



# 3D/4D QSAR технологии для анализа и прогноза биологической активности

Потёмкин Владимир Александрович  
Зав. лабораторией компьютерного  
моделирования лекарственных средств

Мой Соавтор:  
Гришина Мария Александровна,  
Главный научный сотрудник  
лаборатории компьютерного  
моделирования лекарственных средств,  
Высшая медико-биологическая школа ЮУрГУ



# Создание лекарства



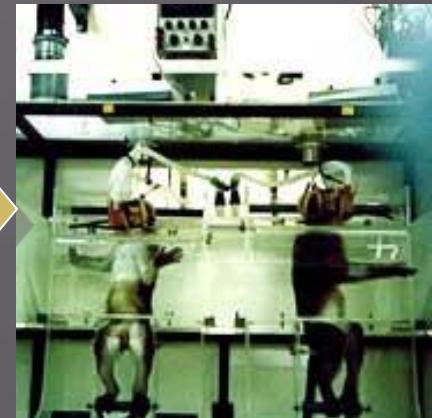
Химическая лаборатория



*in vitro* тесты  
5 000-10 000  
соединений



*in vivo* тесты – 250 соединений



Производство



Создание технологии



Клинические испытания

Фаза I – 10 соединений

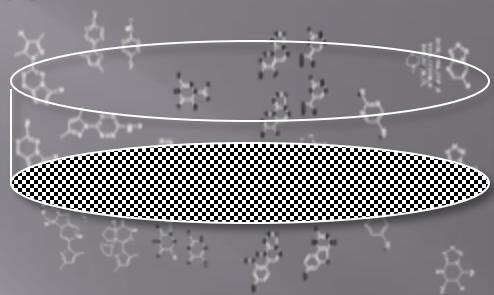
Фаза IV – 1 соединение



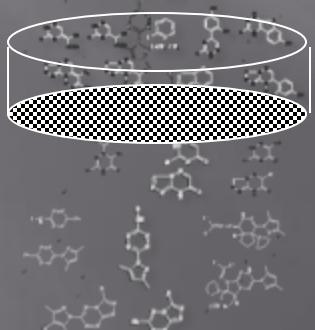
Аптека

Средние затраты на 1 лекарство \$350 миллионов,  
Но порой достигает \$5.5 миллиардов

# Роль компьютерных методов



Виртуальный Скрининг,  
Методы интеллектуального  
анализа данных



3R conception: to  
Reduce, Remove and  
Refine animal usage



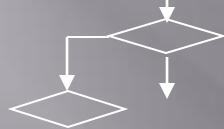
Молекулярный дизайн

# Методы и дескрипторы



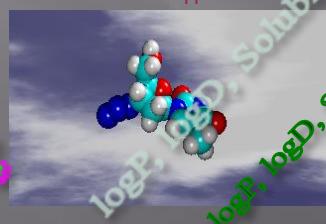
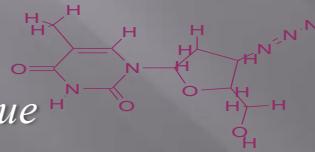
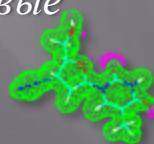
- Регрессионные
- Полиномиальные
- Нелинейные
- PLS

$$BA = f(x_1, x_2, \dots, x_n)$$

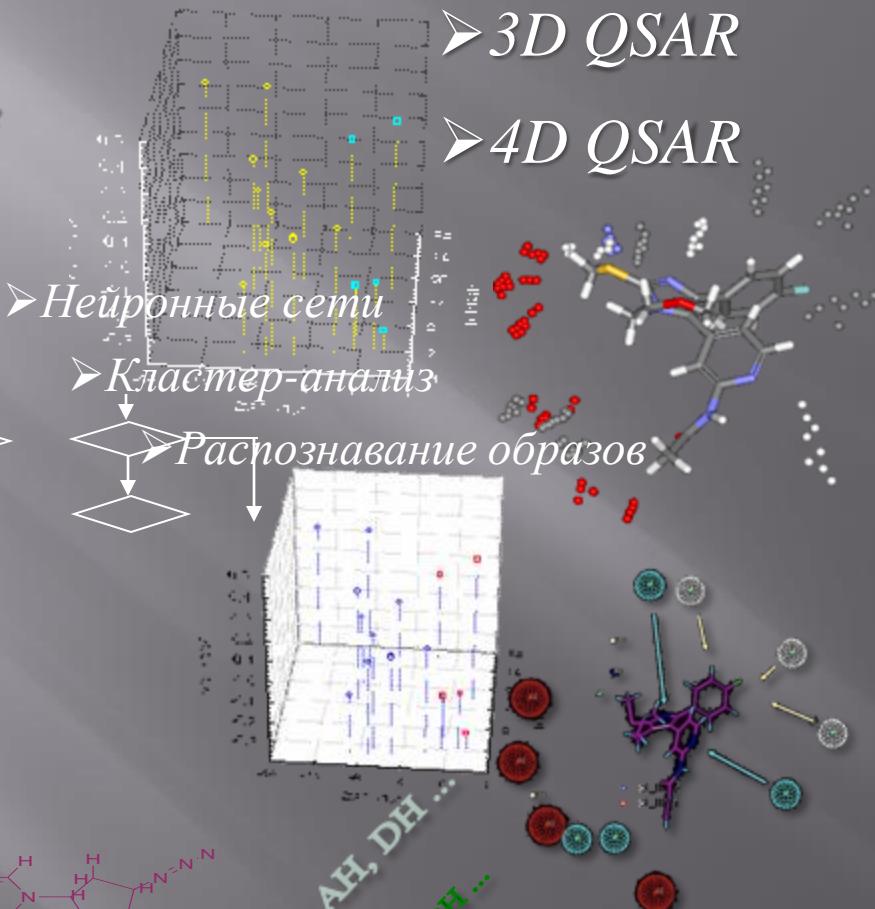


## Дескрипторы

- Топологические
- Геометрические
- Квантовые
- ADME



*logP, logD, Solubility, AH, DH ...*



➤ 3D QSAR

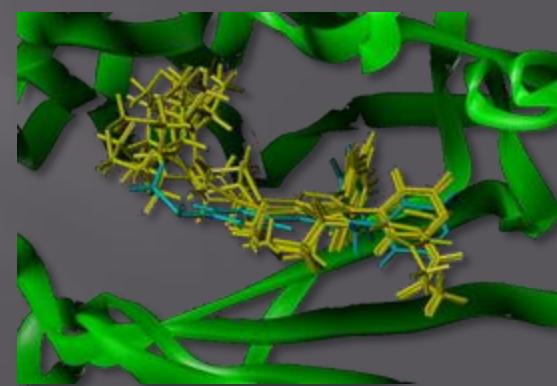
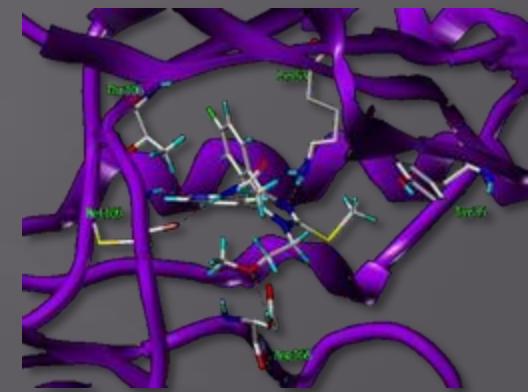
➤ 4D QSAR

➤ Нейронные сети

➤ Кластер-анализ

➤ Распознавание образов

➤ Молекулярный докинг



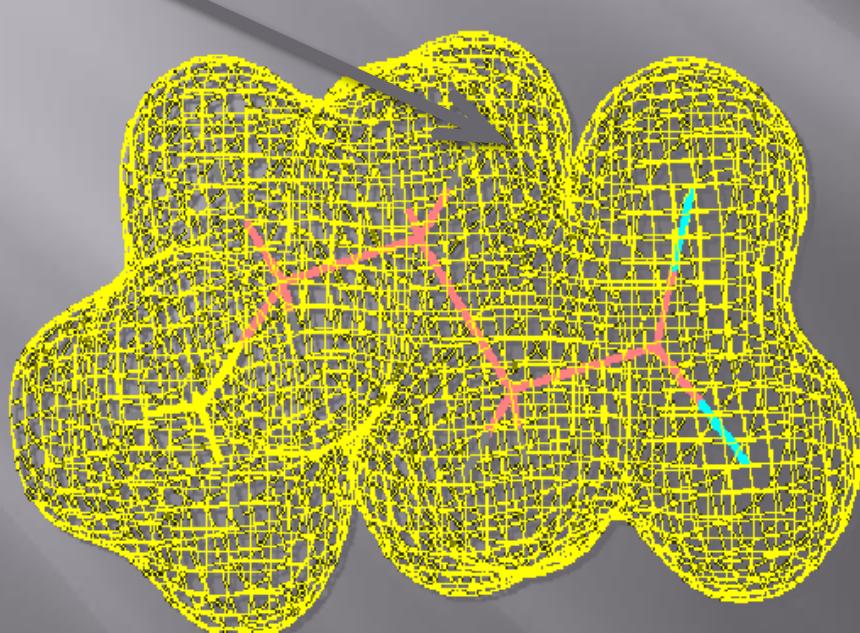


## 3D QSAR

# 3Dimensional Quantitative Structure-Activity Relationship

Анализ экстериора

т.е. молекулярной поверхности

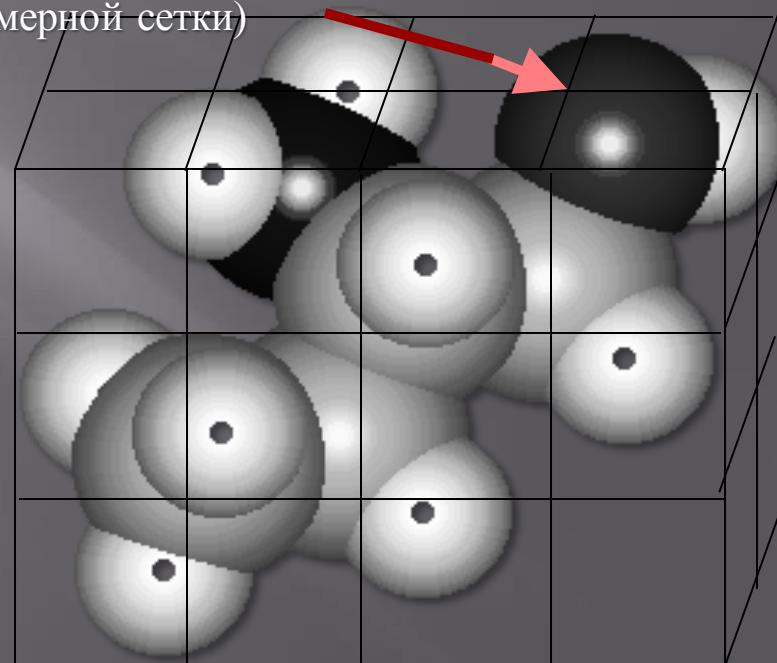


$$BA = a + bE$$

Например: PARM  
(Pseudo-Atomic Receptor Model)

Анализ интериора

т.е. внутреннего устройства молекулы  
(часто это пересечения атомов с точками  
3-мерной сетки)



Например: HASL  
(Hypotetical Active Site Lattice)

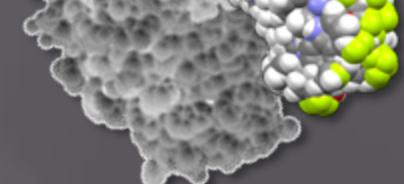
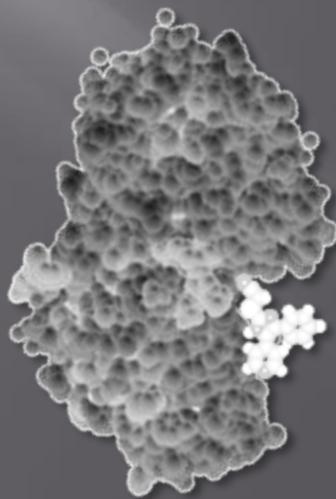
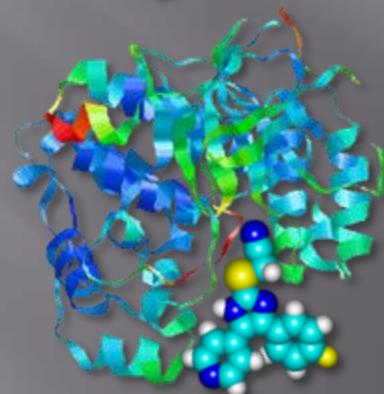
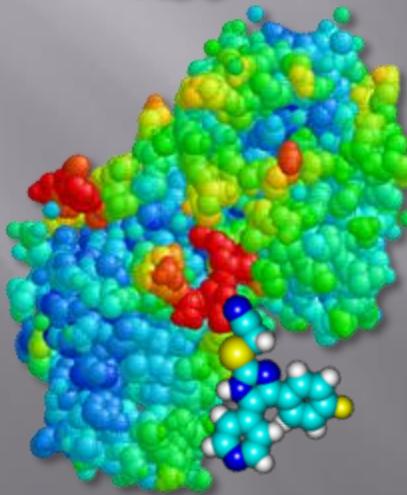
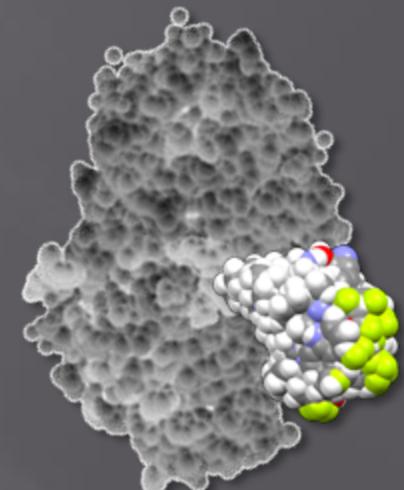
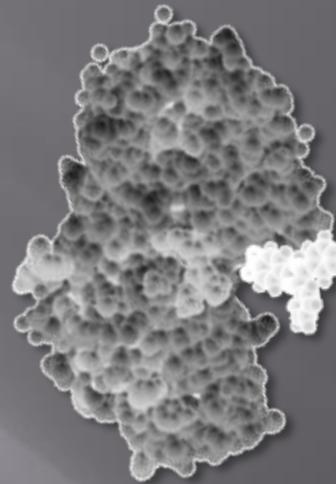
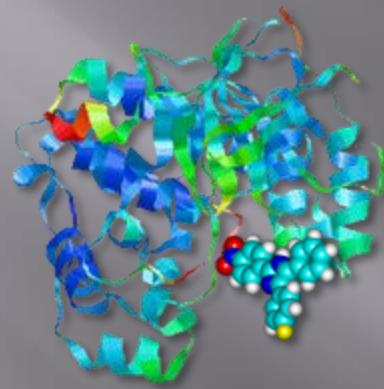
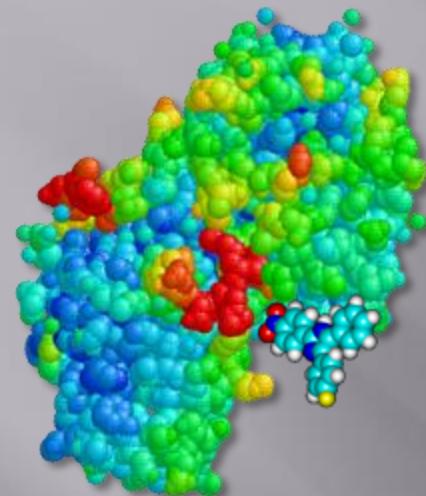
CoMFA

(Comparative Molecular Field Analysis)

$$BA = \sum k_i \delta_i$$



## Принцип комплементарности в основе алгоритмов BiS (Biological Substrate Search) и CiS (Cinderella's Shoe)



1. BiS/MC (multi-conformational), свидетельство о регистрации N 2007613594 07.09.2007 Гришина М.А., Потемкин В.А
2. "BiS\_p", свидетельство о регистрации N 2010615221 13.08.2010. Гришина М.А., Потемкин В.А
3. "CiS\_p", свидетельство о регистрации N 2010615222 13.08.2010. Гришина М.А., Потемкин В.А

# 3D QSAR BiS и CiS алгоритмы



## ВВОД СТРУКТУР

Выборка молекул с известными  
активностями, декартовы  
координаты атомов



## STEP 1

Создание псевдоатомной модели



## STEP 2

Ориентация молекул



## STEP 3

Поправки потенциалов  
взаимодействий, радиусов и зарядов  
псевдо-атомов



## ВЫВОД

- Ориентированные молекулы
- Энергии и силы взаимодействия в системе “псевдо-рецептор - лиганд”
- Уравнение взаимосвязи биологической активности и характеристик взаимодействий в системе “псевдо-рецептор - лиганд”
- Рекомендации к молекулярному дизайну лекарств

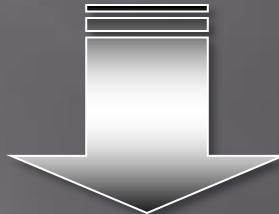
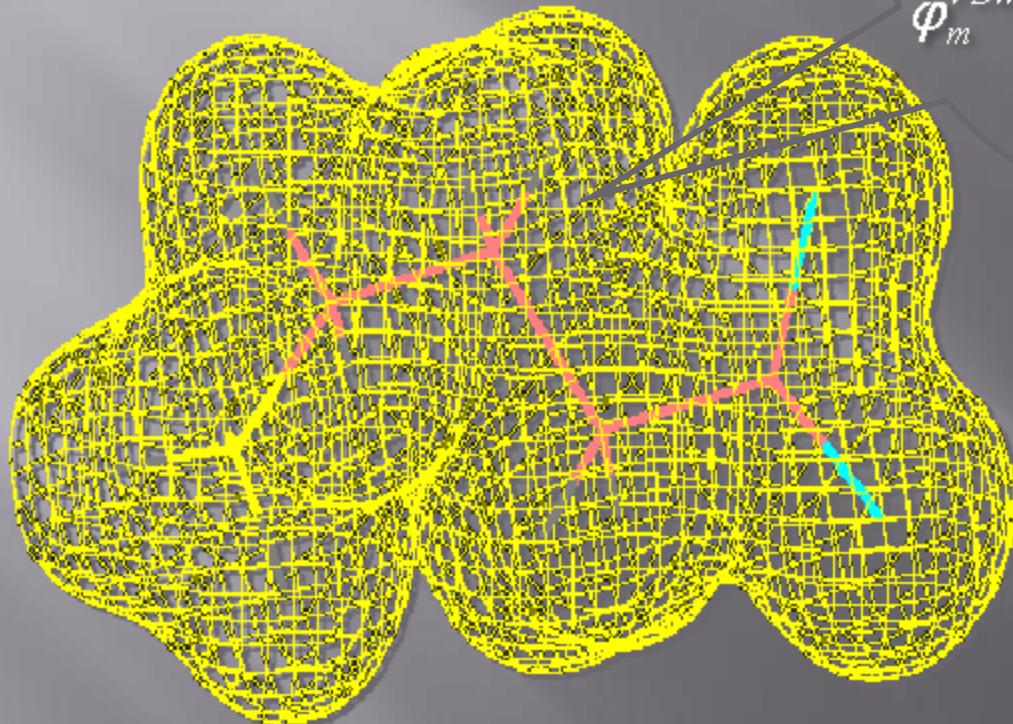
# STEP 1

Создание псевдоатомной модели



$$\boldsymbol{\varphi}_m^q = \sum_{i=1}^N \frac{q_i}{R_{im}} \mathbf{k}$$

$$\boldsymbol{\varphi}_m^{VDW} = -2 \sum_{i=1}^N V_{im} \frac{2^3 r_i^3}{R_{im}^6}$$



Характеристики псевдо-атомов

$$q_m = -\frac{\boldsymbol{\varphi}_m^q}{\sum_{i=1}^N \frac{\mathbf{k}}{R_{im}}}$$

$$\mathbf{r}_m = \sqrt[3]{-\frac{\boldsymbol{\varphi}_m^{VDW}}{2^3 \sum_{i=1}^N \frac{2V_{im}}{R_{im}^6}}}$$

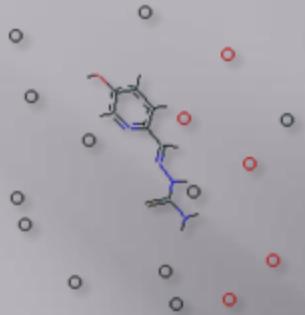
Potemkin V., Grishina M. Principles for 3D/4D QSAR classification of drugs// Drug Discovery Today.- 2008.- V. 13.- No. 21/22.- P. 952 – 959.

Potemkin V.A., Pogrebnov A.A., Grishina M.A. Technique for Energy Decomposition in the Study of "Receptor-Ligand" Complexes// J. Chem. Inf. Model.- 2009.- Vol. 49.- № 6.- P. 1389 - 1406.

## STEP 2

Ориентация в псевдо-атомном поле

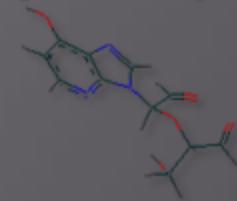
Первая молекула и ее поле



Поле первой молекулы



Вторая молекула



## STEP 3

Поправка потенциалов поля, зарядов  
и радиусов псевдо-атомов

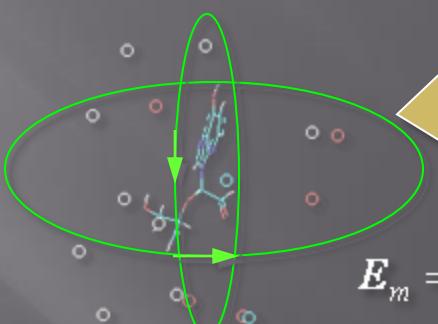
$$\varphi_m^q = \varphi_m^{q'} + \varphi_m^{q''}$$

$$\varphi_m^{VDW} = \varphi_m^{VDW'} + \varphi_m^{VDW''},$$

$$q_m = -\frac{\varphi_m^q}{\sum_{i=1}^N \frac{k}{R_{im}}}$$

$$r_m = \sqrt[3]{-\frac{2V_{im}}{2^3 \sum_{i=1}^N \frac{R_{im}^6}{R_{im}^6}}}$$

Оптимизация расположения  
второй молекулы в поле первой



$$E_m = \sum_{i=1}^N \left( \frac{k q_i q_m}{R_{im}} - 2V_{im} \frac{(r_m + r_i)^6}{R_{im}^6} + V_{im} \frac{(r_m + r_i)^{12}}{R_{im}^{12}} \right)$$

$$P = 1 - \prod_{m=1}^M (1 - p_m)$$

$$p_m = \exp\left(-\frac{E_m}{RT}\right)$$



## ВЫВОД

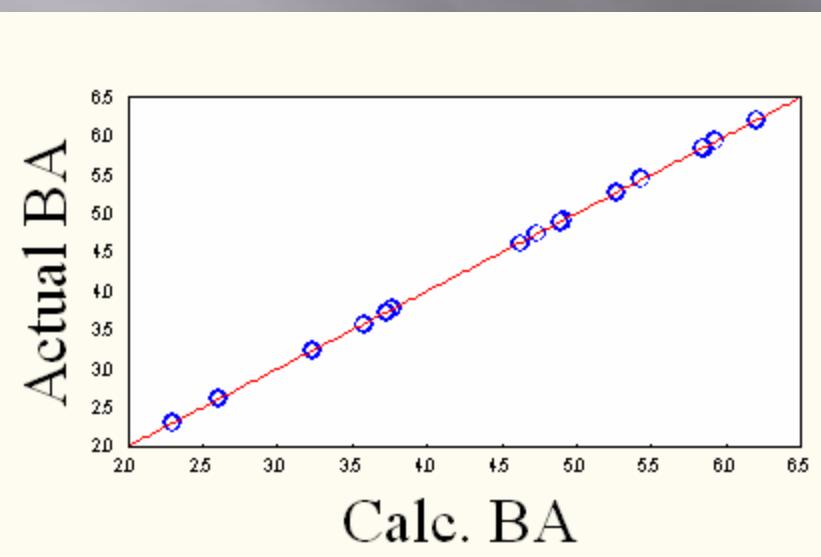
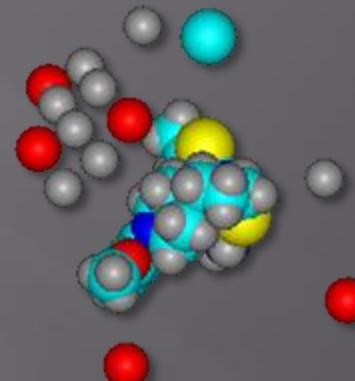
- Ориентированные молекулы
- Энергии и силы взаимодействия в системе “псевдо-рецептор - лиганд”
- Уравнение взаимосвязи биологической активности и характеристик взаимодействий в системе “псевдо-рецептор - лиганд”
- Рекомендации к молекулярному дизайну лекарств



Алгоритмы BiS и CiS позволяют

Построить псевдо-атомную модель рецептора

$$BA = \mathbf{a}_0 + \sum_{i=1}^K \mathbf{a}_i E_{mi} + \sum_{j=1}^L \mathbf{b}_j F_{mj}$$



Уравнение взаимосвязи биологической активности и характеристик взаимодействий в системе “псевдо-рецептор - лиганд”

$$E_m = \sum_{i=1}^N \left( \frac{k \mathbf{q}_i \mathbf{q}_m}{R_{im}} - 2V_{im} \frac{(r_m + r_i)^6}{R_{im}^6} + V_{im} \frac{(r_m + r_i)^{12}}{R_{im}^{12}} \right) + \frac{k_m x^2}{2}$$

Potemkin V., Galimova O., Grishina M. Cinderella's Shoe for virtual drug discovery screening and design// Drugs of the Future.- 2010.- Vol. 35.- P. 14 – 15.

$$k_m = \frac{\kappa T}{A^2}$$

# Анализ интериора 3D/4D QSAR алгоритм ConGO

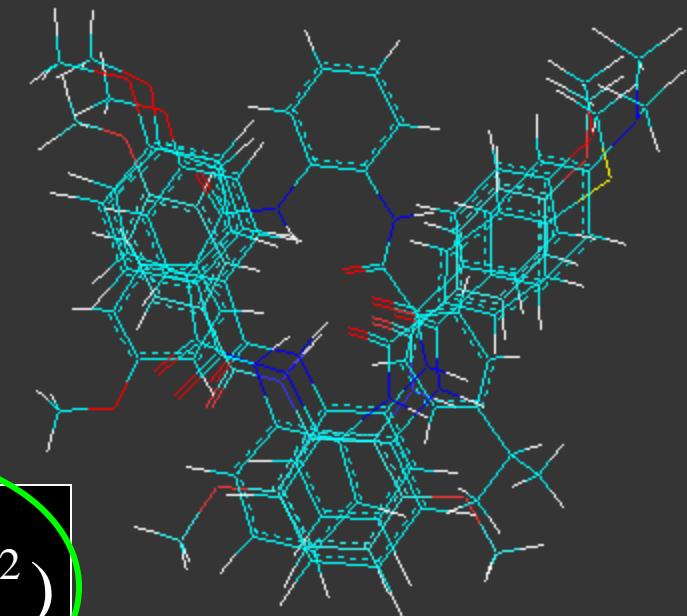


$$\Delta\Phi_i = \frac{1}{M-1} \sum_{m=1}^M \frac{(|\Psi_{mi}|^2 - |\psi_{mi}|^2)^2}{(|\Psi_{mi}|^2 + |\psi_{mi}|^2)^2}$$

$$BA_k = (c_1 + c_2 \sum_{t=1}^T |\psi_{kt}|^2 - c_3 \sum_{u=1}^U |\psi_{ku}|^2)$$

Фармакофорный фрагмент

Анти-фармакофорный фрагмент





## Алгоритм AlteQ для оценки электронной плотности

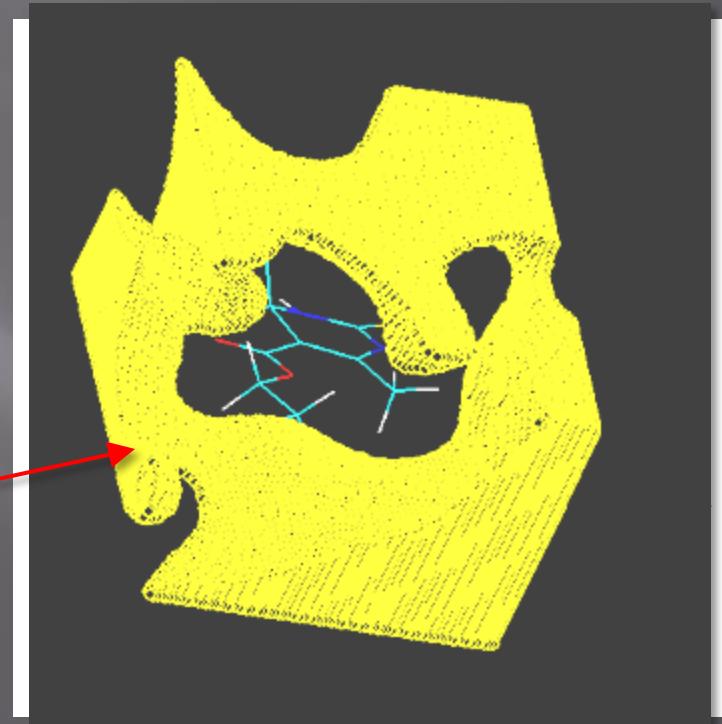
$$|\psi|^2 = \sum a_i 10^{b_i R}$$

Области с высокой  
погрешностью оценки  
электронной плотности

**AlteQ – не более 0.4%**  
**молекулярного пространства**

**DFT B3LYP/6-311G(d,p) ~ 20%**

$$\left| \rho_m - \rho_{m_{quant}} \right| \geq 10$$

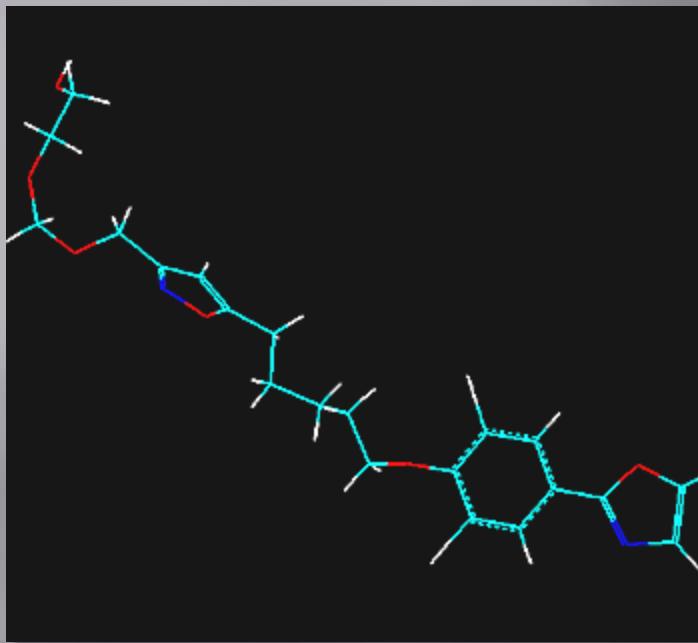


Salmina E., Grishina M.A., Potemkin V.A. An approximation of the Cioslowski–Mixon bond order indexes using the AlteQ approach// J. Comp.-Aid. Mol. Des.- 2013.- V. 27. No. 9.- P. 793-805.

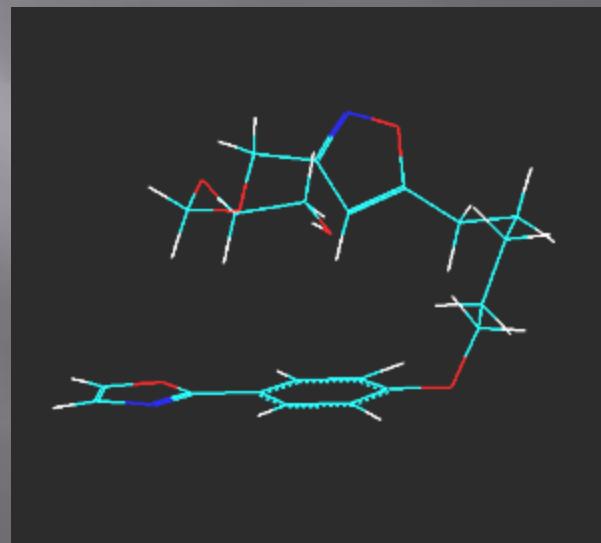
V. Potemkin, M. Grishina //J. of Comp-Aid. Mol. Des., 2008

Grishina M., Potemkin V. et al // Comp. Theor. Chem.- 2016.- Vol. 1091.- P. 122–136

# Ингибитор риновируса HRV14 (1hri)



Активный конформер  
(геометрия из ко-кристалла  
“рецептор - лиганд”)

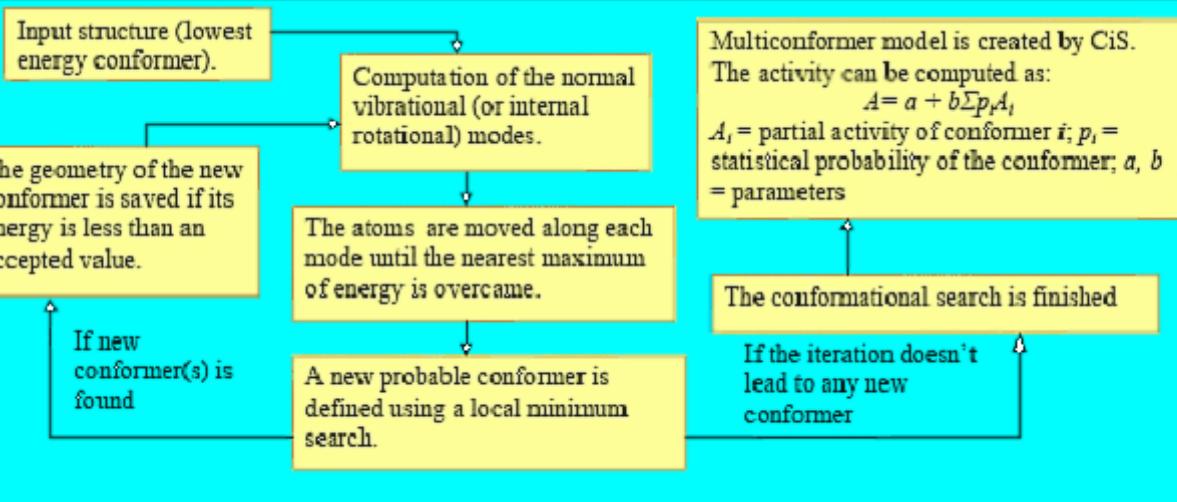


Конформеры , отвечающие  
минимуму потенциальной  
энергии (энантиоконформеры)  
(силовое поле ММЗ)

\* <http://www.wwpdb.org>

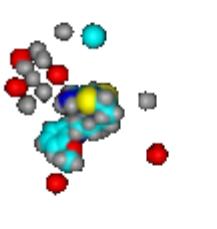


# Алгоритм MultiGen для 4D QSAR анализа

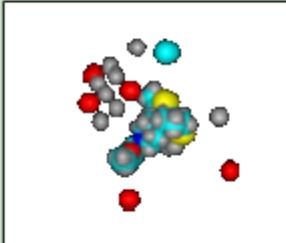


Major conformers for the most and the least active compounds of the set are shown. Statistical probabilities and biological activities for each of them are shown under figures.

The most active molecule:

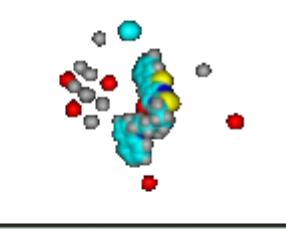


$pIC_{50} = -23.24$   
 $p = 0.482$

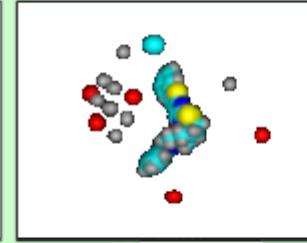


$pIC_{50} = 42.00$   
 $p = 0.482$

The least active molecule:

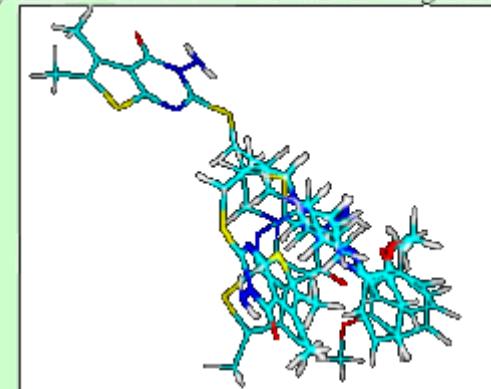


$pIC_{50} = 12.61$   
 $p = 0.472$

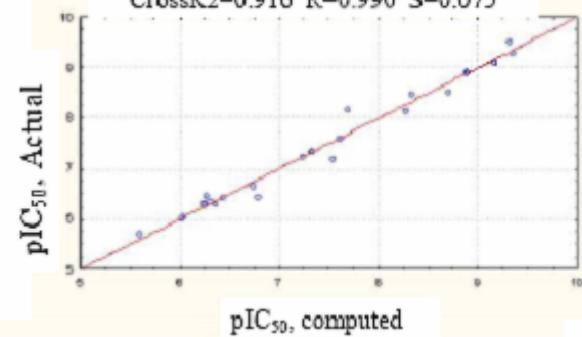


$pIC_{50} = -1.80$   
 $p = 0.472$

The result of superimposing are generally in agreement with the results of NMR-investigations.



LOO-plot for 5-HT<sub>1A</sub> receptor agonists  
CrossR2=0.916 R=0.990 S=0.075





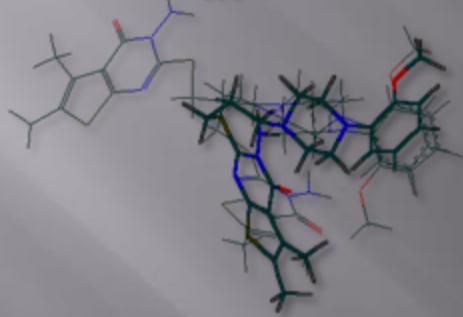
# 4D QSAR

Конформеры с низкой и высокой активностью

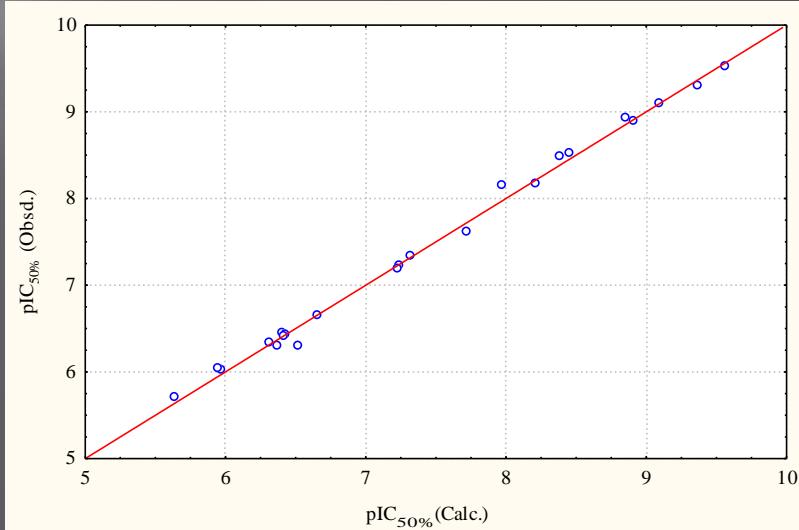
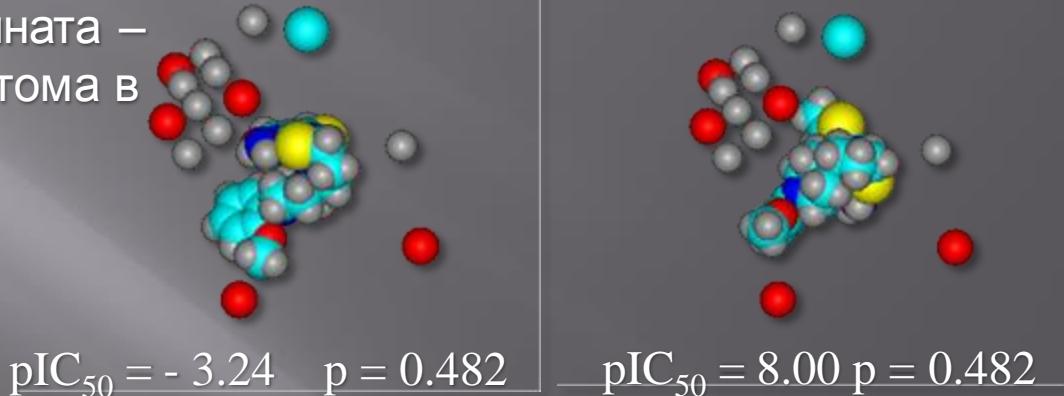
3 декартовы координаты, 4<sup>th</sup> координата – вероятность нахождения каждого атома в данной точке пространства

Мультиконформационная модель

агониста 5-HT<sub>1A</sub>-рецептора



$$BA = \sum b a_i$$



Potemkin V.A., Pogrebnoy A.A., Grishina M.A. Technique for Energy Decomposition in the Study of "Receptor-Ligand" Complexes// J. Chem. Inf. Model.- 2009.- Vol. 49.- № 6.- P. 1389 - 1406.

## Созданы модели для прогноза, виртуального скрининга и молекулярного дизайна для более 50 видов биологической активности



- Acyl-CoA-cholesterol transferase inhibitory activity
- Adrenoreceptor inhibitory activity
- (Anti-Hypertensive Activity)'
  - Alpha-R-receptor inhibitory activity
  - HIV1-protease inhibitory activity
  - Analgetic activity
  - Anti-Bacterial activity
  - Anti-Oxidant activity
  - Anti-Inflammatory activity (combined action)
  - COX1 inhibitory activity
  - COX2 inhibitory activity
  - LOX5 inhibitory activity
  - Anti-Tumor Alkylic activity
  - Anti-Tumor Topoisomerase I inhibitory activity
  - Anti-Tumor Topoisomerase II inhibitory activity
  - Anti-Tumor Dihydrofolate reductase inhibitory activity
  - Anti-Tumor DNA anti-metabolic activity
  - Anti-Tumor Antimitotic activity
  - Anti-Arrhythmic activity
  - Tuberculostatic Dihydrofolate reductase inhibitory activity
  - Anti-Tumor Cyclin-dependent kinase 4 inhibitory activity
  - Anti-Psychotic activity (diazepine site)
  - Human factor XA Inhibitory activity
  - HT51A inhibitory activity

- Anti-Inflammatory activity (p38 MAP-kinase inhibitors)
  - Anti-Inflammatory activity in vivo
  - Anti-Inflammatory activity in vivo (peritonitis)
  - Anti-Influenza A activity
  - Anti Awesky Disease activity
  - Anti-Influenza (Birds) activity
  - Anti Issyk-Kul Haemorrhagic Fever activity
  - Anti Infectious Laryngotracheitis activity
  - Anti-Encephalitic activity
  - Anti Crymean Haemorrhagic Fever activity
  - Anti Karelian Fever activity
  - Anti Rift valley Fever activity
  - Anti-Influenza B activity
  - Progestagenic activity
  - Gamma-radioprotector activity (mechanism I)
  - Gamma-Radioprotector activity (mechanism II)
  - Alpha-Radioprotector activity
  - Tuberculostatics (combined action)
  - Ecotoxicity
  - Vasorelaxant activity
  - Anti-Adenovirus activity
  - Anti-Influenza activity (Hong-Kong virus)
  - Anti Herpes Simplex virus activity
  - Metabolism at CYP450 3A4
  - Metabolism at CYP450 2D6

# Алгоритмы BiS, CiS, ConGO для виртуального скрининга и интеллектуального анализа данных



Соединения из баз данных компаний TOSLab InterBioScreen

ISIS/Base - [TOSLab\_Collection.db/all]

File Edit Options Object Database Search List Window Help

Forms Query Browse Update

<Root> 8969 of 17938

Search Domain: All

TOSLAB Database of heterocyclic compounds

Chemical structure: CN1C=CC2=C1C(=O)c1ccccc1S(=O)(=O)c3ccccc32

id

C8-0594
C8-0595
C8-0596
C8-0597
C9-0001
C9-0002
C9-0003
C9-0004
C9-0005
C9-0006
C9-0007
C9-0008
C9-0009
C9-0010
C9-0011
C9-0012
C9-0013
C9-0014
C9-0015
C9-0016
C9-0017
C9-0018
C9-0019
C9-0020
C9-0021
C9-0022
C9-0023
C9-0024
C9-0025
C9-0026
C9-0027
C9-0028
C9-0029

C8-0594 C8  
C<sub>24</sub>H<sub>17</sub>ClN<sub>2</sub>O<sub>2</sub>S  
432.93

Department of Technology for Organic Synthesis, Urals State Technical University 19, Mira str., 620002, Ekaterinburg, Russia  
<http://www.toslab.com>  
<http://toslab.com>  
e-mail: Vasily.Bakulev@TOSLab.com

818 of 197892

mail: screen@ibscreen.clg.ru  
Web: [www.ibscreen.com](http://www.ibscreen.com)

compounds

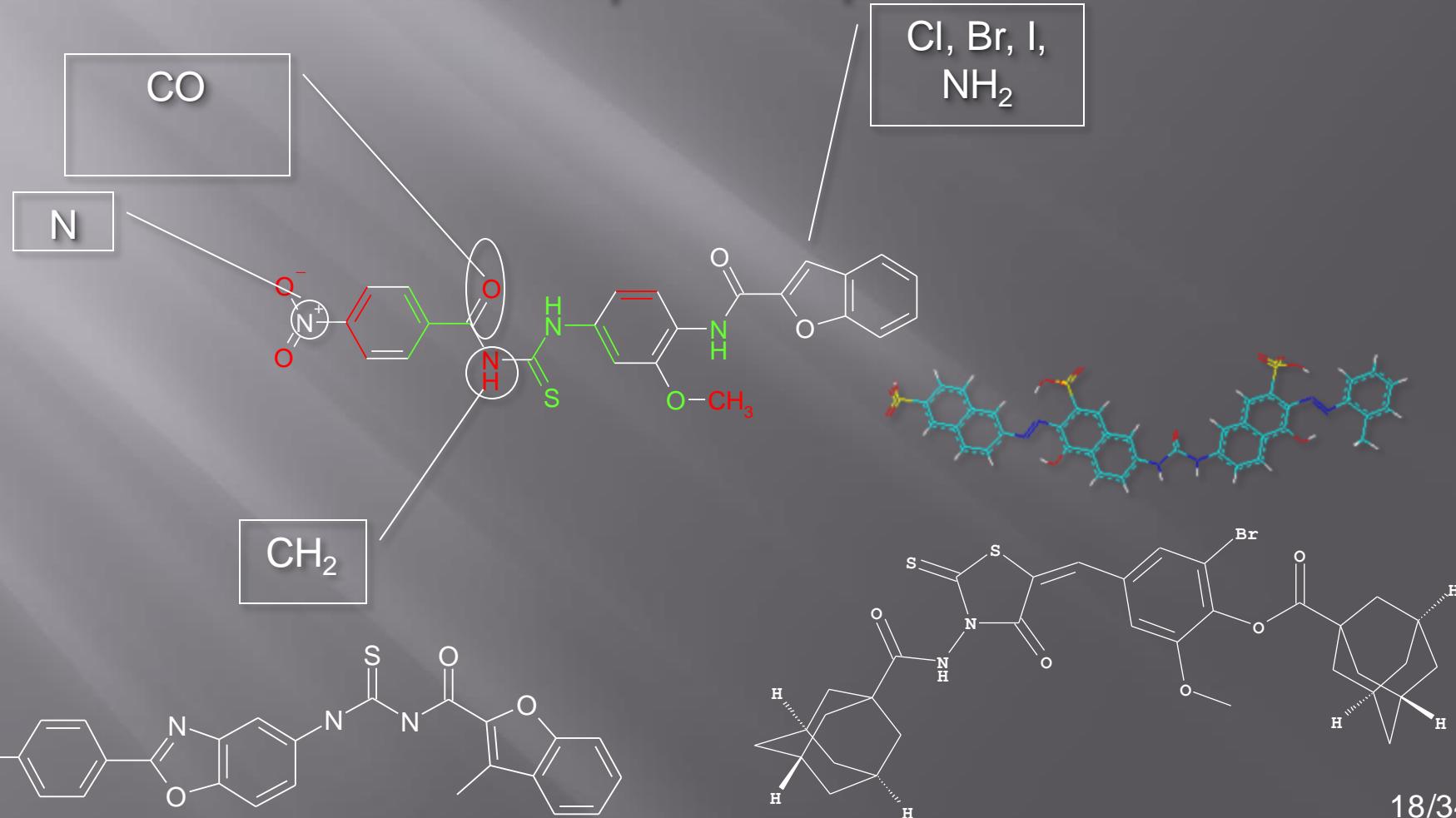
MW: 357.4128  
Index: SC  
TPSA: 55.4000 ClogP: 4.9820

- ингибиторы CDK4
- ингибиторы p38 МАР киназы
- ингибиторы COX-1
- ингибиторы COX-2
- ингибиторы LOX
- ингибиторы HIV1 протеазы
- 5HT<sub>1A</sub> агонисты

# Алгоритмы BiS, CiS, ConGO для молекулярного дизайна новых перспективных лекарств



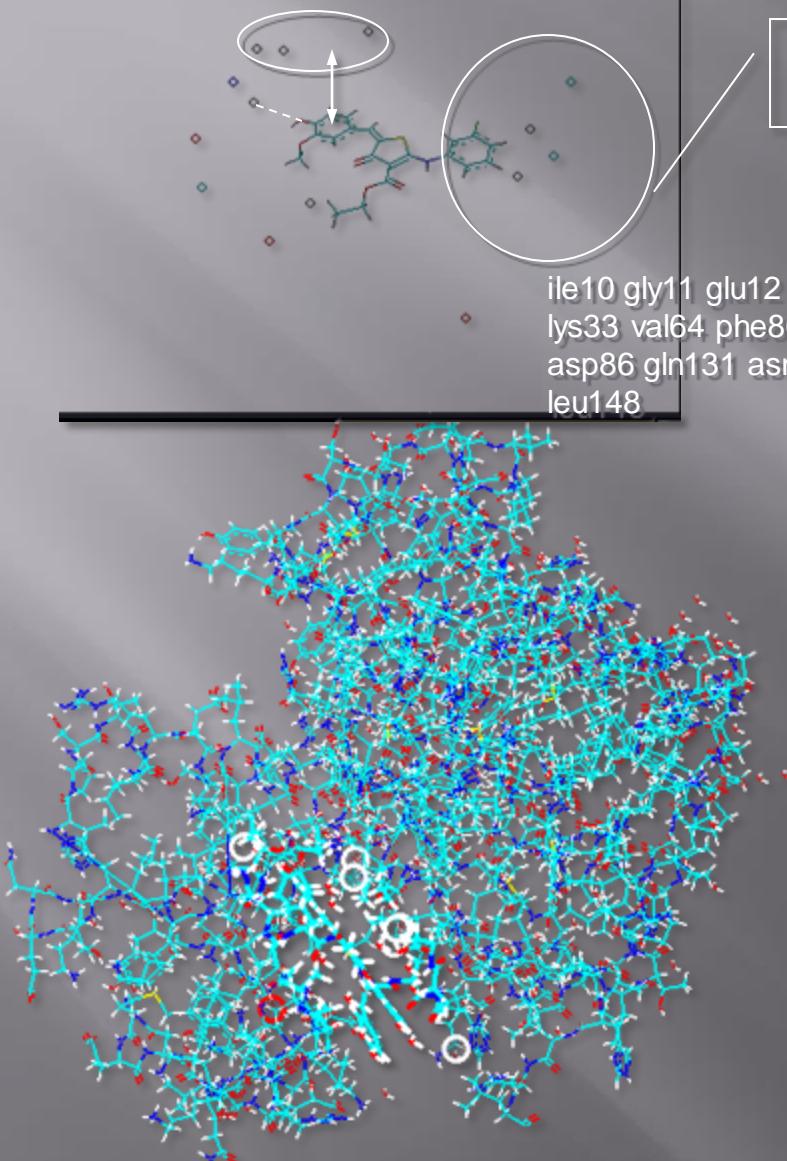
## Ингибитор HIV1-протеазы



# Алгоритмы BiS, CiS, ConGO для молекулярного дизайна новых перспективных лекарств

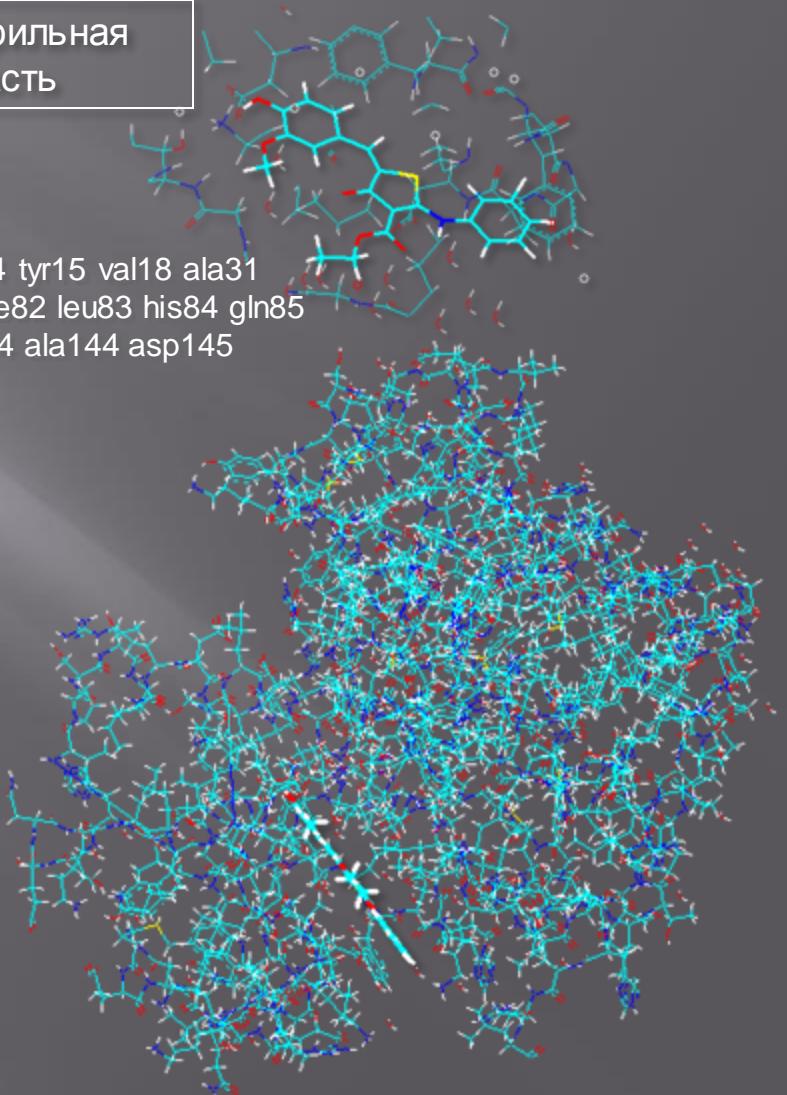
667 признано активными

## Ингибиторы CDK4

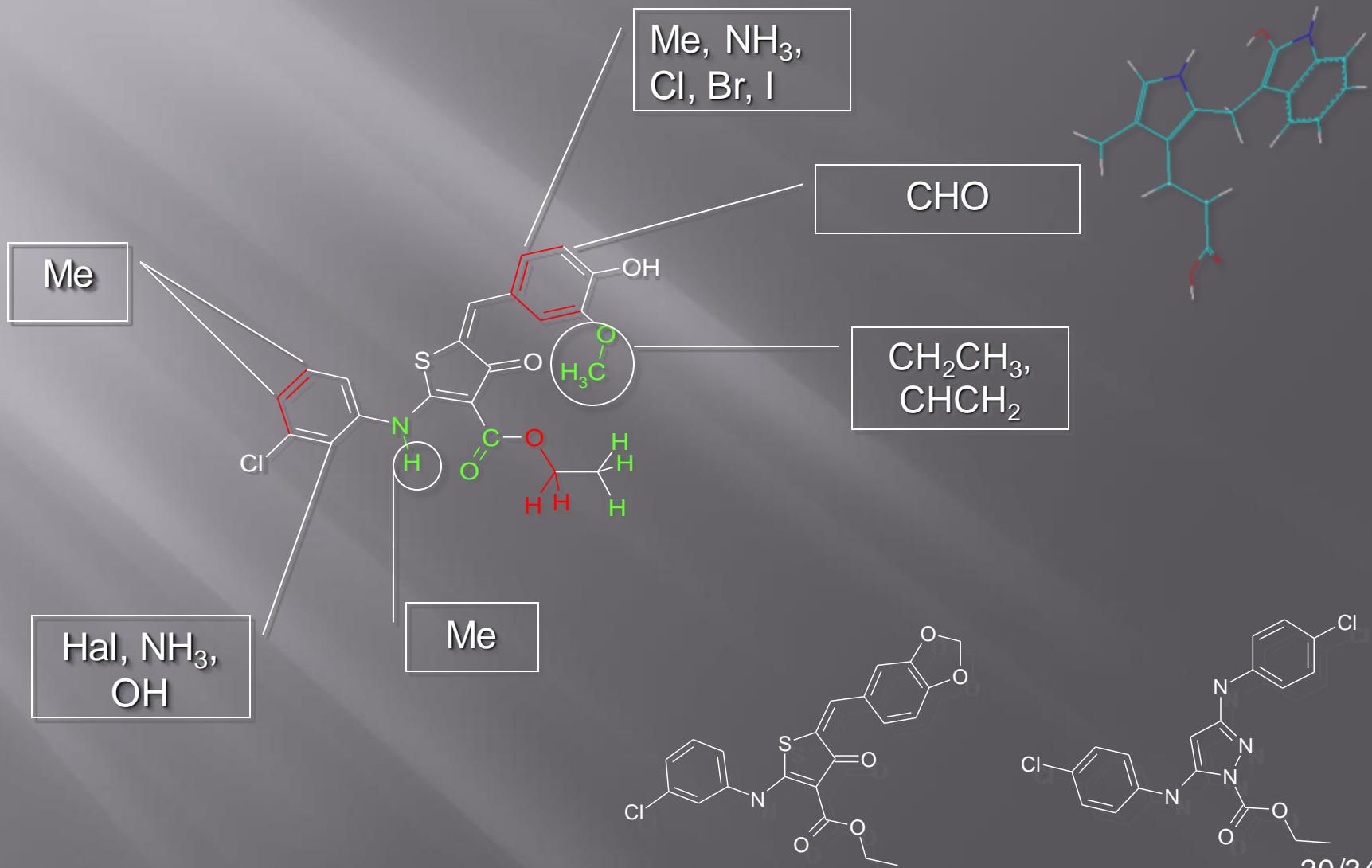


Липофильная  
часть

ile10 gly11 glu12 gly13 thr14 tyr15 val18 ala31  
lys33 val64 phe80 glu81 phe82 leu83 his84 gln85  
asp86 gln131 asn132 leu134 ala144 asp145  
leu148



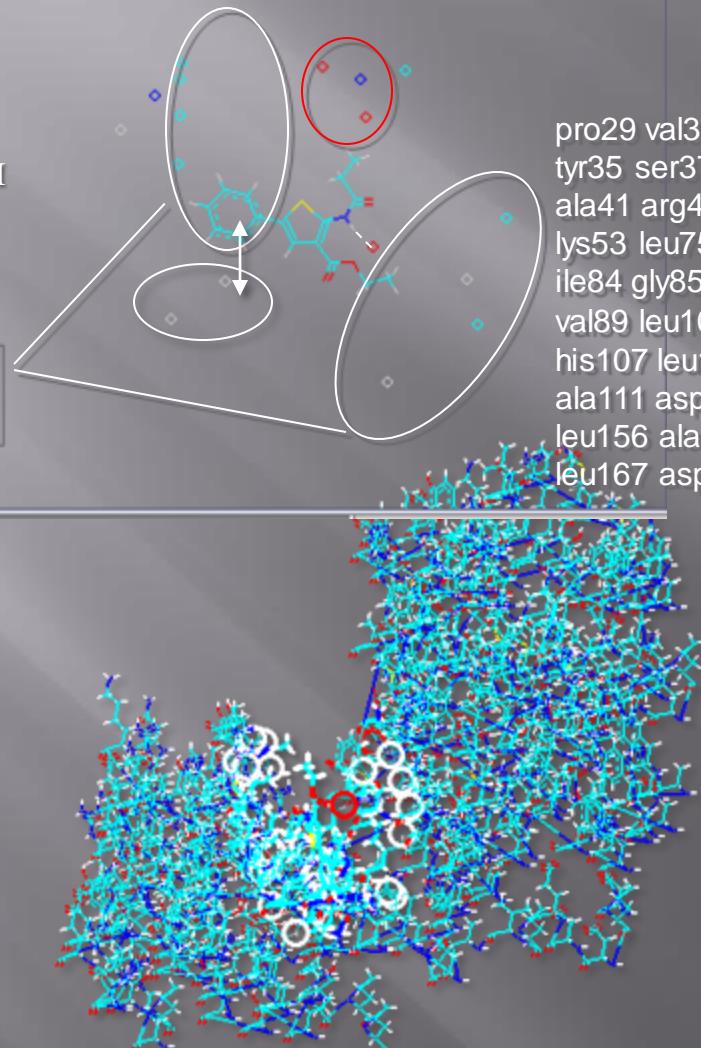
# Ингибиторы CDK4



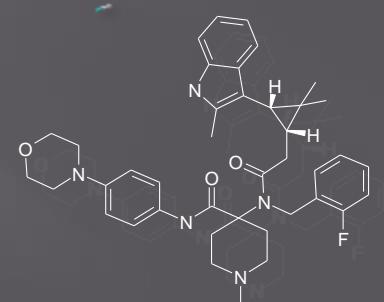
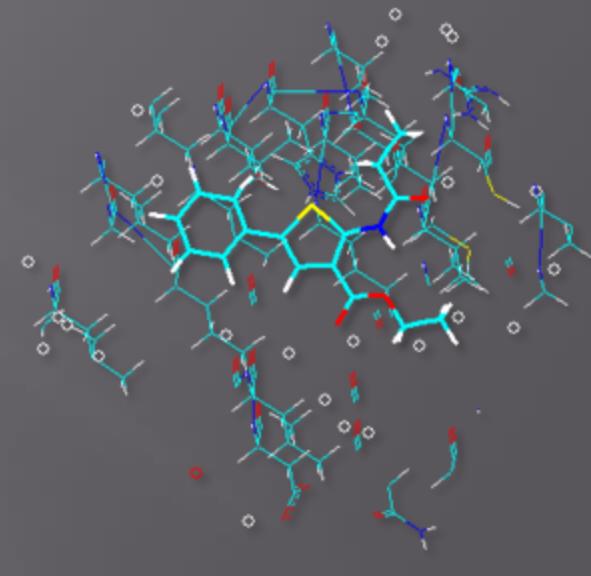
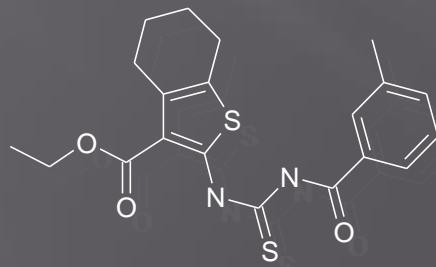
# Ингибиторы p38 МАР киназы

34 признано  
активными

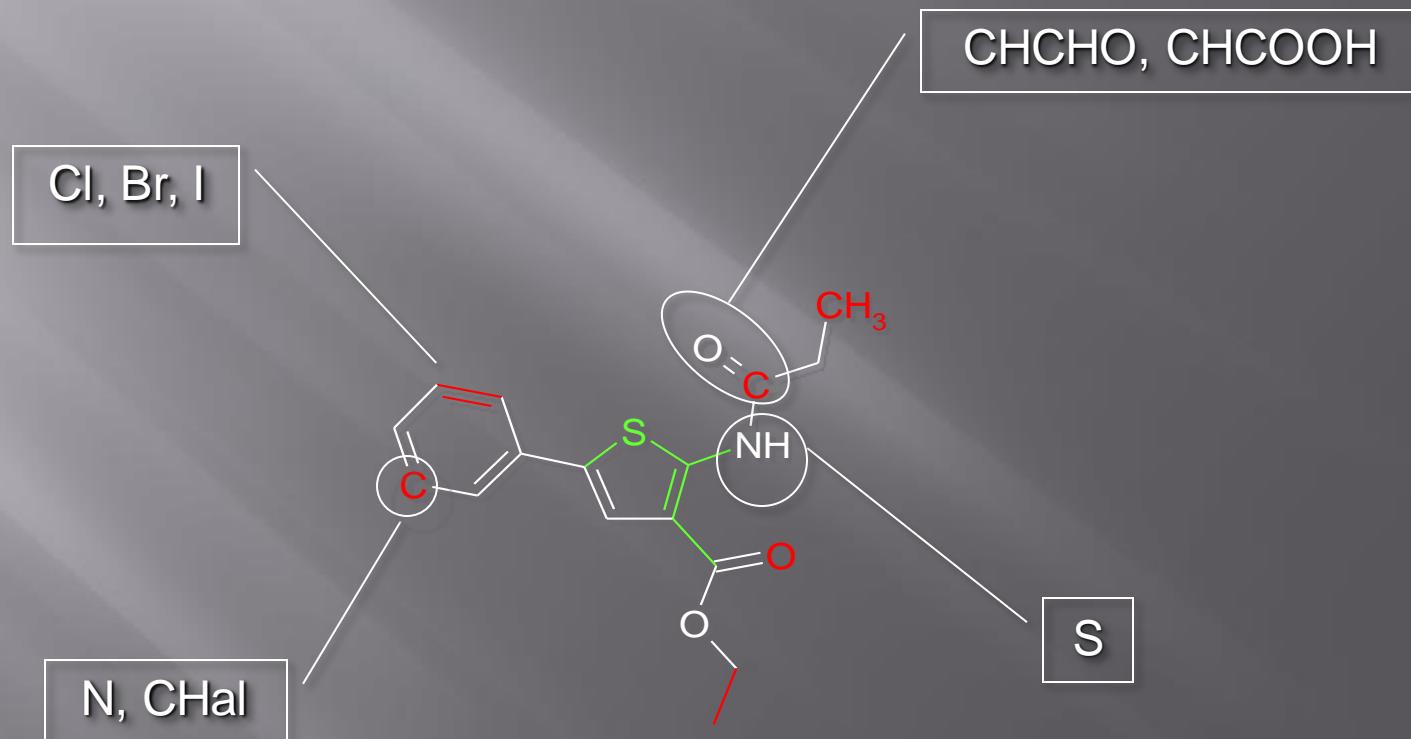
Липофильная  
часть



pro29 val30 gly31 ser32 gly33  
tyr35 ser37 val38 cys39 als40  
ala41 arg49 val50 ala51 val52  
lys53 leu75 lys76 met78 lys79  
ile84 gly85 leu86 leu87 asp88  
val89 leu104 val105 thr106  
his107 leu108 met109 gly110  
ala111 asp150 ser154 asn155  
leu156 ala157 lys165 ile166  
ile167 asp168 phe169

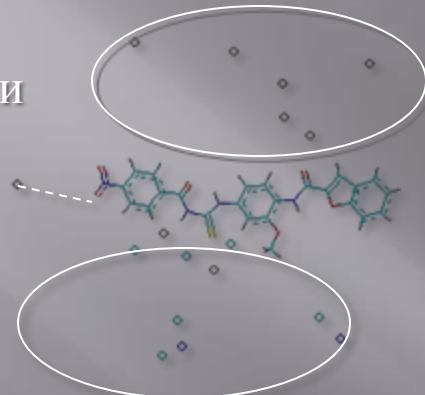


# Ингибиторы p38 МАР киназы



# Ингибиторы HIV1-протеазы

153 признаны  
активными



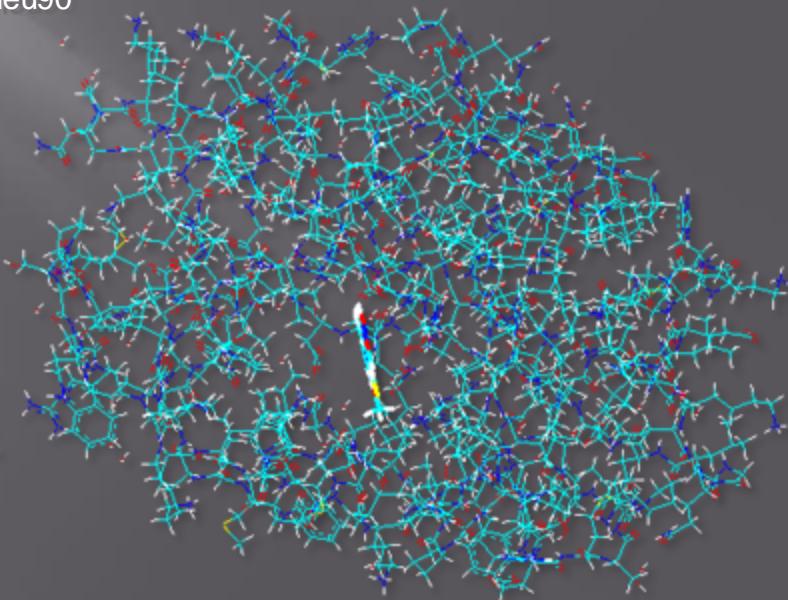
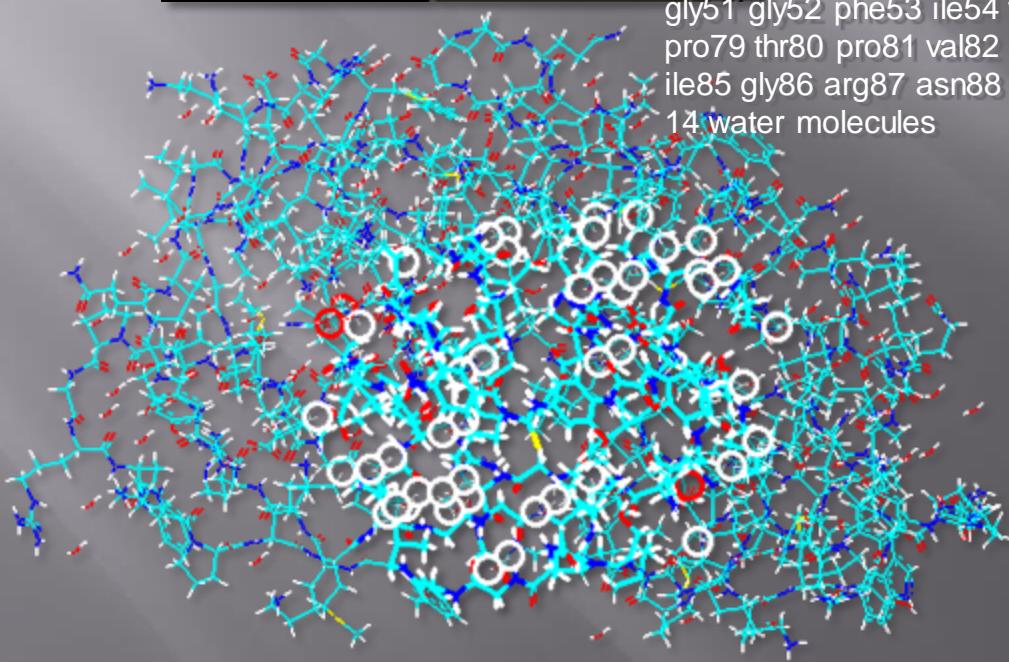
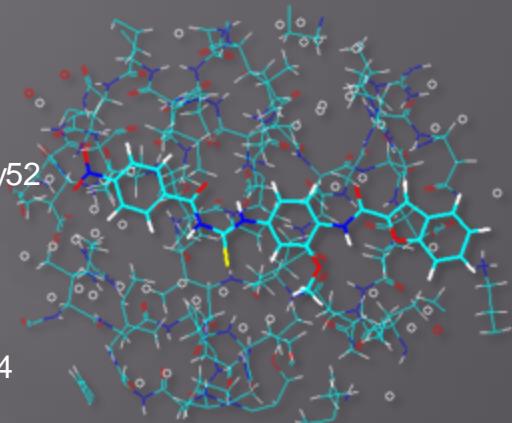
Chain A

arg8 leu23 asp25 thr26 gly27 ala28  
asp29 asp30 thr31 val32 leu33 lys45  
met46 ile47 gly48 gly49 ile50 gly51 gly52  
phe53 ile54 val56 leu76 thr80 pro81  
val82 ile84 ile85 gly86 arg87 asn88  
leu89 leu90

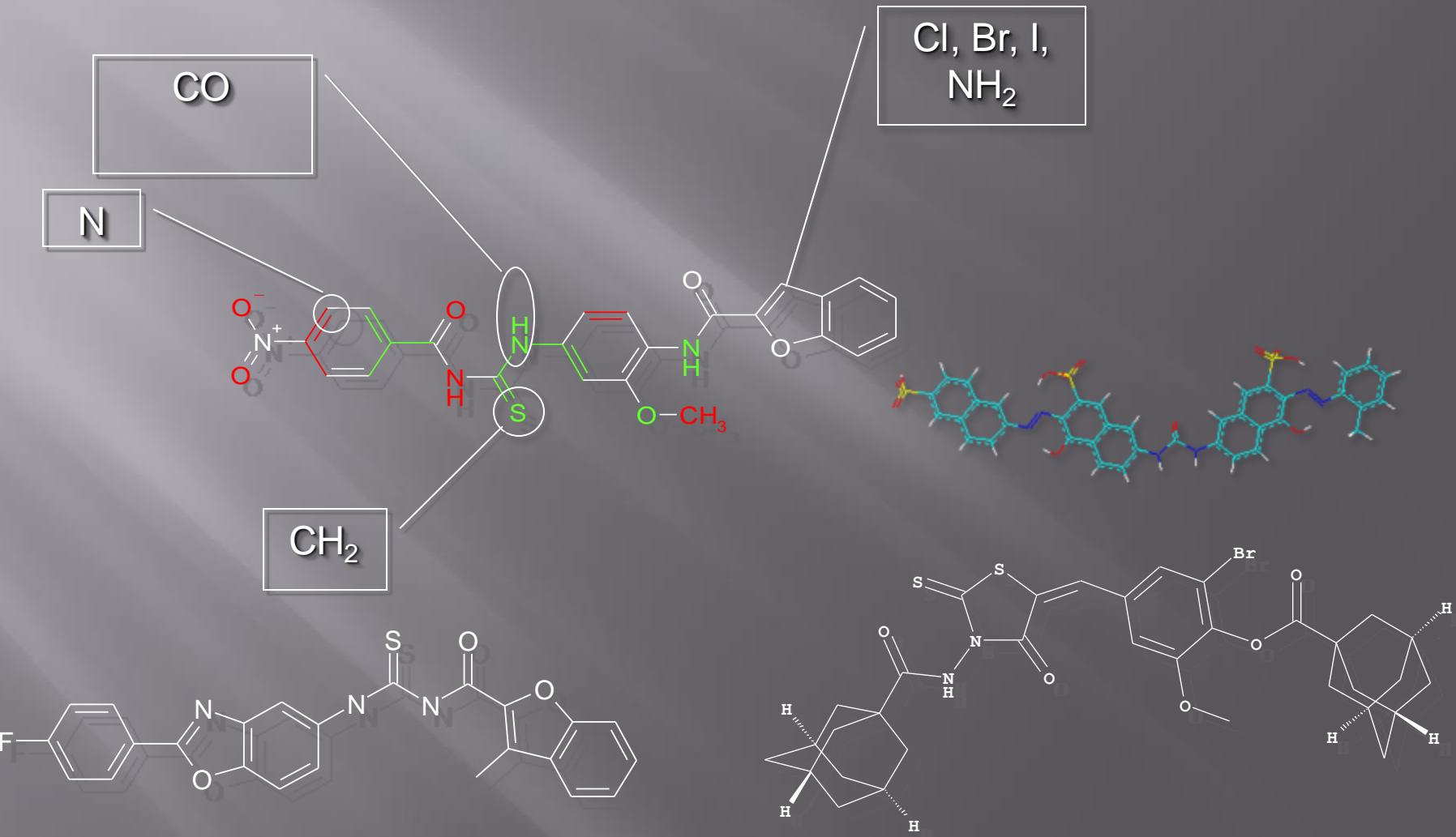
Chain B

leu5 arg8 pro9 leu10 ala22 leu23 leu24  
asp25 thr26 gly27 ala28 asp29 asp30  
thr31 val32 met46 ile47 gly48 gly49 ile50  
gly51 gly52 phe53 ile54 val56 leu76  
pro79 thr80 pro81 val82 asn83 ile84  
ile85 gly86 arg87 asn88 leu90

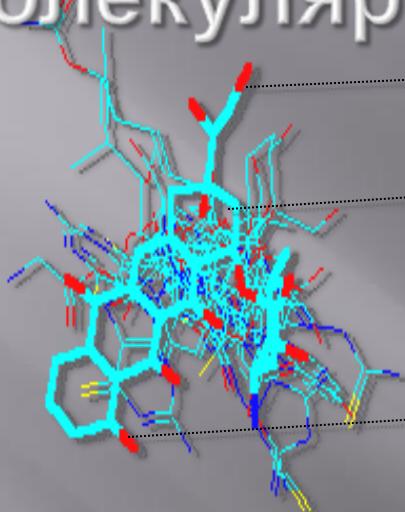
14 water molecules



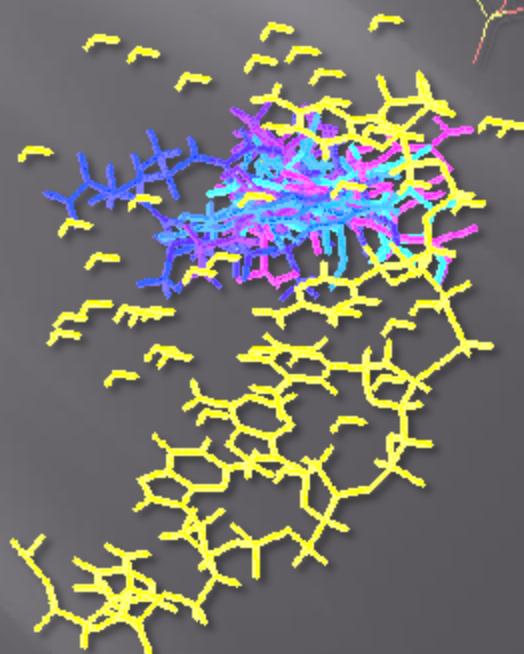
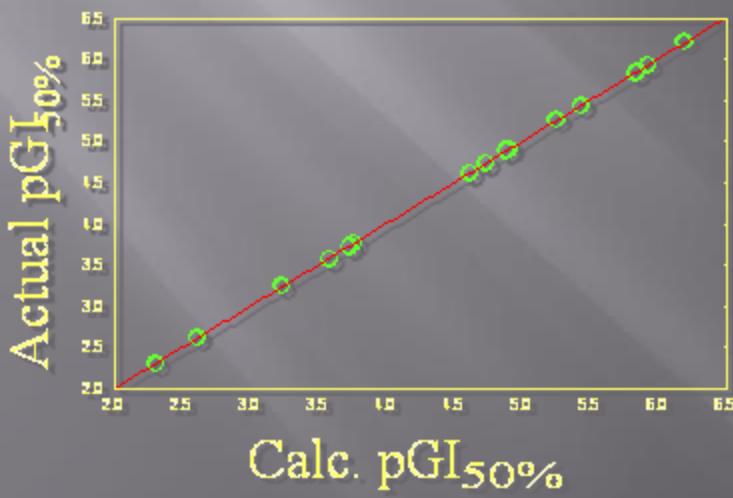
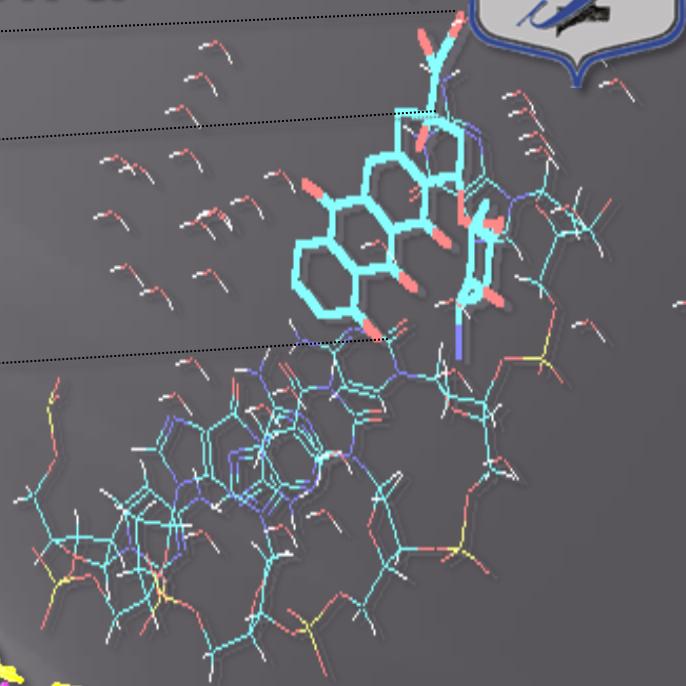
# Ингибиторы HIV1-протеазы



# Алгоритмы BiS, CiS, ConGO для молекулярного докинга



ДНК антиметаболиты





Сравнение данных PCA and и расчетных координат ингибиторов CDK2  
**RMS** – (root mean square) среднеквадратичное отклонение координат, Å  
 $\Delta r_{MAX}$  – максимальное отклонение координат, Å  
 $a_1, a_2, a_3$  – ориентационные отклонения по отношению к главным  
вращательным инвариантам

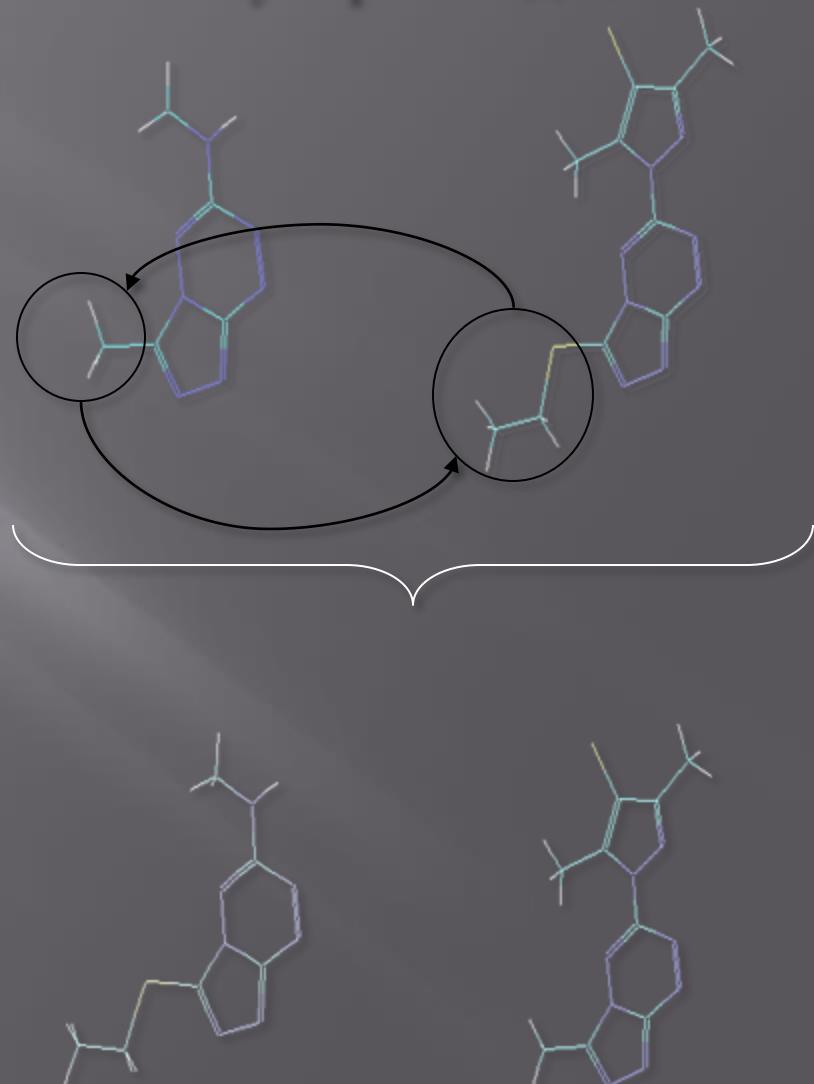
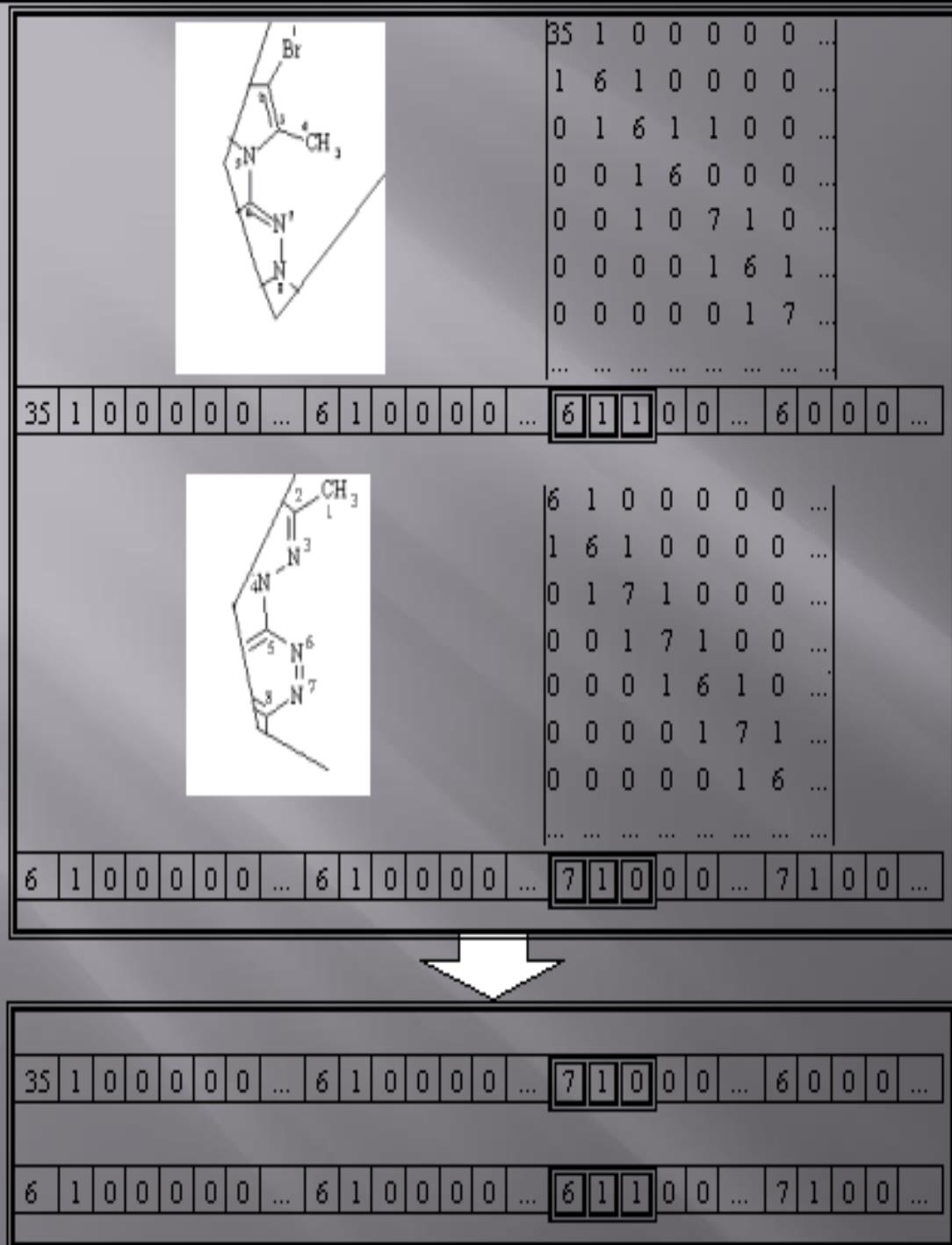
Molecule	RMS	$\Delta r_{MAX}$	$a_1, ^\circ$	$a_2, ^\circ$	$a_3, ^\circ$	RMS*	$\Delta r, \text{\AA}^*$
1agw	0.80	1.24	8.8	8.5	3.6	3.34	4.74
1aq1	0.66	1.23	1.9	4.6	4.3	1.13	1.96
1atp	0.70	1.36	2.6	2.7	0.9	9.25	16.51
1bkx	0.00	0.00	0.0	0.0	0.0	4.85	8.67
1ckp	0.26	0.56	7.0	8.9	4.9	5.81	9.93
2csn	0.74	1.09	7.2	7.2	1.2	0.00	0.00
1fgi	0.81	1.64	3.3	4.3	2.7	3.68	6.20
2hck	0.81	1.63	1.8	1.7	1.0	4.93	8.60
1ian	0.24	0.43	2.5	2.7	1.6	9.72	18.70
1stc	0.72	1.46	7.9	8.6	6.3	0.00	0.00
1ydr	0.97	1.81	1.6	5.7	5.9	0.43	0.72
1yds	0.68	1.11	1.1	4.3	4.5	0.00	0.00
1ydt	1.14	2.17	1.1	5.7	5.5	1.04	2.18

\* Cosgrove D.A., Bayada D.M., Johnson A.P., J. Comput.-Aided Mol. Design, 2000(14), 573 - 591.

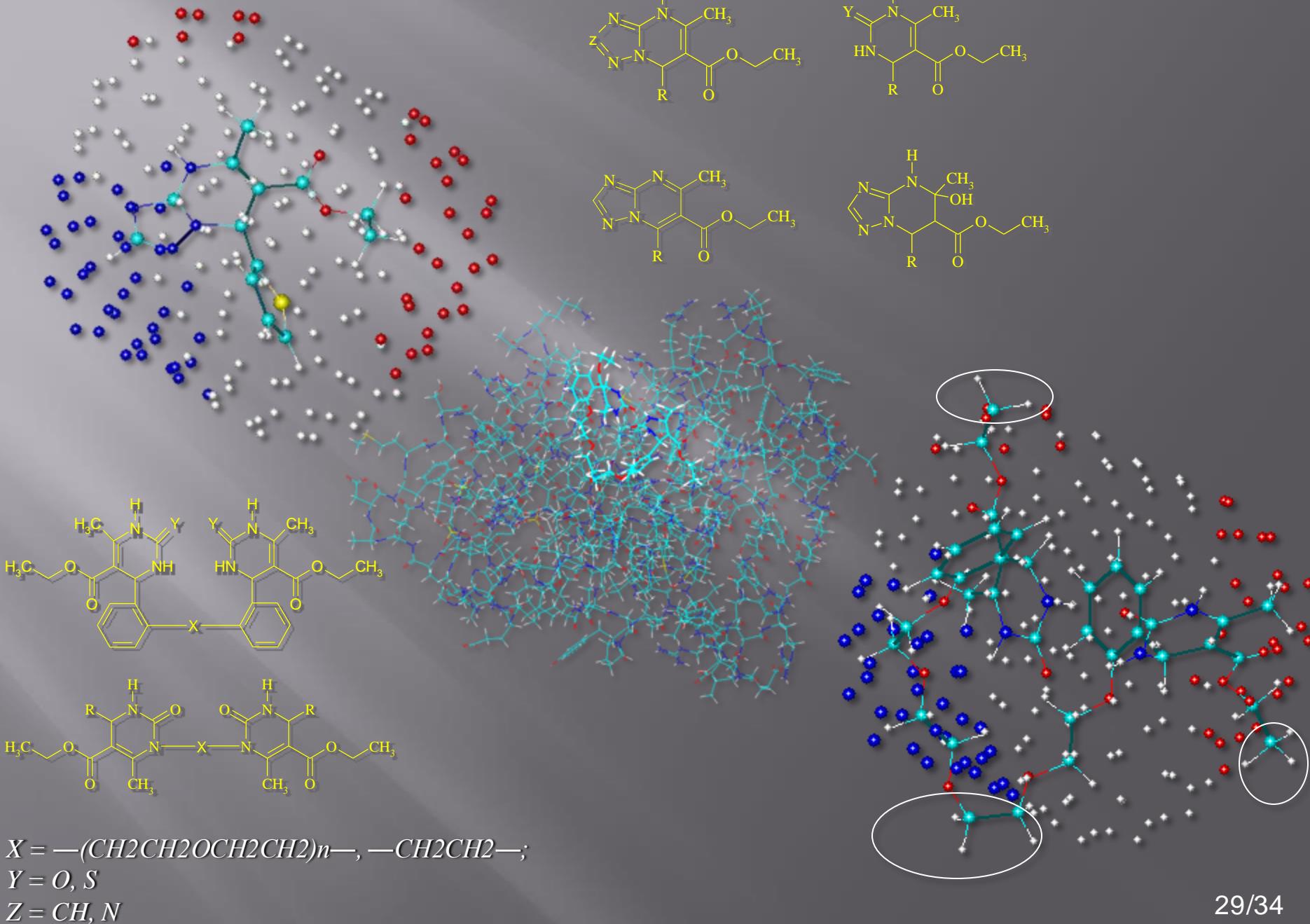
# Генетический алгоритм DesPot для молекулярного дизайна



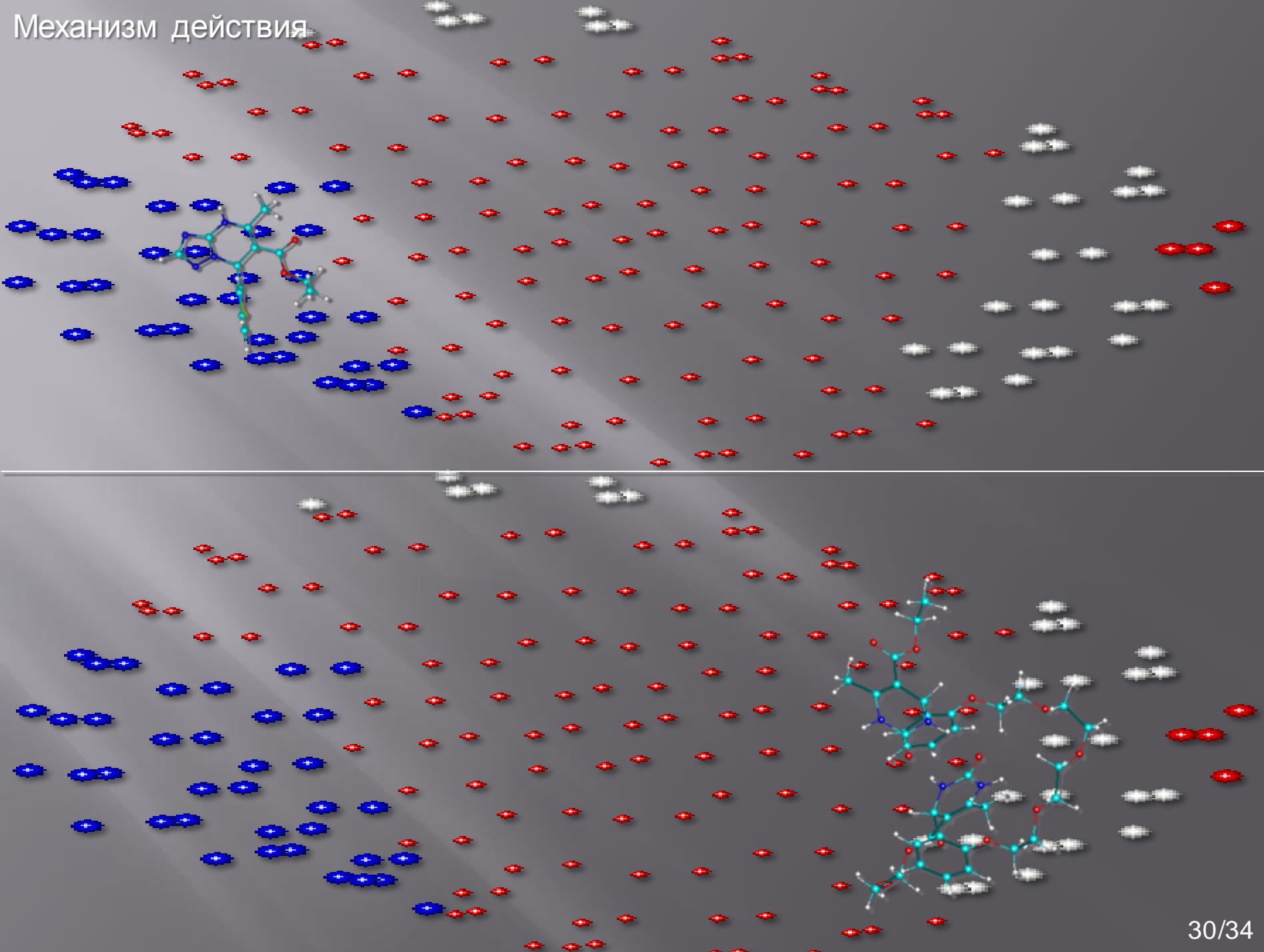
# Генетический алгоритм DesPot для молекулярного дизайна



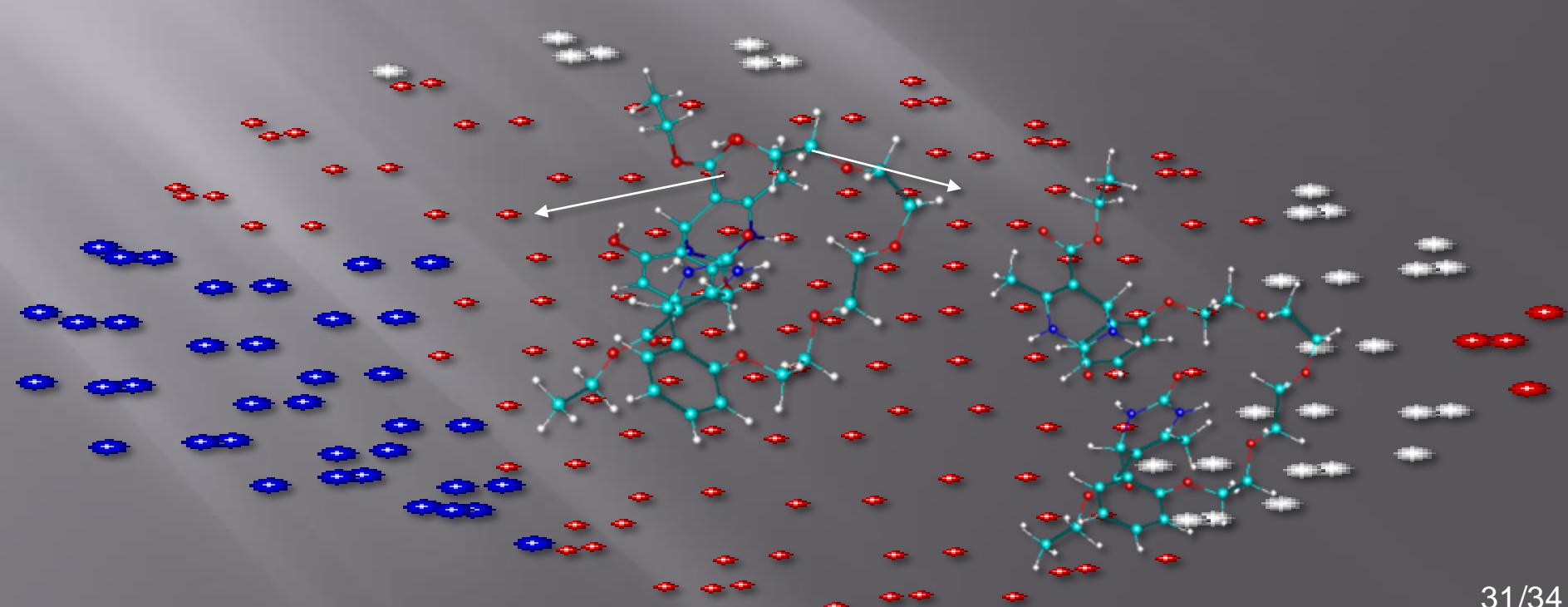
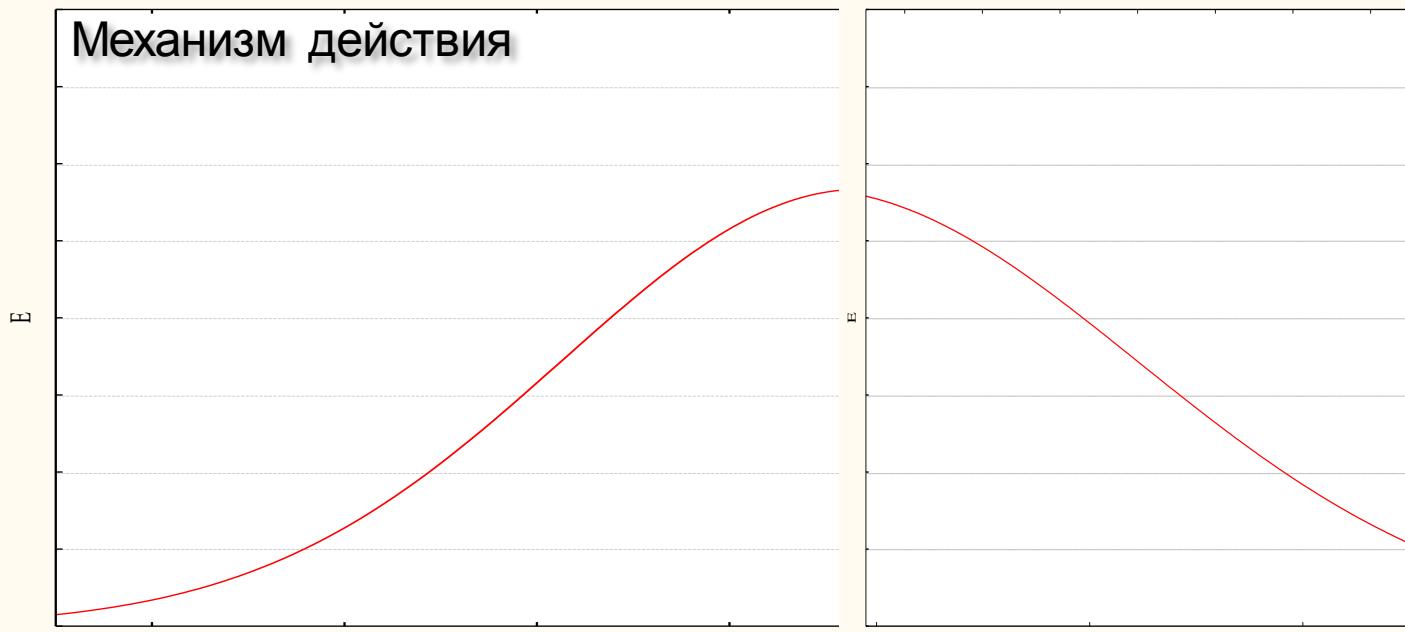
# Туберкулостатики



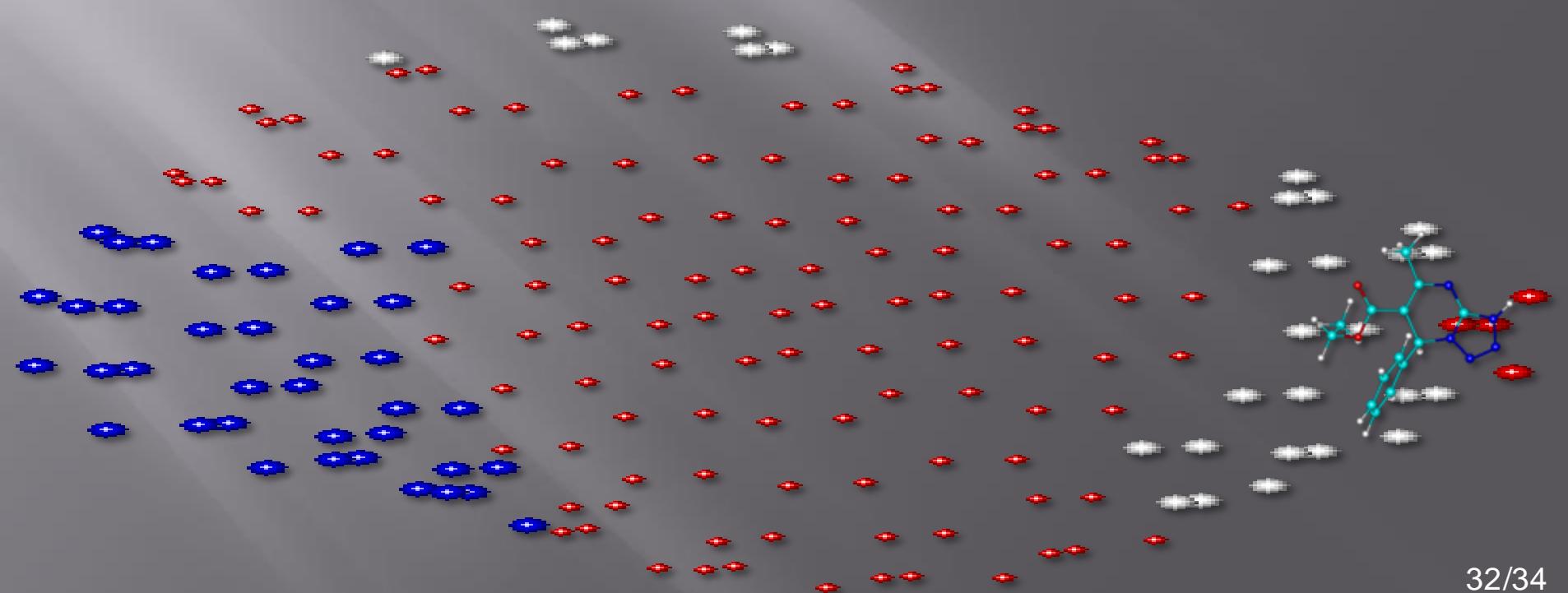
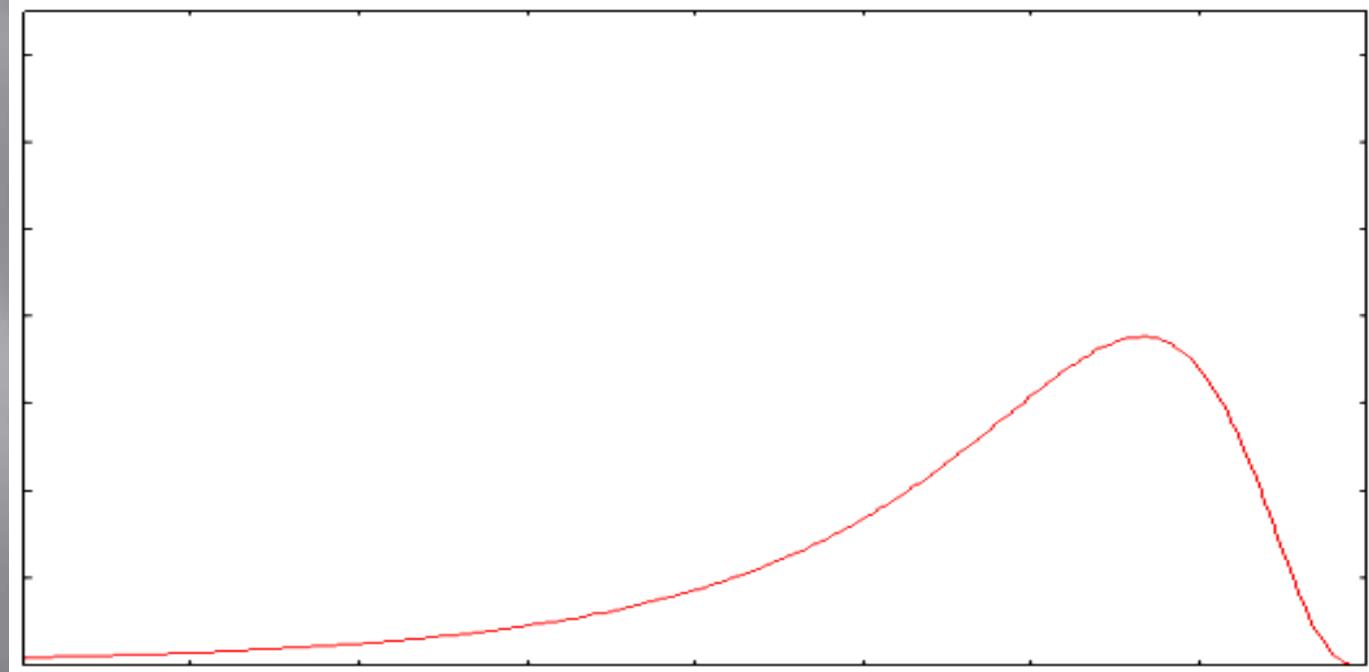
# Механизм действия



# Механизм действия



# Механизм действия



# Приложение для платформы Windows

BiS / MultiConformational

File Create BBI Create B2 Predict Activity Cancel

	BiS1.Pred	BiS.Pred	BiS2.Pred	Conf.Act.Pred.
novokain	-.52975	-.65300	-.75589	-.65300
hydukam	-1.20150	-1.00000	-.99544	-1.00000
meoxam	-1.			
mhloram	-1.			
mmeam	-.			
mmef3am	-1.			
obram	-.			

Threshold  2

novokain

Chemical structure of novokain showing a benzene ring with substituents. Atoms are numbered: 1 (top), 2 (top-left), 3 (top-right), 4 (top-right), 5 (left), 6 (top), 7 (bottom-left), 8 (bottom), 9 (bottom-right), 10 (right), 11 (bottom), 12 (bottom-right), 13 (bottom), 14 (right), 15 (right), 16 (right), 17 (right), 18 (bottom), 19 (right), 20 (right), 21 (right), 22 (right), 23 (right), 24 (right), 25 (right), 26 (right), 27 (right), 28 (right), 29 (right), 30 (right), 31 (right), 32 (right), 33 (right), 34 (right), 35 (right), 36 (right), 37 (bottom-right).

Done

Atom	Replace by	Increase	copy	Atom	Add	Increa:
NP		.00660		Near 37 H	C4	.
4 H	Incr = -4.728551E-02			Near 37 H	C3	.
	----	.04729		Near 37 H	C2	.
C0		.04729		Near 37 H	H	.
O2		.04729		Near 37 H	SI	.
O1		.04729		Near 37 H	IP	.
N3		.04729				



Мой Соавтор:  
Гришина Мария Александровна,  
Доктор химических наук,  
Главный научный сотрудник  
лаборатории компьютерного  
моделирования лекарственных средств,  
Высшая медико-биологическая школа ЮУрГУ

Благодарю за внимание!