About OMICS Group

- OMICS Group is an amalgamation of Open Access Publications and worldwide international science conferences and events. Established in the year 2007 with the sole aim of making the information on Sciences and technology 'Open Access', OMICS Group publishes 500 online open access scholarly journals in all aspects of Science, Engineering, Management and Technology journals. OMICS Group has been instrumental in taking the knowledge on Science & technology to the doorsteps of ordinary men and women. Research Scholars, Students, Libraries, Educational Institutions, Research centers and the industry are main stakeholders that benefitted greatly from this knowledge dissemination. OMICS Group also organizes 500 International conferences annually across the globe, where knowledge transfer takes place through debates, round table discussions, poster presentations, workshops, symposia and exhibitions.

OMICS International Conferences

OMICS International is a pioneer and leading science event organizer, which publishes around 500 open access journals and conducts over 500 Medical, Clinical, Engineering, Life Sciences, Pharma scientific conferences all over the globe annually with the support of more than 1000 scientific associations and 30,000 editorial board members and 3.5 million followers to its credit.

OMICS Group has organized 500 conferences, workshops and national symposiums across the major cities including San Francisco, Las Vegas, San Antonio, Omaha, Orlando, Raleigh, Santa Clara, Chicago, Philadelphia, Baltimore, United Kingdom, Valencia, Dubai, Beijing, Hyderabad, Bengaluru and Mumbai.

Protein-ligand low energy minima pose analysis: docking target functions evaluation with the FLM program

Alexey Sulimov, Ekaterina Katkova, Vladimir Sulimov,

DIM Ψ NTA

Moscow, Russia

Lomonosov
Moscow State University
Research Computer Center

5 components of the rational and smart new drug development

1. ***To be needed bio-target*** **–** medics, pharmaceutists, biologists, biochemists, molecular biologists …
2. Availability of the bio-target 3D structure
3. High accuracy of the protein-ligand binding energy calculations. Error ΔG_{bind} **< 1 kcal/mol**
4. New compounds synthesis symplicity
5. Availability of reliable test systems for experimental measurements of inhibitors activity

Application of molecular modeling should improved effectiveness of rational drug design

- Decrease time of new inhibitors design
- Increase diversity of new inhibitors
- Decrease the number of new compounds syntheses
- Decrease time of the new drug development
- Decrease expenses of R&D

The protein-ligand binding free energy

$\Delta G_{bind} = \Delta H - T\Delta S$

ΔG_{bind} – *the protein-ligand binding free energy*

ΔH – *binding enthalpy,*

$-T\Delta S$ – *binding entropy*

$\Delta G = kT \ln(K_i)$, K_i – *inhibition constant*

K_i – *measured in experiment*

Main molecular modeling tools for calculation of the ΔG_{bind} :

- **Docking** – ligand positioning in the target protein, estimation of ΔG_{bind}. Docking is the *most popular method* for Drug Design. There are many docking programs: AutoDock, DOCK, ICM, GOLD, FlexX, FlexE, BUDE, Glide, Surflex-Dock, Lead-Finder, SOL, TTDock, *etc.*

- **Molecular Dynamics** – calculation of trajectories of all protein-ligand atoms and all water molecules; ΔG_{bind} calculation – energy averaging along the trajectories.

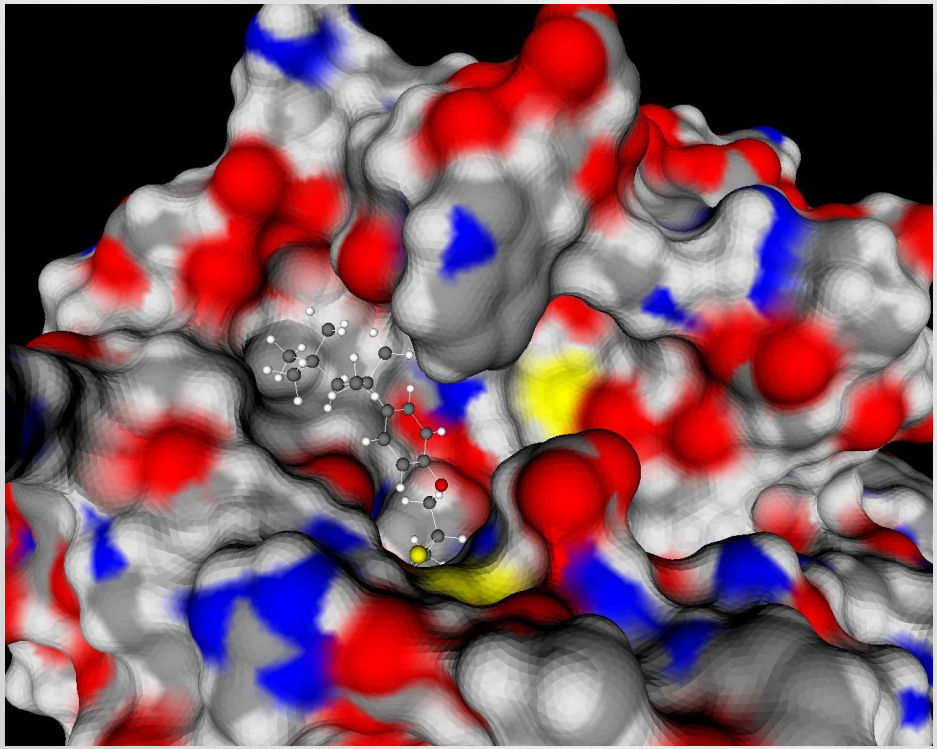
Docking versus MD

- Docking is the *most popular method* for Drug Design, Quick, Virtual screening of many thousands of ligands. Score – estimation of ΔG_{bind} – ACCURACY IS BAD.

- *Theoretically* MD is the *most precise method* of ΔG_{bind} calculation, **too slow** for virtual screening, many tricks in calculations – alchemy. Accuracy is not enough for an arbitrary protein-ligand complex.

Docking Paradigm:

The ligand position in the target protein active site corresponds to ***the global minimum*** of the protein-ligand energy function



Docking – finding the ***global minimum*** of the **target energy function**

Docking problems

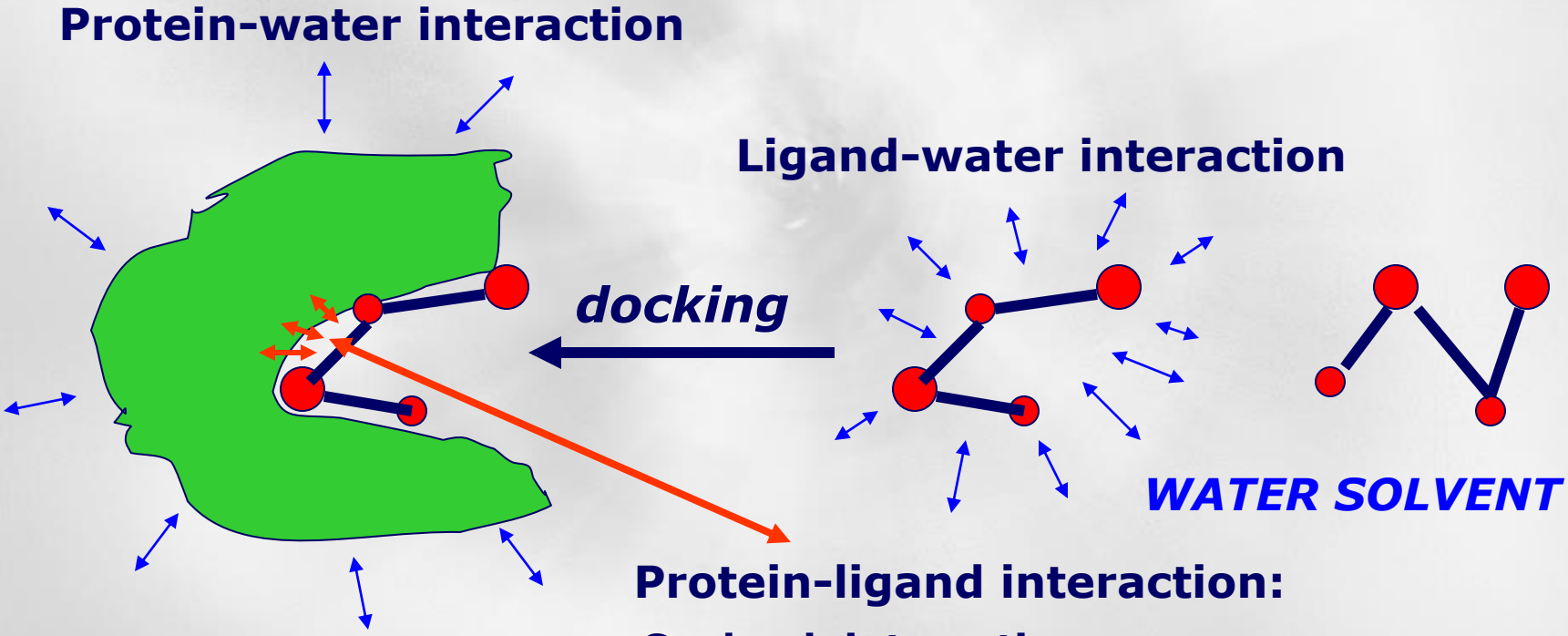
- Positioning accuracy is not high enough: there are many examples of the native ligand docked with RMSD > 2 Å
- Accuracy of ΔG_{bind} calculation is not high enough
- $T\Delta S$ ~ *(the number of ligand torsions)* – *bad approximation*
- Fitting parameters are used in many docking programs – impossible to estimate ***accuracy a priory***
- It is impossible to optimize lead compound: to distinguish between weak, medium and strong inhibitors on the base of docking results
- Accuracy of ΔG_{bind} calculations must be better than **1 kcal/mol**

Key Programs for Drug Design:

Docking – the correct ligand positioning in the active site of the target-protein

Scoring – the correct estimation of the protein-ligand interaction energy

High accuracy ~ 1 kcal/mol ~ 0.05 eV



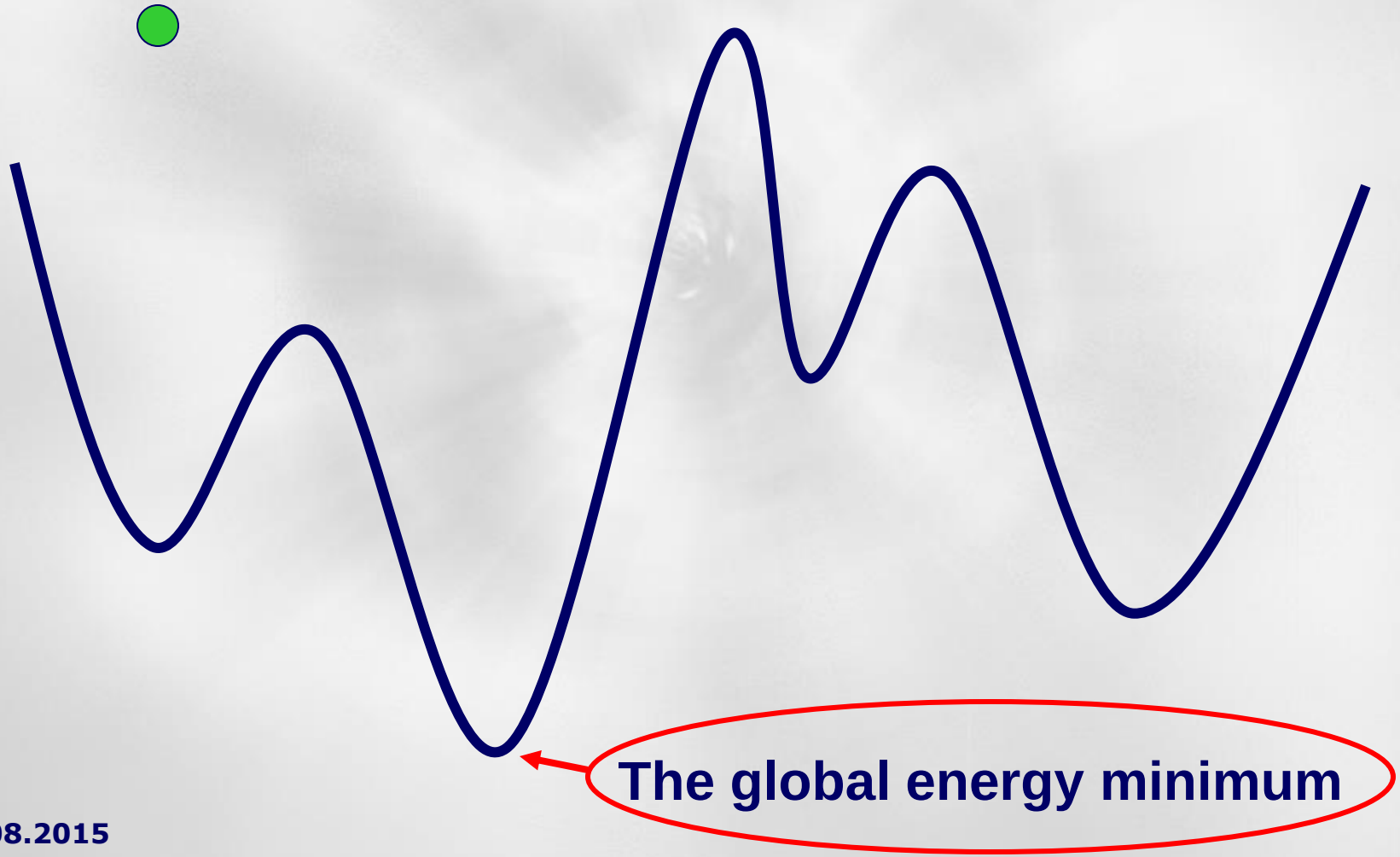
Protein-ligand interaction:

- **Coulomb interactions**
- **Van der Waals interactions**

Is it possible to use Docking for accurate ΔG_{bind} calculations? Goals of our research:

- To find the global minimum (minimum minimorom) for protein-ligand energy target function
- Different target functions
- No fitting parameters
- Detailed investigation of protein-ligand low energy minima
- Employment of Supercomputers for docking
- To find causes of low docking accuracy
  - Accuracy of ligand positioning
  - Accuracy of binding energy calculations

The ligand movement in the protein active site



Docking

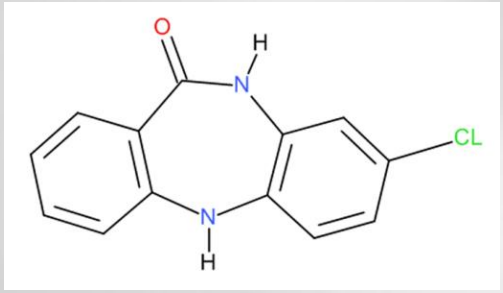
- To find the global protein-ligand energy minimum
- To find a spectrum of protein-ligand low energy minima
- The global minimum and nearest in energy local minima give main contribution to the binding energy

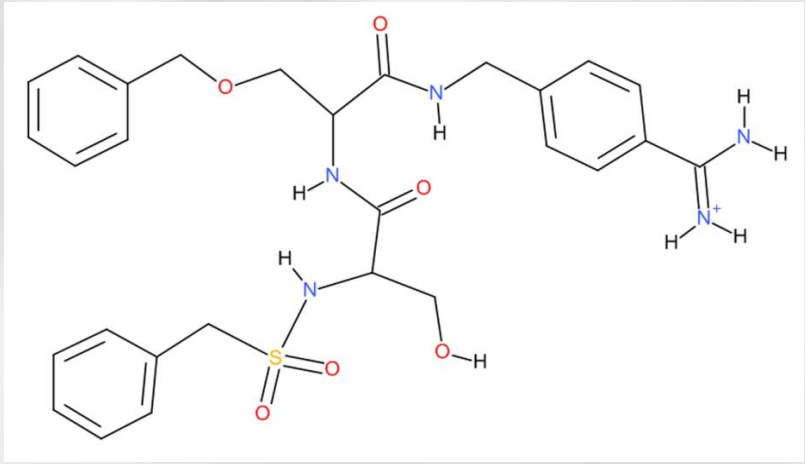
FLM docking program – **F**ind **L**ocal **M**inima

- FLM does not use any preliminary calculated energy grid
- Rigid protein – *for the present investigation*
- Local energy optimization in respect to all ligand atoms from a random initial position
- Vacuum or implicit solvent models
- Force Field MMFF94 - *for the present investigation*
- Parallel multi-processors calculations: **8191 cores** several hours of the Lomonosov supercomputer
- Search for the low energy minima spectrum (1024 lowest energy different minima)
- Monte Carlo exhaustive minima search

Investigation of docking accuracy

- Positioning accuracy,
- The *docking paradigma* check
- ΔG_{bind} calculation accuracy
- A set of 16 protein-ligand complexes from Protein Data Bank, RMSD ≤ 2 Å, different ligand size and flexibility with known inhibition constants K_i





74 atoms, 19 torsions

Docking accuracy depends on:

- Force field – inter- and intra-molecular interactions
- Solvent (water) model
- Models of the target protein and ligands
- The free energy calculation method and approximations
- Algorithms of calculations
- Computer resources for docking of one ligand

Five target energy functions

- MMFF94 local optimization in vacuum
- MMFF94 +solvent in the PCM (Polarized Continuum Model) model
- MMFF94 + solvent in the Surface-GB model
- PM7 (MOPAC) local optimization in vacuum
- PM7(MOPAC) + solvent in COSMO model

PM7 – new quantum-chemical semiempirical method:

•Improved dispersion interactions

•Improved Hydrogen Bonds description

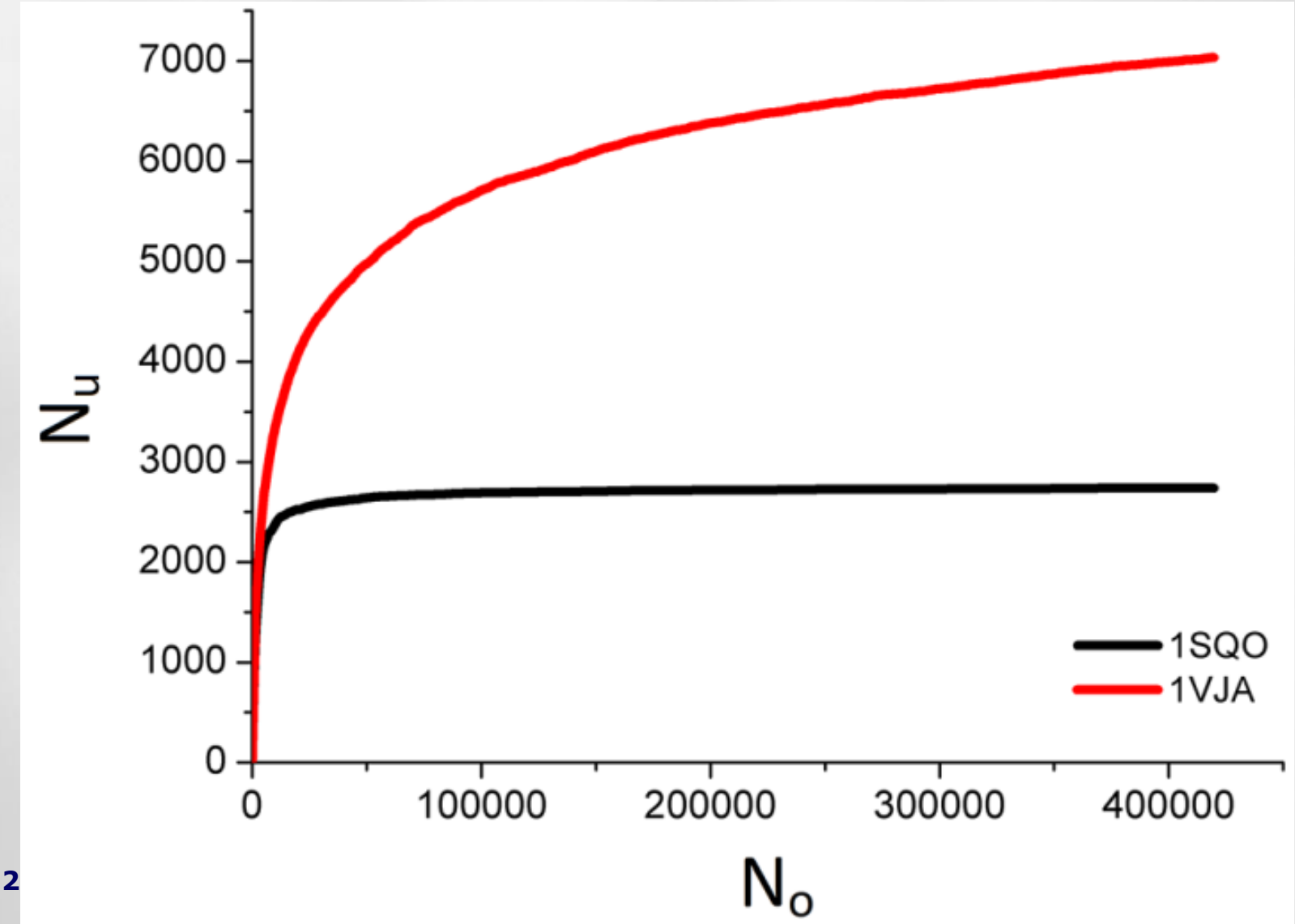
Minima sets found by FLM {1}

- {1}MMFF94 – low energy minima selection in vacuum
- {1}MMFF94+PCM – minima energies recalculation with PCM
- {1}PM7 – local optimization from these minima with MOPAC in vacuum
- {1}PM7+COSMO – minima energies recalculation with c COSMO (MOPAC)

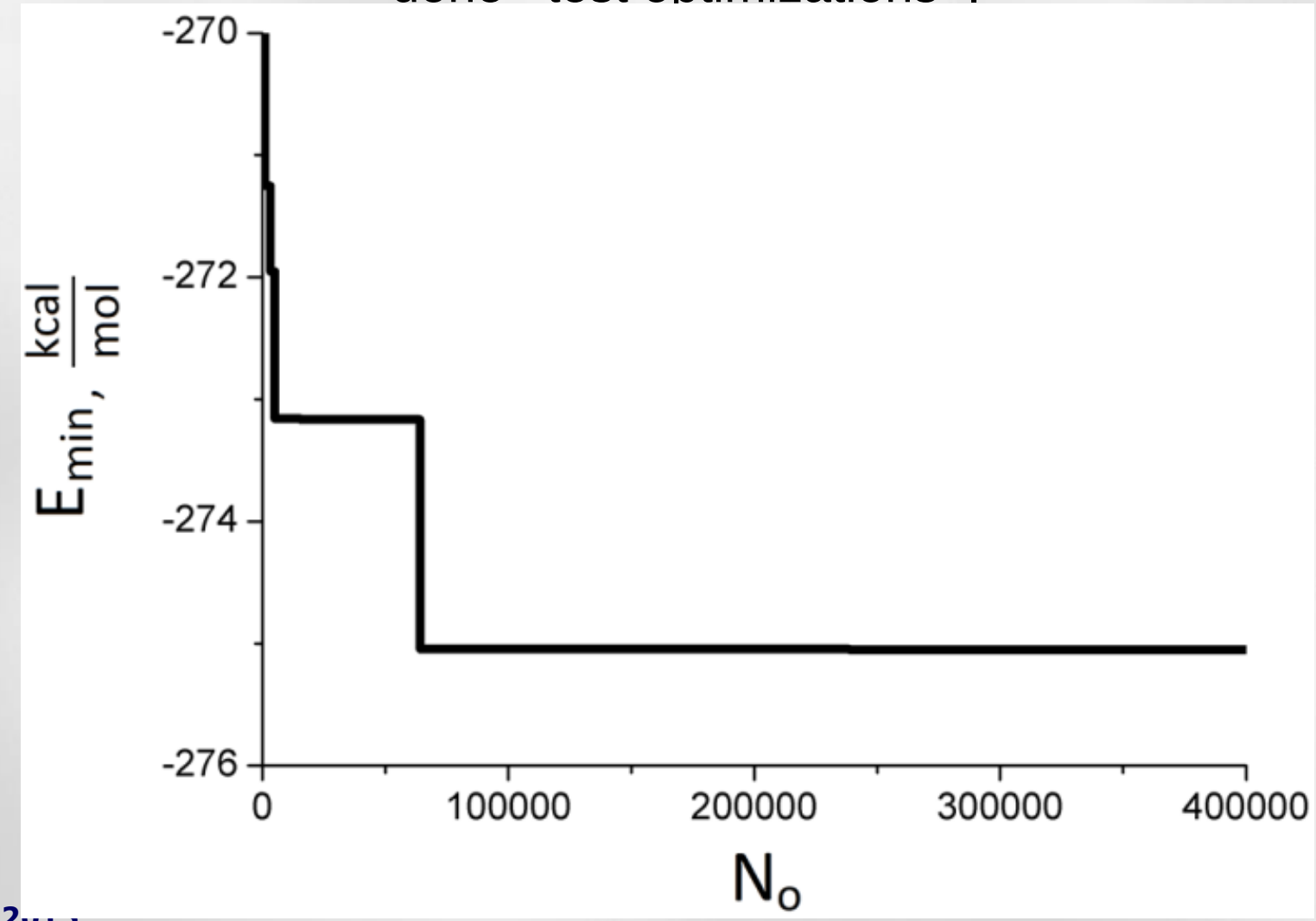
Minima sets found by FLM {2}

- {2}MMFF94+PCM – low energy minima selection with PCM solvent model
- {2}MMFF94 – minima energies recalculation in vacuum
- {2}MMFF94+S-GB - minima energies recalculation with Surface-GB solvent model

The local minima set updates number (*Nu*) depending on done «test optimizations» number (*No*) for 1SQO (black line) and 1VJA (red line) protein-ligand complexes



Potential energy of the global minimum (*Emin*) depending on done «test optimizations» number (*No*) for 1VJA protein-ligand complex. This energy was updated for the last time after 64205 done «test optimizations».



Minima indexes

- **Each energy minimum has an integer index corresponding to its position in the minima list sorted by their energies in ascending order. The lowest energy minimum has index equal to 1.**

- **IN** – (**I**ndex of **N**ative) is the index of the locally optimized native ligand.

- **INN** – (**I**ndex of **N**ear **N**ative) is the index of the minimum having RMSD from the non-optimized native ligand position less than 2 Å

IN / INN indexes for different target functions

PDB ID	{1}MMFF	{1}MMFF +PCM	{1}PM7	{1}PM7+ COSMO	{2}MMFF	{2}MMFF +PCM	{2}MMFF +SGB
4FT0	36 / 20	8 / 7	37 / 12	1 / 1	180 / 99	164 / 159	8 / 6
4FT9	45 / 28	1 / 1	25 / 6	1 / 1	194 / 125	3 / 1	1 / 1
4FSW	5 / 5	6 / 6	40 / 40	12 / 13	110 / 102	134 / 140	21 / 3
4FTA	*inf* / *inf*	4 / *inf*	379 / *inf*	1 / *inf*	*inf* / *inf*	186 / 187	97 / 97
4FV5	204 / 131	3 / 3	253 / 194	2 / 1	186 / 134	6 / 3	5 / 5
4FV6	*inf* / *inf*	1 / *inf*	49 / *inf*	1 / *inf*	86 / 289	3 / 68	1 / 24
1DWC	*inf* / 670	245 / 25	689 / 661	158 / 141	245 / 114	250 / 35	107 / 8
1TOM	*inf* / *inf*	13 / *inf*	*inf* / *inf*	1 / *inf*	*inf* / *inf*	13 / 4	7 / 1
1C5Y	1 / 1	2 / 1	7 / 1	2 / 1	1 / 1	2 / 1	1 / 1
1F5L	1 / 1	1 / 1	43 / 16	69 / 30	1 / 1	10 / 1	1 / 1
1O3P	20 / 18	21 / 1	5 / 1	3 / 1	69 / 62	274 / 1	130 / 2
1SQO	1 / 1	2 / 1	1 / 1	1 / 1	1 / 1	54 / 1	5 / 1
1VJ9	46 / 1	86 / 51	32 / 1	26 / 8	6 / 1	11 / 18	10 / 14
1VJA	42 / 3	7 / 1	7 / 4	6 / 4	4 / 49	1 / 2	1 / 1
2P94	36 / 2	19 / 1	23 / 6	7 / 1	22 / 1	35 / 1	21 / 1
3CEN	96 / 1	18 / 1	13 / 1	3 / 1	90 / 1	35 / 1	13 / 1

"*inf*" for IN means that all (1024) found by FLM low-energy minima have energies below the energy of the optimized native ligand

"*inf*" for INN means that all found by FLM low-energy minima have RMSD from the native position above 2 Å

Ligand positioning with *MMFF94 in vacuo* target function

- The docking paradigm *is confirmed* only for *3 complexes out of 16 (6 complexes out of 30) – 20% complexes.* For these complexes ***IN / INN = 1 / 1***:
  - *The locally optimized native ligand pose has lowest energy among energies of all minima found by FLM*
  - *The minimum with lowest energy (the global minimum of the target function) found by the FLM program is close to the ligand native pose*

Protein	PDBID	$E_{global\ min.} - E_{opt.nat.}$, kcal/mol
CHK1	4FT0	-29.5
	4FT9	-18.3
	4FSW	-15.4
	4FTA	-91.2
ERK2	4FV5	-36.6
	4FV6	-19.3
thrombin	1DWC	-41.5
	1TOM	-62.3
urokinase	1C5Y	0.0
	1F5L	0.0
	1O3P	-5.2
	1SQO	-0.1
	1VJ9	-6.6
	1VJA	-4.7
factor Xa	2P94	-9.5
	3CEN	-14.0

Difference between the found global minimum of potential energy $E_{global\ min}$ and the energy $E_{opt.nat.}$ of locally optimized native ligand

1b9v	212
1br5	11
1c5y	2
1dwc	11
1efy	30
1f5l	1
1hpv	4
1i7z	1
1j01	4
1k1j	9
1lqd	11
1lzg	3
1mq6	19
1o3p	12
1ppc	12
1sqo	1
1tom	22
1vj9	17
1vja	46
2p94	7
2pax	2
3cen	7
3kiv	7
3pax	3
4fsw	1
4ft0	1
4ft9	5
4fta	2
4fv5	2
4fv6	4

The number of local minima within 5 kcal/mol near the global minimum

The docking target function: MMFF94 in vacuum

Conclusions: ligand positioning

- The energy target function with implicit solvent model is better than the energy target function in vacuum: indexes **IN** and **INN** become lower

- PM7 with solvent (COSMO) is better than MMFF94 with solvent (PCM)

- Docking program FLM is the tool for testing valuability of energy function for docking application
- Docking program FLM is the tool for testing effectiveness of the global energy minimum search

To improve docking accuracy:

- To find better force field
- To use implicit water models
- Force fields should be substituted by quantum chemistry – semiempirical method PM7

LET US MEET AGAIN..

We welcome you to our future conferences of OMICS International

2nd International Conference and Expo

on

Drug Discovery & Designing

On

October -31 November-02, 2016 at **Istanbul, Turkey**

http://drug-discovery.pharmaceuticalconferences.com/

- **…Surely every medicine is an innovation; and he that will not apply new remedies, must expect new evils…**

Francis Bacon

(1561-1626)

OF INNOVATIONS

Thank you!