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.

Is it possible to improve considerably the accuracy of docking programs?

Vladimir Sulimov



DIMONTA

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 *in vitro*

The protein-ligand binding free energy

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

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

ΔH – binding enthalpy,

$-T\Delta S$ – binding entropy

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

K_i – measured in experiment

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

Main molecular modeling tools for calculation of 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, Surflex-Dock, 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.

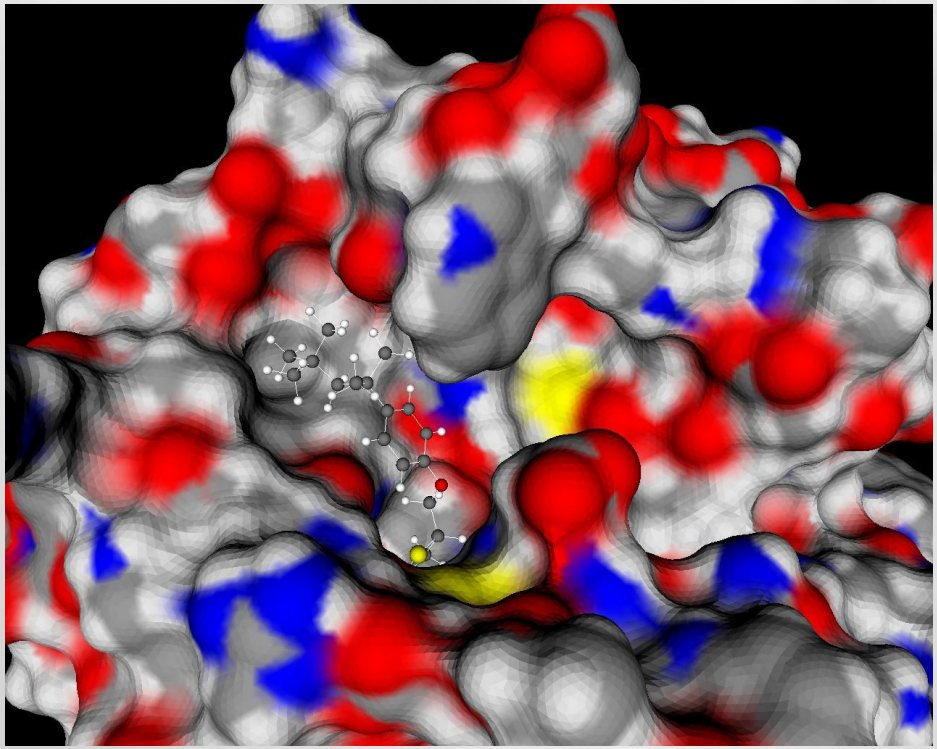
Docking versus MD

- 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 However it is **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

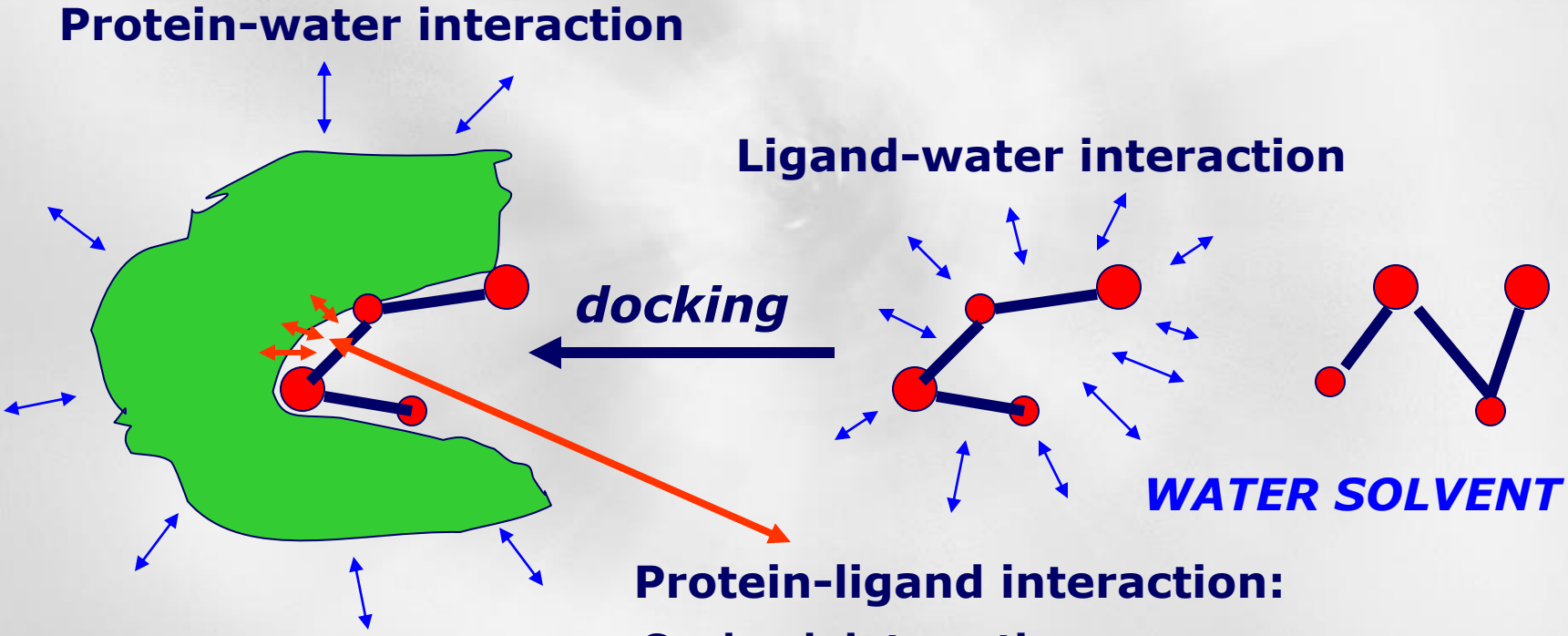
- Positioning accuracy is not high enough: there are many examples of the native ligand docked with RMSD > 2 Å
- Accuracy of ΔG_{bind} calculation is not high enough
- $T\Delta S$ ~ *(the number of ligand torsions)* – *bad approximation*
- Fitting parameters are used in many docking programs – impossible to estimate ***docking 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?

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

FLM 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

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

- Taking into account the ligand deformation energy
- ΔG_{bind} : the configuration integral is calculated in the frame of the *Multi-Well approximation*: the energy surface is approximated by a set of harmonic wells – a set of harmonic oscillators
- Atomic vibrations are taken into account
- Accurate calculation of the binding enthalpy ΔH and the entropy – $T\Delta S$ both in the frame of the same model: $\Delta G_{bind} = \Delta H - T\Delta S$

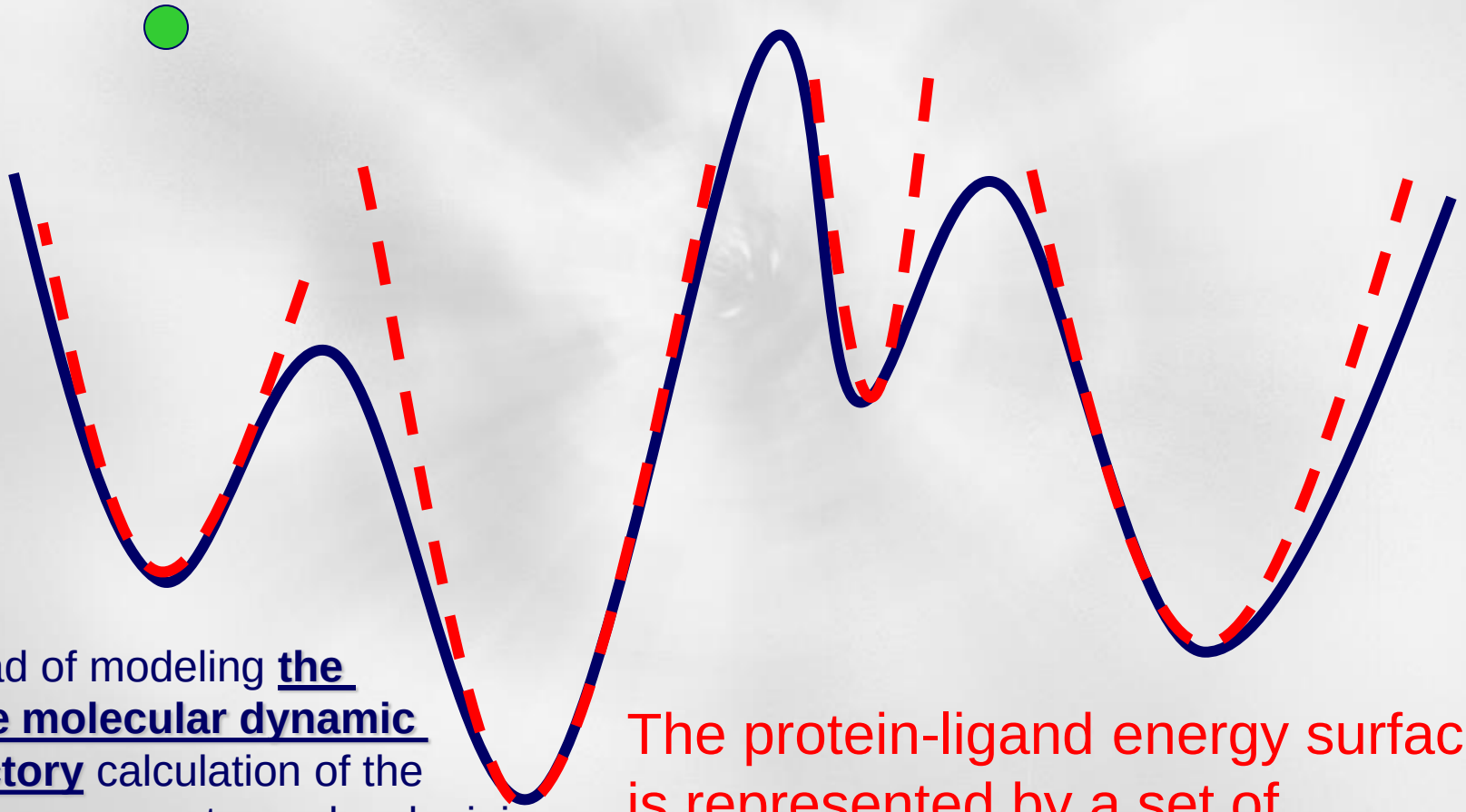
Gibbs free energy

$$G = -kT \ln Z$$

The ligand movement in the protein active site

$$Z = \frac{1}{(2\pi\hbar)^{3n}} \int e^{-(U+W)/kT} \, dx_1 \ldots dx_{3n} dp_1 \ldots dp_{3n}$$

Configuration integral



Instead of modeling **the whole molecular dynamic trajectory** calculation of the ligand movement near local minima

The protein-ligand energy surface is represented by a set of independent harmonic wells

Free energy of a molecular system in the multi-well approximation

$$G = -kT \ln Z = -kT \ln \left(\sum_i \exp\left(-\frac{E_0^i}{kT} \right) Z_\nu^i Z_t^i Z_r^i \right)$$

i – index of the wells

ν – vibration, t – translation, r- rotation

If the only global minimum is taken into account:

$$G = E_0^1 + G_\nu + G_t + G_r$$

Configuration integrals

$$Z_v^i = \prod_{l=1}^{3n} \frac{e^{-\hbar\omega_l^i / 2kT}}{1 - e^{-\hbar\omega_l^i / kT}}$$

vibrations

$$Z_t^i = e * \frac{(2\pi MkT)^{\frac{3}{2}}}{h^3 \rho}$$

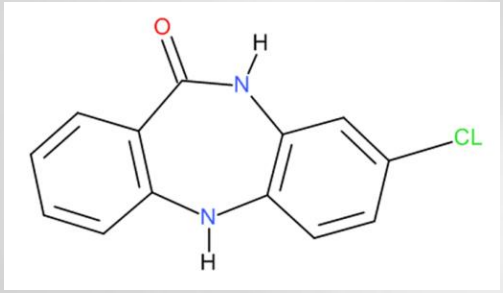
Translation as a whole

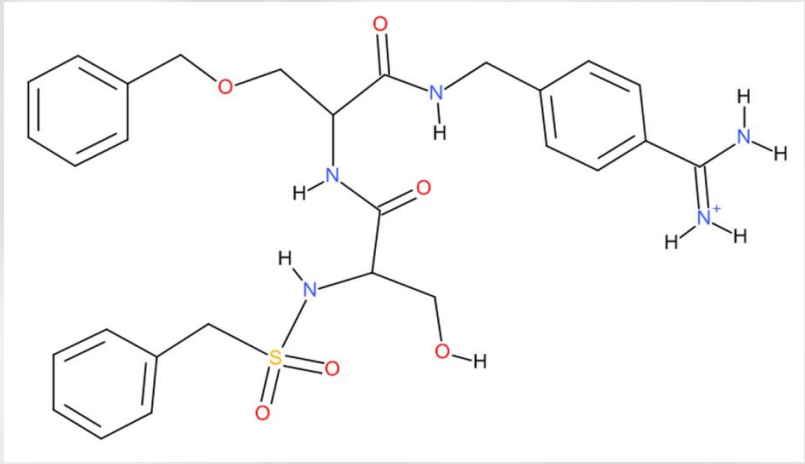
Rotation as a whole

$$Z_r^i = \frac{(8\pi^2 kT)^{\frac{3}{2}}}{h^3} \sqrt{\pi I_A I_B I_C}$$

Investigation of docking accuracy

- Positioning accuracy,
- To check the ***docking paradigm***
- 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

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

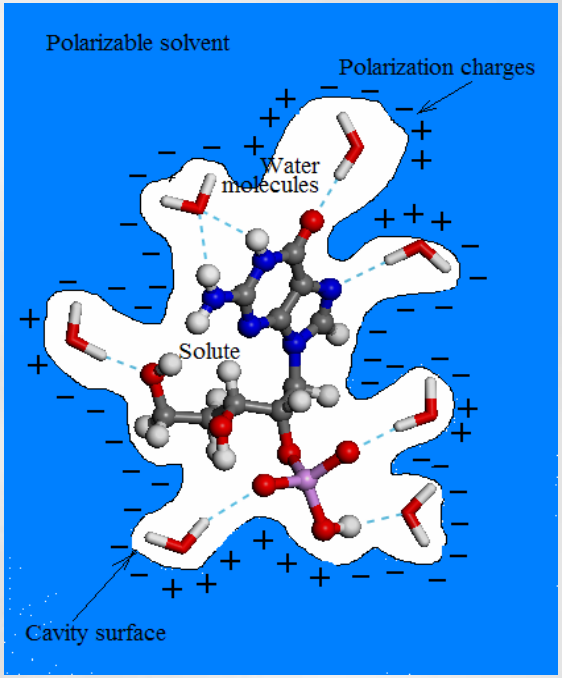
Ligand positioning with *MMFF94 in vacuo* target function

- The docking paradigm *is confirmed* only for *3 complexes out of 16 (6 complexes out of 30) – 20% complexes. For these complexes:*
  - *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*

Conclusions: ligand positioning

- The docking paradigm is correct for only 20% tested protein-ligand complexes for docking target function MMFF94 in vacuum
- The energy target function with implicit solvent model is better than the energy target function in vacuum
- PM7 with solvent (COSMO) is better than MMFF94 with solvent (PCM)

PCM – Polirized Continuum Model



Energy of polar interaction with water

$$\Delta G_{pol} = \frac{1}{2}\sum_i Q_i \int_{SES} \frac{\sigma(\vec{r})}{\left|\vec{R}_i - \vec{r}\right|} dS$$

Integral equation for polarized charges

$$\sigma(\vec{r}) = \frac{(1-\varepsilon)}{2\pi(1+\varepsilon)}\left(\sum_i \frac{Q_i\left((\vec{r}-\vec{R}_i)\cdot\vec{n}\right)}{\left|\vec{r}-\vec{R}_i\right|^3} + \int_{SES} \frac{\sigma(\vec{r}')\left((\vec{r}-\vec{r}')\cdot\vec{n}\right)}{\left|\vec{r}-\vec{r}'\right|^3} dS'\right)$$

Strong interactions of the ligand and the protein with water (ε =78) – strong screening of Coulomb interactions between protein and ligand atoms

MMFF94 *in vacuum* binding energy and its components

PDB ID	ΔG_{exp} kcal/mol	ΔG_{bind} kcal/mol	ΔE, kcal/mol	ΔG_v, kcal/mol	ΔG_t, kcal/mol	ΔG_r, kcal/mol	ΔG_{all}, kcal/mol
4FT0	-10.1	82.2	63.8	-3.9	10.6	10.3	1.4
4FT9	-10.9	-48.5	-63.7	-5.0	10.4	9.9	-0.1
4FSW	-6.8	-44.2	-60.2	-3.7	10.3	9.4	0.0
4FTA	-9.8	-9.8	-30.4	-0.2	10.6	10.1	0.1
4FV5	-10.9	-79.3	-102.1	0.5	10.7	10.6	1.0
4FV6	-12.3	-74.6	-96.8	-0.4	10.8	10.7	1.1
1DWC	-10.5	-128.5	-144.9	-4.2	10.9	10.5	0.2
1TOM	-11.8	-224.1	-248.7	2.7	10.7	10.5	0.7
1C5Y	-6.0	-16.3	-34.8	-0.7	10.0	8.7	0.5
1F5L	-7.5	34.7	17.4	-2.3	10.2	9.3	0.1
1O3P	-9.4	23.8	3.9	-2.0	10.6	10.2	1.1
1SQO	-10.6	-4.6	-24.8	-0.4	10.3	9.6	0.7
1VJ9	-10.7	-25.9	-50.7	2.4	11.0	10.8	0.6
1VJA	-10.9	-31.0	-51.1	-2.1	10.9	10.5	0.8
2P94	-13.0	-42.0	-68.2	-5.5	10.9	10.8	0.0
3CEN	-11.7	-49.2	-69.5	-1.9	10.9	10.6	0.7

Binding energies: Energy [complex (global minimum)] – Energy [protein] – Energy [ligand(global minimum)]

protein	PDB ID	experiment	MMFF94	MMFF94PCM	PM7	PM7COSMO
chk1	4FT0	-10,07395188	63,841268	0,041591	-39,47881	-48,67162
	4FT9	-10,92016068	-63,719203	-9,976033	-111,64202	-48,76744
	4FSW	-6,808284844	-60,198612	-5,889005	-108,48227	-46,41323
	4FTA	-9,786382101	-30,356895	-5,439266	-126,54093	-58,60897
erk2	4FV5	-10,92016068	-102,009676	-7,42143	-168,4395	-54,36695
	$FV6	-12,33855309	-96,91791	-7,753661	-164,91241	-72,6775
trombin	1dwc	-10,50877778	-146,157975	-33,431619	-194,69286	-70,11618
	1tom	-11,774118	-248,286861	-49,658169	-258,09958	-73,67144
urokinase	1c5y	-5,958183512	-34,833614	-79,149206	-33,81052	-52,83622
	1f5l	-7,483047102	17,401276	-52,264153	-29,98013	-81,95168
	1o3p	-9,443057187	3,593206	-40,299811	-32,5346	-64,5975
	1sqo	-10,57543735	-24,781983	-50,418942	-58,75228	-69,56098
	1vj9	-10,71289378	-49,570314	-50,034524	-91,68809	-76,58262
	1vja	-10,92016068	-49,86702	-51,913132	-95,42522	-81,39721
factor xa	2p94	-13,01222594	-68,200522	-15,739433	-153,53035	-68,94207
	3cen	-11,71540693	-69,478796	-18,165638	-133,99803	-63,11649
	correlation with experim		0,407213426	-0,360650906	0,598182678	0,330909633

Ligand deformation energy = (Energy of the free ligand in the binding conformation) - (Energy of the free ligand global minimum)

Protein	PDBID	E^{ld}, kcal/mol
urokinase	1C5Y	4.591
	1F5L	5.301
	1O3P	19.85
	1SQO	11.54
	1VJ9	47.51
	1VJA	47.16
factor Xa	2P94	23.33
	3CEN	20.25

Conclusions: binding free energy in the multi-well approximation

- There are still significant disagreements between calculated and measured binding free energies

- MMFF94 in vacuum – the worst case
- MMFF94 + PCM water model – the intermediate case
- PM7 +COSMO water model – the best case

- ***Desolution energy gives*** considerable contribution to binding energy. It can be as high as 200 kcal/mol – screening of protein-ligand Coulomb interactions.

$$U_W(\text{Protein - Ligand}) - U_W(\text{Protein}) - U_W(\text{Ligand})$$

Conclusions: binding free energy in the multi-well approximation

- The most valuable energy component is the potential energy of the global minimum
- contributions of vibrational degrees of freedom or multiple minima accounting are small amendments
- entropy contribution to the binding energy is comparable to the enthalpy contribution
- binding free energy components, corresponding to translational and rotational degrees of freedom, are practically constant for all tested protein-ligand complexes
- The ligand deformation energy can be as high as several dozen kcal/mol

To improve docking accuracy:

- Improvement positioning results in the improvement of the binding energy
- To find better force field
- To use implicit water models
- Force fields should be substituted by quantum chemistry – semiempirical method PM7

Success – Join efforts of:

- **Mathematicians** – effective algorithms of the generalized global energy minimum search on multidimentional surface with a large number of dimentions
- **Physicists, chemists** – force fields, quantum chemistry, inter- and intra-molecular interactions in water
- **Programmers** – effective programs, parallel multi-processors high performance supercomputing
- **Chemists –** design of inhibitors candidates
- **Biochemists, biophysicists, medics** – experimental confirmation

Discussion questions?

1. Are there docking programs without fitting parameters?
2. What force field is the best for docking and MD?
3. Why the local optimization of the native ligand results sometimes in a large displacement – RMSD > 1.5 Å?
4. Is it possible to calculate the solution energy in water with accuracy better than 1 kcal/mol?
5. Is it possible to perform docking with explicit water model?
6. Can mobility of the protein atoms improve docking accuracy?
7. Is quantum-chemical docking feasible?

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!