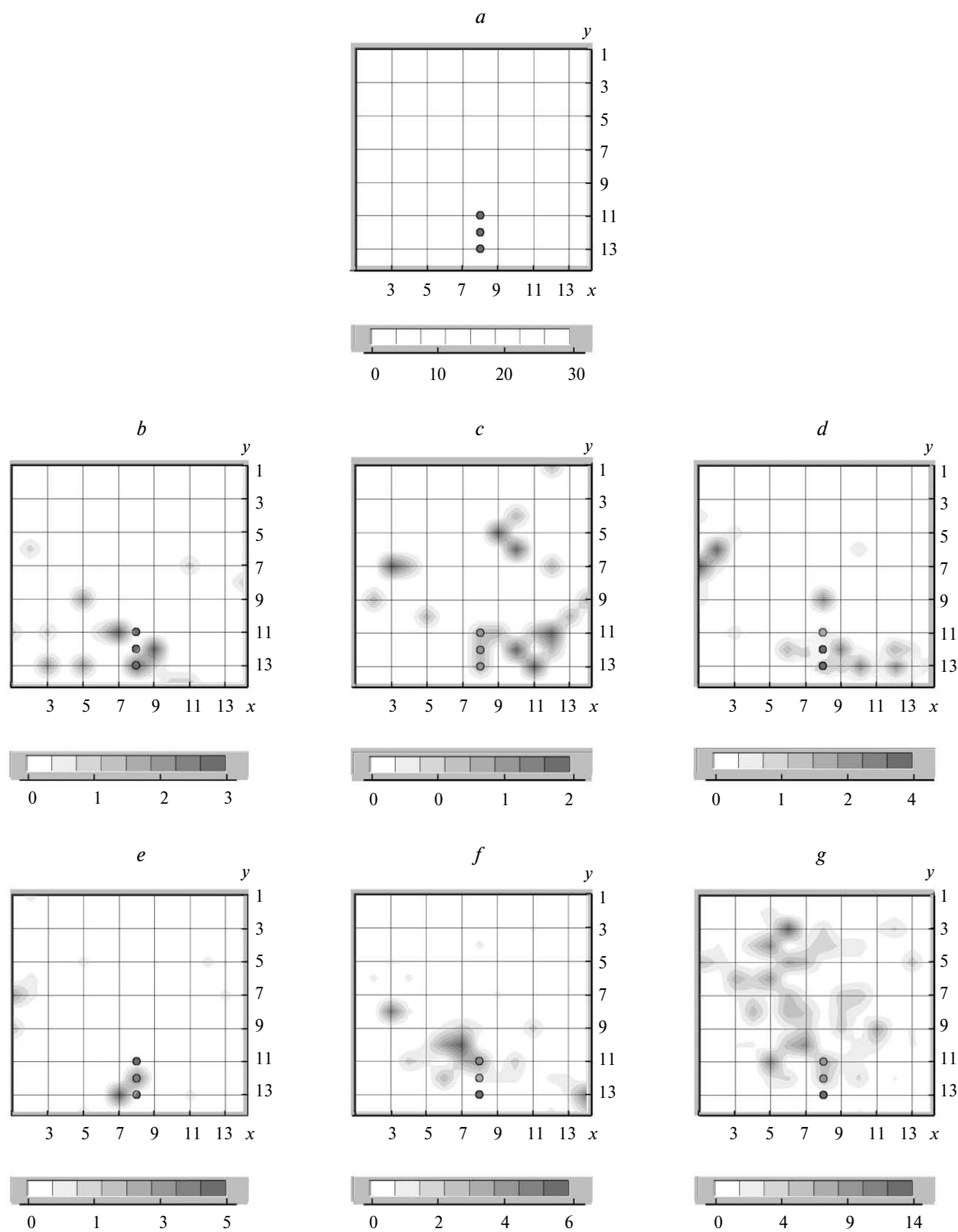


Fig. 4. Localization of compounds from the tested sample II (a) against the background of representative target-specific areas: lactamase inhibitors (b), squalene synthase inhibitors (c), HIV integrase inhibitors (d), DNA gyrase inhibitors (e), aldose reductase inhibitors (f), potassium channel inhibitors (g).

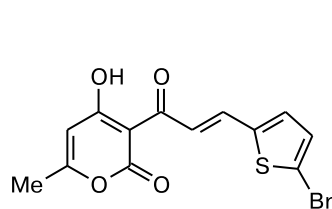
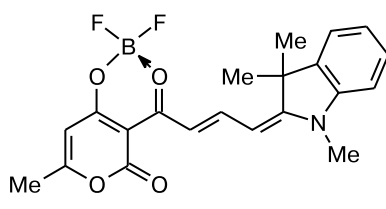
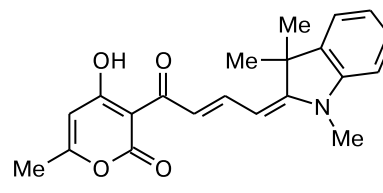
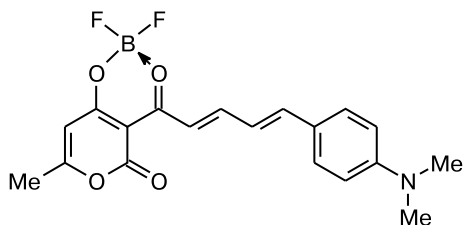
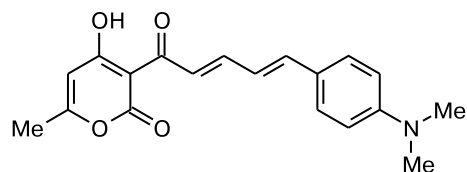
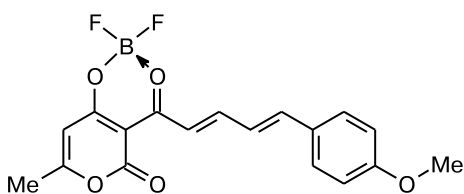
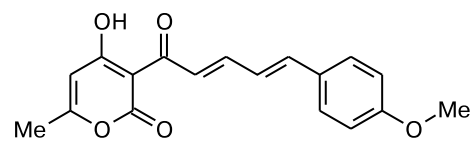
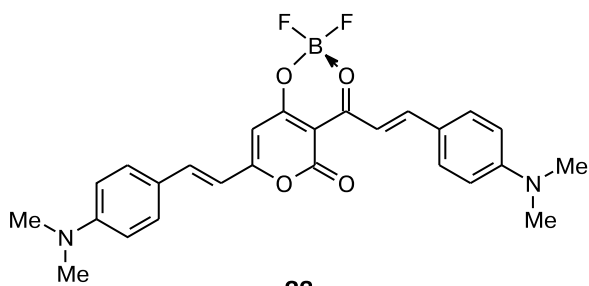
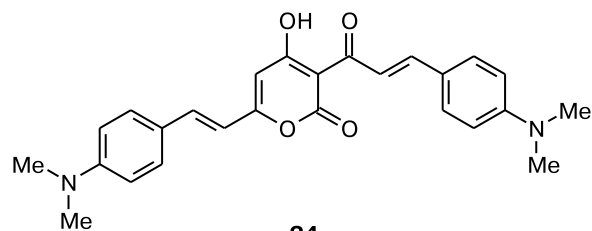
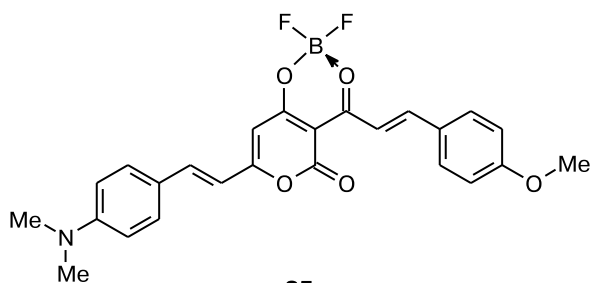
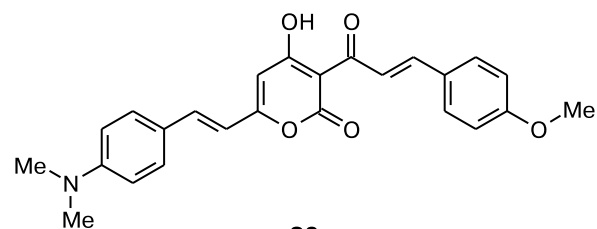
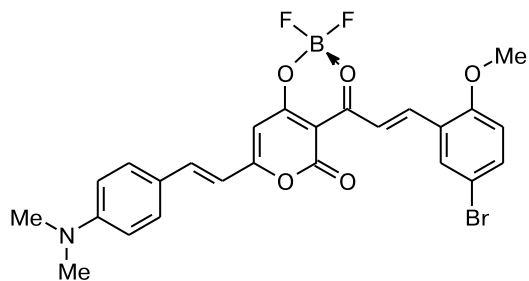
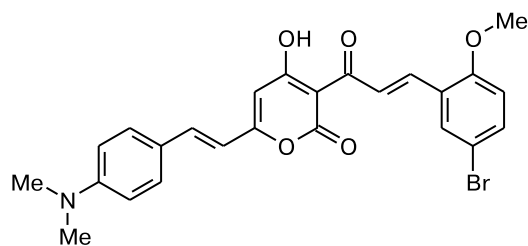
**16****17****18****19****20****21****22****23****24****25****26****27****28**

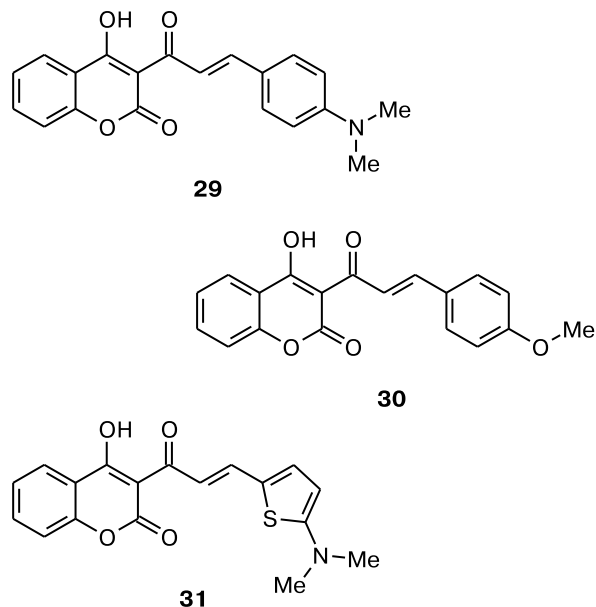
Table 2. Inhibition of the cytotoxicity^a and catalytic activity of HIV-1 IN in 3'-processing and chain transfer

Compound	IC ₅₀ ^b /μmol L ⁻¹	
	3'-Processing	Chain transfer
7	13±8	7±2
8	>100	56±1
9	12±6	4±2
10	>100	>100
11	9±3	16±4
12	64	>100
13	>100	>100
14	19±4	11±6
15	9±2	3±2
16	65±4	23±3
17	>100	>100
18	>100	>100
19	>100	100
20	>100	>100
21	11±3	4±3
22	>100	39±14
23	>100	>100
24	>100	>100
25	>100	>100
26	>100	>100
27	>100	>100
28	>100	>100

^a Determined by the MTT assay; in all the cases, the cytotoxicity exceeds 10 μmol L⁻¹.

^b From at least three independent runs.

Based on the experimental data obtained, we can make some conclusions about a relationship between the structure and inhibitory properties of the cinnamoyl derivatives of dehydroacetic acid. The poorest inhibitory activity was exhibited by the compounds



with bulky substituents. The boron difluoride complexes with electron-withdrawing substituents in the phenyl and hetaryl fragments in position 3 proved to be most active. These complexes have the lowest IC₅₀ values, being most active in the inhibition of 3'-processing and chain transfer reactions. The presence of a chelate boron complex is decisive for the activity of the inhibitors.

To sum up, computer-assisted screening in combination with experimental *in vitro* testing for antiviral activity in HIV-infected cells revealed a novel class of IN inhibitors, namely, dehydroacetic acid derivatives. Further structural optimization of these compounds seems to be promising in a search for the most efficient inhibitors of HIV-1 integrase.

Table 3. Parameters of the binding of some compounds to the active site of integrase (molecular docking data)

Compound	MW ^a	RB ^b	HBA ^c	HBD ^d	AlogP ^e	S+logP ^f	PSA ^g	eHits_Score	GOLD_Score
7	441.89	2	8	0	4.358	2.53	91.59	-1.08	51.99
9	382.95	2	4	0	4.41	4.11	46.83	-2.59	49.89
10	335.15	3	4	1	2.87	3.88	63.60	-1.96	42.13
11	322.04	2	4	0	3.86	3.60	46.83	-1.68	46.55
12	274.24	3	4	1	2.33	3.35	63.60	-1.74	44.45
13	353.15	3	5	0	3.90	3.42	78.30	-2.15	42.36
14	305.35	4	5	1	2.36	2.92	95.08	-2.01	47.72
15	388.98	2	4	0	4.18	3.84	75.06	-1.38	50.76
16	341.18	3	4	1	2.64	3.53	91.84	-2.12	48.39
21	360.12	4	5	0	4.11	4.33	56.06	-1.75	54.98
22	312.32	5	5	1	2.57	3.65	72.83	-1.79	50.12

^a Molecular weight.

^b The number of rotating bonds.

^c The number of hydrogen bond acceptors.

^d The number of hydrogen bond donors.

^e Atom-based logP calculations.

^f S+logP (Simulations plus logP model).

^g Polar surface area.

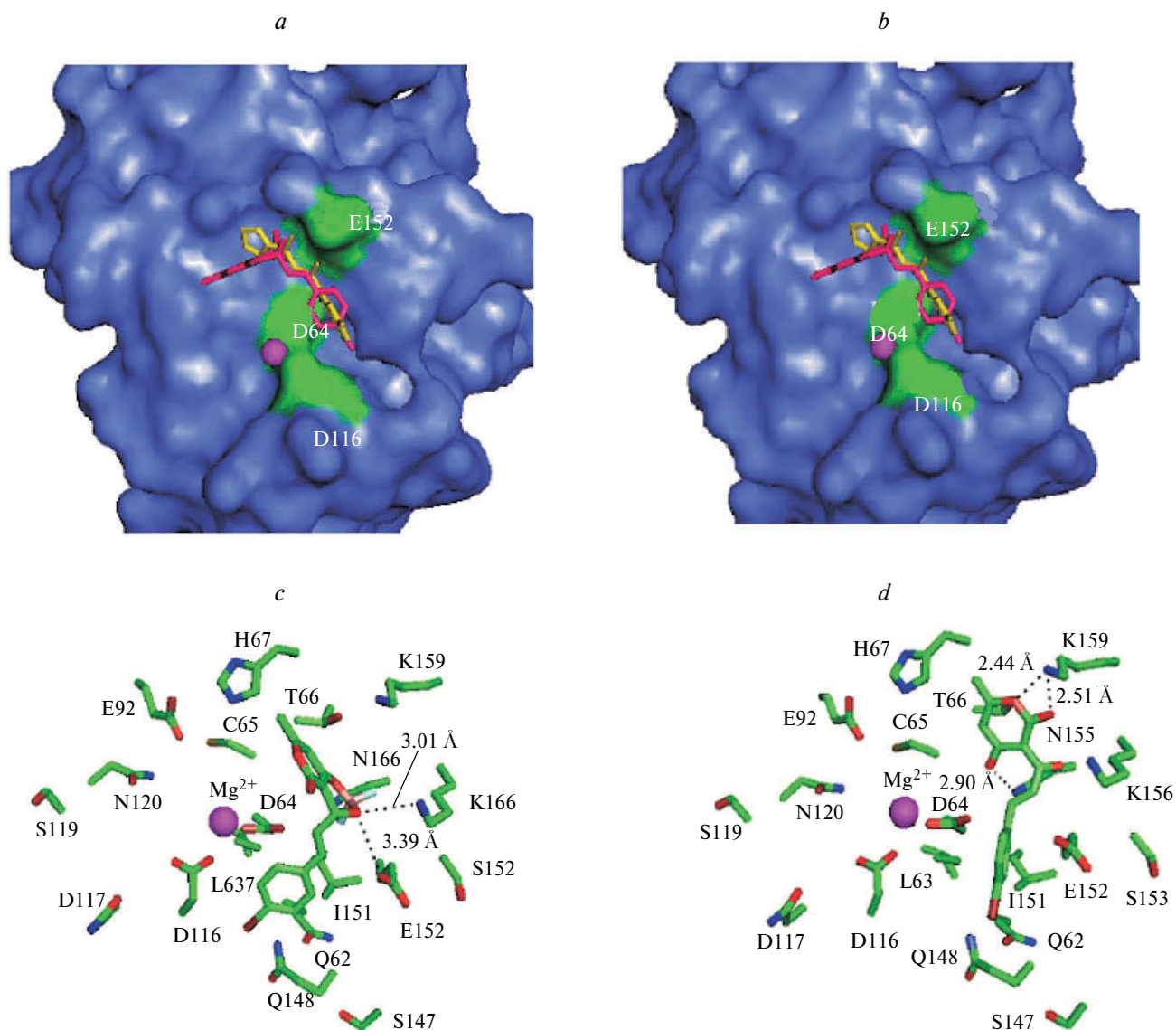


Fig. 5. Binding conformation of ligands **9** and **10** in the catalytic domain of HIV-1 IN: compounds **3** (a) and **4** (b) (both pink) over the crystalline ligand 5CITEP (yellow) on the binding surface (green) and the Mg²⁺ ion (purple); detailed representation of the interactions between integrase and ligands **9** (c) and **10** (d).

Note. Figure 5 is available in full color in the on-line version of the journal (<http://www.springerlink.com>).

Experimental

Computer-assisted screening was performed with the Smart-Mining program designed at the Chemical Diversity Research Institute (www.chemrar.ru). Other commercial programs for the design of computer models include NeuroSolution (NeuroDimension, Inc., USA; www.nd.com), SVMlight (Joachims T.; www.svmlight.joachims.org), and ChemoSoft (Chemical Diversity Labs, Inc. (USA)—CDRI (Russia); www.chemdiv.com).

Biochemical tests for antiviral activity in HIV-infected cells were carried out at the Department of Pharmaceutical Sciences of the University of Southern California (Dr. Nouri Neamati, USA).³³

Compounds and enzymes. All the compounds were dissolved in DMSO and the resulting solutions were kept at −20 °C.

Samples of γ-[³²P]-ATP were purchased from Amersham Biosciences or ICN. Wild-type integrase was provided by Dr. Robert Craigie (Laboratory of Molecular Biology, NIDDK, NIH (Bethesda, MD)).

Preparation of oligonucleotide substrates. The 21-unit oligonucleotide (21top (5′-GTGTGGAAAATCTCTAGCAGT-3′) and 21bot (5′-ACTGCTAGAGATTTTCCA CAC-3′)) was purchased from Norris Cancer Center Microsequencing Core Facility (University of Southern California) and purified by exposing to UV light on polyacrylamide gel. To estimate the inhibitory activity of 3′-processing and chain transfer reactions with 5-terminal labeled substrates, the 5-terminal groups of the oligonucleotide 21top were labeled with T4-polynucleotide kinase (Epicentre, Madison, WI) and [γ-³²P]-ATP. The kinase was thermally inactivated, whereupon a 1.5 *M* excess of the oligonucle-

otide 21bot was added. The mixture was heated to 95 °C and then gradually cooled to ~20 °C. The annealed two-strand oligonucleotide was separated from the unbound material by column chromatography (micro Bio-Spin 25 column, USA Scientific, Ocala, FL).

Estimation of *in vitro* inhibitory activation. For estimation of inhibitory activation against 3'-processing and chain transfer reactions, wild-type integrase was preincubated to a final concentration of 200 nmol L⁻¹ with an inhibitor in a reaction buffer (0.05 M NaCl, 0.001 M HEPES buffer (pH 7.5), 5 · 10⁻⁵ M EDTA, 5 · 10⁻⁵ M dithiothreitol, 10% glycerol (wt v⁻¹), 7.5 · 10⁻³ M MnCl₂, bovine serum albumin (0.1 mg mL⁻¹), 0.01 M 2-mercaptoethanol, 10% DMSO, and 0.025 M MOPS (pH 7.2)) at 30 °C for 30 min. Then 20 · 10⁻⁹ M 5-terminal ³²P-labeled oligonucleotide was added and the resulting mixture was additionally incubated for 1 h. The reaction mixture was precipitated by adding a 50-fold excess of a dye (98% deionized formamide, 0.01 M EDTA, 0.025% xylene cyanol, and 0.025% bromophenol blue). A sample (5 µmol/L) of the reaction mixture was analyzed by electrophoresis in polyacrylamide gel (0.09 M trisborate (pH 8.3), 0.002 M EDTA, 20% acrylamide, and 8 M urea). The gels were dried *in vacuo*, placed in a PhosphorImager cassette, and scanned with a Typhoon 8610 Variable Mode Imager instrument (Amersham Biosciences). Quantitative estimation was done with the Image Quant 5.2 software. The inhibition value (in percent) was calculated by the following formula:

$$\%I = 100 \cdot [1 - (D - C)/(N - C)],$$

where *C*, *N*, and *D* are the percentages of the fraction of the 21-unit substrate converted into a 19-unit substrate (*via* 3'-processing) or a DNA and DNA+IN chain transfer product in the absence and in the presence of the inhibitor, respectively. The IC₅₀ values were determined by plotting the logarithm of the inhibitor concentration as a function of %*I* versus the concentration required to inhibit the catalytic activity by 50%.

Cytotoxicity assays. The cancer cell line HCT-116 of the human large intestine was represented as a monolayer culture (8 · 10³ compartments). The cells were cultured in 96-well plates. The compounds were dissolved in the RPMI-1640 medium (containing 10% bovine serum albumin) to a concentration of 10 µmol L⁻¹. After 72-h culturing, a solution of MTT (5 mg mL⁻¹, 20 µL) was added to each well and the compartments were incubated for 4 h. Then the supernatant was poured from each compartment and associated crystalline compartments with MTT were dissolved in DMSO (150 µL per well) with stirring at ~20 °C. Absorbance at 570 nm was measured on an absorbance microplate reader.

Molecular docking. All compounds were examined for docking at an active IN site to determine their biologically active conformations and possible binding mechanism. Molecular docking calculations were performed with the GOLD v.3.2 and eHITS software on the crystal structures of IN. The GOLD program with a standard set of parameters was run on a Linux multiprocessor workstation and a Silicon Graphics Onyx 24-processor workstation. The eHITS program was started up on Sun Grid Compute Utility dual-processor nodes accessible on <http://Network.com>.

Structures of all the inhibitors were modeled with the Catalyst program (Accelrys, Inc.) on a Linux multiprocessor workstation and a Silicon Graphics Onyx 24-processor workstation. The Poling algorithm embedded in Catalyst was used for

modeling conformations of each compound. All possible unique conformations were generated in an energy range more than 20 kcal mol⁻¹ wide with the latest-generation procedure implemented in Catalyst. For docking modeling of each ligand, its lowest energy conformation was used.

Structure of the catalytic domain of integrase. Modeling of docking was performed with the crystal structure of IN containing the bound inhibitor 5CITEP (1-(5-chloroindol-3-yl)-3-hydroxy-3-(2*H*-tetrazol-5-yl)prop-2-enone) from the Protein Data Bank (PDB No. 1QS4). Some residues were unbound in each chain of the crystal structure. Chain *A*, which is bound to 5CITEP and promotes the formation of a chelate complex of the Mg²⁺ ion with the amino acid residues D64 and D116, was chosen as a docking target. Four unbound residues of chain *A* were modeled. From chain *B* of 1QS4, the residues Y143 and N144 were obtained, while the residues I141 and P142 were modeled on the basis of the 1BIS IN structure through a linking bridge. Complete modeling and minimization of the protein was performed with the Insight II software (Accelrys, Inc.). All the water molecules were removed from the crystal structure, while the Mg²⁺ ions were left intact. Then hydrogen atoms and acid-base residues in the ionic forms were added to the active site. The acidity of the medium (pH 7.0), at which the protein was protonated, was maintained throughout the docking calculations. The ligand (5CITEP) was removed to make its binding site accessible during the modeling.

Parameters of the binding of compounds 8–10 to the active site of integrase according to the data from molecular docking calculations. For each tested compound, the molecular weight, the number of rotating bonds, the number of hydrogen bond donors, the number of hydrogen bond acceptors, and S+log*P* were calculated with the ADMET Predictor software (Simulations Plus, Inc.). The quantity Alog*P* 98 was calculated with the Discovery studio program (Accelrys, Inc.).

References

1. S. P. Gupta, M. S. Babu, N. Kaw, *J. Enzyme Inhib.*, 1999, **14**, 109.
2. G. Patonay, J. Salon, J. Sowell, L. Strekowski, *Molecules*, 2004, **9**, 40.
3. K. D. Volkova, V. B. Kovalska, A. I. Tatarts, I. D. Patsenker, D. V. Krivorotenko, S. M. Yarmoluk, *Dyes Pigm.*, 2007, **72**, 285.
4. T. K. Chiu, D. R. Davies, *Curr. Top. Med. Chem.*, 2004, **4**, 965.
5. Yu. Yu. Agapkina, T. A. Prikazchikova, M. A. Smolov, M. B. Gottikh, *Usp. Biol. Khim. [Advances in Biological Chemistry]*, 2005, **45**, 87 (in Russian).
6. Y. Pommier, N. Neamati, *Adv. Virus Res.*, 1999, **52**, 427.
7. Y. Pommier, C. Marchand, N. Neamati, *Antiviral Res.*, 2000, **47**, 139.
8. R. Dayam, N. Neamati, *Curr. Pharm. Des.*, 2003, **9**, 1789.
9. R. Dayam, T. Sanchez, N. Neamati, *J. Med. Chem.*, 2005, **48**, 8009.
10. J. Deng, T. Sanchez, I. Q. Al-Mawsawi, R. Dayam, R. A. Yunes, A. Garofalo, M. B. Bolger, N. Neamati, *Bioorg. Med. Chem.*, 2007, **15**, 4985.
11. D. J. Hazuda, P. Felock, M. Witmer, A. Wolfe, K. Stillmock, J. A. Grobler, A. Espeseth, L. Gabryelski, W. Schleif, C. Blau, M. D. Miller, *Science*, 2000, **287**, 646.

12. D. J. Hazuda, N. J. Anthony, R. P. Gomez, S. M. Jolly, J. S. Wai, L. Zhuang, T. E. Fisher, M. Embrey, J. P. Guare, Jr., M. S. Egbertson, J. P. Vacca, J. R. Huff, P. J. Felock, M. V. Witmer, K. A. Stillmock, R. Danovich, J. Grobler, M. D. Miller, A. S. Espeseth, L. Jin, I.-W. Chen, J. H. Lin, K. Kassahun, J. D. Ellis, B. K. Wong, W. Xu, P. G. Pearson, W. A. Schleif, R. Cortese, E. Emini, V. Summa, M. K. Holloway, S. D. Young, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 11233.
13. V. F. Traven, A. V. Manaev, T. A. Chibisova, *Dyes Pigm.*, 2003, **58**, 41.
14. A. V. Manaev, K. V. Tambov, V. F. Traven, *Zh. Org. Khim.*, 2008, **44**, 1064 [*Russ. J. Org. Chem. (Engl. Transl.)*, 2008, **44**].
15. V. F. Traven, A. V. Manaev, I. V. Voevodina, I. N. Okhrimenko, *Izv. Akad. Nauk, Ser. Khim.*, 2008, 1479 [*Russ. Chem. Bull., Int. Ed.*, 2008, **57**, 1508].
16. A. V. Manaev, I. N. Okhrimenko, V. F. Traven, *Izv. Akad. Nauk, Ser. Khim.*, 2008, 1701 [*Russ. Chem. Bull., Int. Ed.*, 2008, **57**].
17. Ya. A. Ivanenkov, E. V. Bovina, K. V. Balakin, *Usp. Khim.*, 2009, **78**, 503 [*Russ. Chem. Rev. (Engl. Transl.)*, 2009, **78**].
18. W. P. Walters, M. T. Stahl, M. A. Murcko, *Drug Discovery Today*, 1998, **3**, 160.
19. J. Bajorat, *Nat. Rev. Drug Discovery*, 2002, **1**, 882.
20. K. V. Balakin, Ya. A. Ivanenkov, A. V. Skorenko, S. N. Kovalenko, I. A. Zhuravel, V. P. Chernykh, *Ukr. Zh. Org. Farm. Khim. [Ukrainian Journal of Organic and Pharmaceutical Chemistry]*, 2004, **2**, 47 (in Russian).
21. Y. Nikolsky, K. V. Balakin, Y. A. Ivanenkov, A. A. Ivashchenko, N. P. Savchuk, *PharmaChem*, 2003, **4**, 68.
22. N. P. Savchuk, S. E. Tkachenko, K. V. Balakin, *Methods and Principles in Medicinal Chemistry*, Vol. **23**, *Cheminformatics in Drug Discovery*, Ed. T. I. Oprea, Wiley—VCH, Weinheim, 2005, 287.
23. K. V. Balakin, D. Sc. (Chem.) Thesis, Ivanovo State Univ., Ivanovo, 2005 (in Russian).
24. Y. Nikolsky, K. V. Balakin, Y. A. Ivanenkov, A. A. Ivashchenko, N. Ph. Savchuk, *J. Pharm. Chem.*, 2003, **4**, 68.
25. Y. A. Ivanenkov, K. V. Balakin, A. V. Skorenko, S. E. Tkochenko, N. Ph. Savchuk, A. A. Ivashchenko, Y. Nikolsky, *Chem. Today*, 2003, **21**, 72.
26. S. V. Trepalin, V. A. Gerasimenko, A. V. Kozyukov, N. Ph. Savchuk, A. A. Ivashchenko, *J. Chem. Inf. Model.*, 2002, **42**, 249.
27. Y. A. Ivanenkov, L. M. Khandarova, in *Pharmaceutical Data Mining: Approaches and Applications for Drug Discovery*, Ed. K. V. Balakin, John Wiley and Sons, Inc., New York, 2009, 457.
28. K. V. Balakin, Y. A. Ivanenkov, N. P. Savchuk, in *Chemo-genomics: Methods and Applications*, Ed. E. Jacoby, Humana Press—Springer, New York, 2009, p. 21.
29. Y. A. Ivanenkov, E. V. Bovina, K. V. Balakin, *Usp. Khim.*, 2009, **78** [*Russ. Chem. Rev. (Engl. Transl.)*, 2009, **78**, 503].
30. I. Pletnev, Y. Ivanenkov, A. Tarasov, in *Pharmaceutical Data Mining: Approaches and Applications for Drug Discovery*, Ed. K. V. Balakin, John Wiley and Sons, Inc., New York, 2009, 425.
31. L. Guttman, *Psychometrika*, 1954, **19**, 149.
32. R. B. Catell, S. A. Vogelmann, *Multi Behav. Res.*, 1977, **12**, 289.
33. C. S. Peirce, *A Theory of Probable Inference*, *Studies in Logic*, Little, Brown, and Co., Boston, 1883, 126.
34. A. Pendri, N. A. Meanwell, K. M. Peese, M. A. Walker, *Expert Opin. Ther. Pat.*, 2011, **21**, 1173.
35. K. Ramkumar, K. V. Tambov, R. Gundla, A. V. Manaev, V. Yarovenko, V. F. Traven, N. Neamati, *Bioorg. Med. Chem.*, 2008, **16**, 8988.

*Received June 14, 2011;
in revised form October 28, 2011*