Canada NRC·CNRC

UNIVERSITY OF ALBERTA
EDMONTON·ALBERTA·CANADA

National Institute for Nanotechnology

Brainstorming meeting on "Solvation, non-uniform polarizability, and local field effects in solids, liquids, life, and devices"

Stewart Blusson Quantum Matter Institute
University of British Columbia, Vancouver, BC, 11 – 14 April 2016

Molecular theory of solvation for chemistry, biochemistry, biophysics, and nanomaterials

Andriy Kovalenko

National Institute for Nanotechnology, and University of Alberta
Edmonton, Canada

Brainstorming meeting on "Solvation, non-uniform polarizability, and local field effects in solids, liquids, life, and devices”

Stewart Blusson Quantum Matter Institute
University of British Columbia, Vancouver, BC, 11 – 14 April 2016

Molecular theory of solvation for chemistry, biochemistry, biophysics, and nanomaterials

Andriy Kovalenko

National Institute for Nanotechnology, and University of Alberta
Edmonton, Canada

Canada
NRC·CNRC

UNIVERSITY OF ALBERTA
EDMONTON·ALBERTA·CANADA

National Institute for Nanotechnology

Brainstorming meeting on "Solvation, non-uniform polarizability, and local field effects in solids, liquids, life, and devices”

Stewart Blusson Quantum Matter Institute
University of British Columbia, Vancouver, BC, 11 – 14 April 2016

Molecular theory of solvation for chemistry, biochemistry, biophysics, and nanomaterials

Andriy Kovalenko

National Institute for Nanotechnology, and University of Alberta
Edmonton, Canada

-30°C

Canada
NRC·CNRC

UNIVERSITY OF ALBERTA
EDMONTON·ALBERTA·CANADA

National Institute for Nanotechnology

Brainstorming meeting on "Solvation, non-uniform polarizability, and local field effects in solids, liquids, life, and devices”

Stewart Blusson Quantum Matter Institute
University of British Columbia, Vancouver, BC, 11 – 14 April 2016

Molecular theory of solvation for chemistry, biochemistry, biophysics, and nanomaterials

Andriy Kovalenko

National Institute for Nanotechnology, and University of Alberta
Edmonton, Canada

-30°C

+30°C

Canada NRC·CNRC

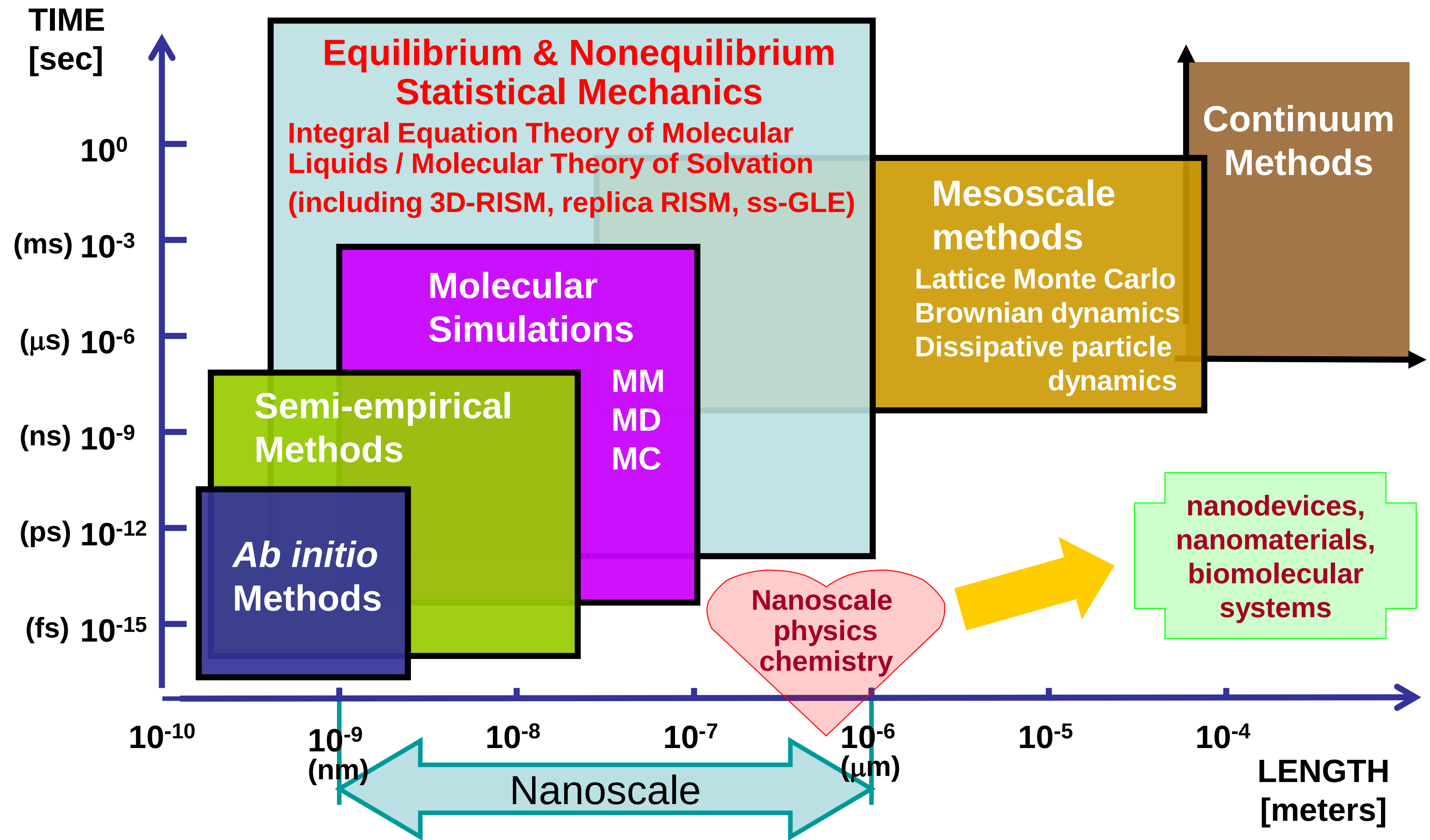
UNIVERSITY OF ALBERTA
EDMONTON·ALBERTA·CANADA

National Institute for Nanotechnology

ANDRIY KOVALENKO GROUP
Multiscale Theory and Modeling

The Kovalenko Group
research program:

Platform of Theory, Modeling, and Simulation on Multiple Space and Time Scales



Equilibrium & Nonequilibrium Statistical Mechanics

Integral Equation Theory of Molecular Liquids / Molecular Theory of Solvation (including 3D-RISM, replica RISM, ss-GLE)

Molecular Simulations

MM
MD
MC

Semi-empirical Methods

***Ab initio* Methods**

Mesoscale methods

Lattice Monte Carlo
Brownian dynamics
Dissipative particle dynamics

Continuum Methods

Nanoscale physics chemistry

nanodevices, nanomaterials, biomolecular systems

Contents

Essentials of statistical-mechanical molecular theory of solvation, or integral equation theory of liquids

Integro-differential equations for one-particle density

Fluctuating liquid (liquid interface) with particle density $\rho_a(r)$
in external potential field $u_a(r)$:

- mechanical and chemical balance (between the phases)

Born-Green-Yvon equation (part of BBGKY hierarchy) involves two-body correlations:
distribution functions $g_{ab}(r_1,r_2)$

$$\nabla_1 \log\rho_a(\mathbf{r}_1) = -\nabla_1 \beta u_a^{\text{ext}}(\mathbf{r}_1) + \sum_b \int \mathrm{d}\mathbf{r}_2 \rho_b(\mathbf{r}_2) g_{ab}(\mathbf{r}_1,\mathbf{r}_2;[\rho]) \nabla_2 \beta u_{ab}(\mathbf{r}_1,\mathbf{r}_2)$$

Lovett-Mou-Buff-Wertheim equation involves two-body correlations:
direct correlation functions $c_{ab}(r_1,r_2)$

$$\nabla_1 \log\rho_a(\mathbf{r}_1) = -\nabla_1 \beta u_a^{\text{ext}}(\mathbf{r}_1) + \sum_b \int \mathrm{d}\mathbf{r}_2 c_{ab}(\mathbf{r}_1,\mathbf{r}_2) \nabla_2 \rho_b(\mathbf{r}_2)$$

Ornstein-Zernike type integral equation theory

J.-P. Hansen and I. McDonald, Theory of Simple Liquids, 3rd ed. Elsevier, 2006.

Helmholtz Free Energy

$$\begin{aligned} A &= -k_B T \log(Q_N(V,T)) \\ &= A^{\mathrm{id}} + A^{\mathrm{ex}} = k_B T \left[N\{\log(\rho) + 3\log\Lambda - 1\} - \log \frac{Z_N(V,T)}{V^N} \right] \end{aligned}$$

Partition Function

$$Q_N(V,T) = \frac{\Lambda^{-3N}}{N!} \int dr^N \exp(-\beta U(r_1, \cdots, r_N)) = \frac{\Lambda^{-3N}}{N!} Z_N(V,T)$$

Configurational Integral

$$Z_N = \int dr^N \exp(-\beta U(r_1, \cdots, r_N)), \quad \beta = \frac{1}{k_B T}$$

Ornstein-Zernike type integral equation theory

$$Z_N = \int dr^N \exp(-\beta U(r_1, \cdots, r_N))$$

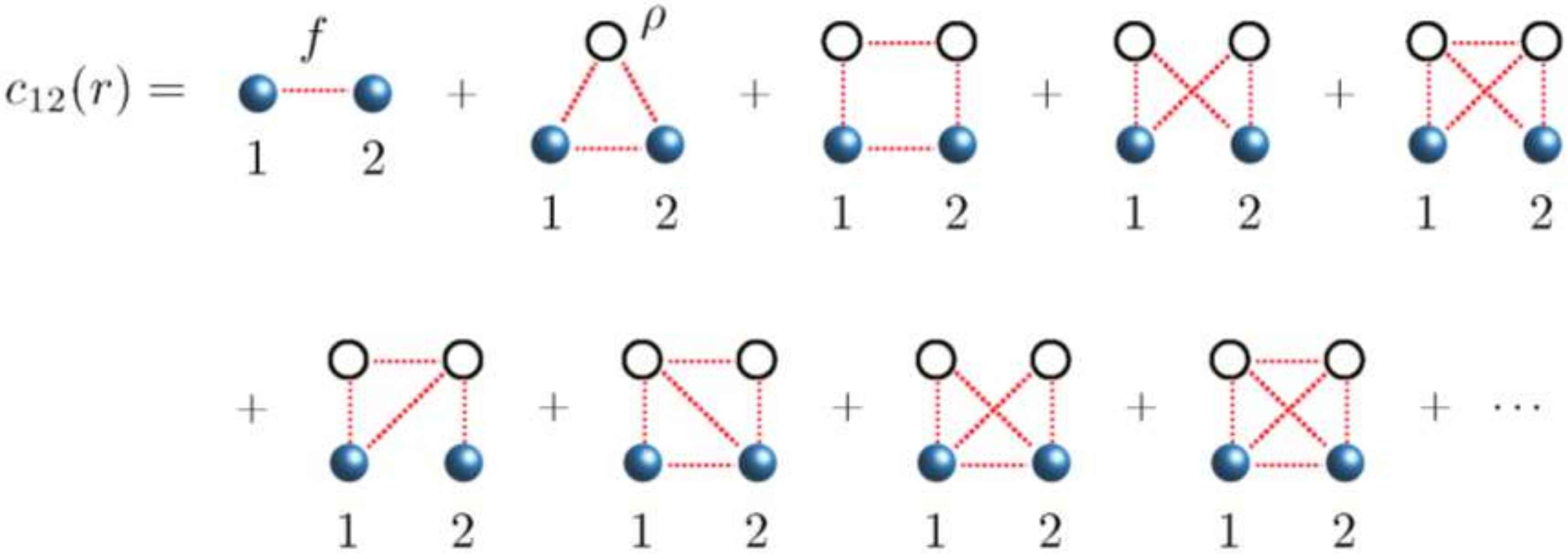
Pairwise-additive Interaction Potential: $U(r_1, \cdots, r_N) = \sum_{i \neq j} u_{ij}(r_{ij})$

$$Z_N = \int dr^N \prod_{i \neq j} \exp(-\beta u_{ij}(r_{ij})) = \int dr^N \prod_{i \neq j} \{\exp(-\beta u_{ij}(r_{ij})) - 1 + 1\}$$

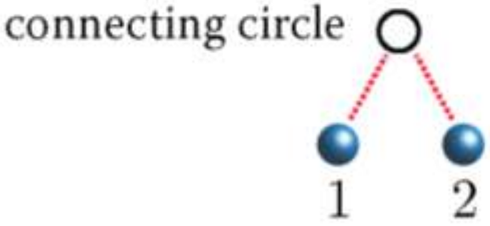
Mayer *f*-function: $f_{ij}(r_{ij}) \equiv \exp(-\beta u_{ij}(r_{ij})) - 1$

$$Z_N = \int dr^N \prod_{i \neq j} \{f_{ij}(r_{ij}) + 1\}$$

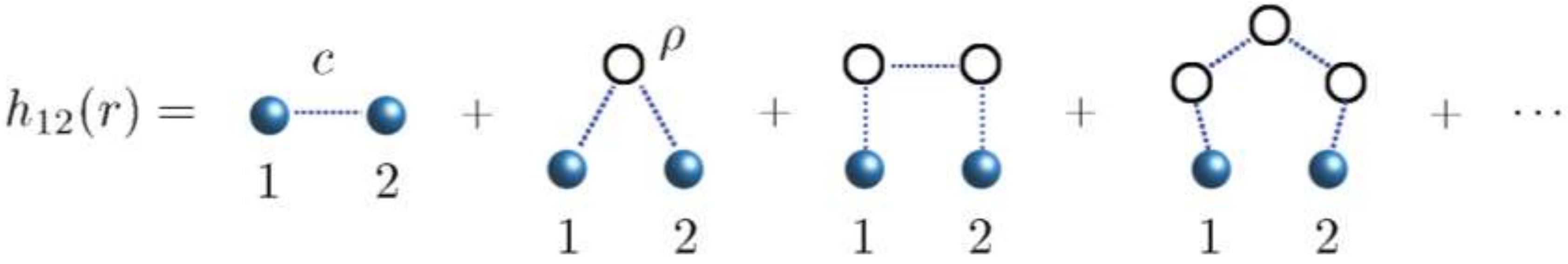
Ornstein-Zernike type integral equation theory



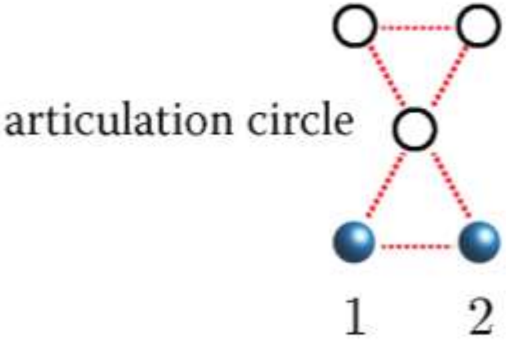
Direct Correlation Function = { all simple diagrams that consist of two root circles(1 and 2), field ρ-circles, and *f*-bonds, and are free of connecting circles }



Ornstein-Zernike type integral equation theory

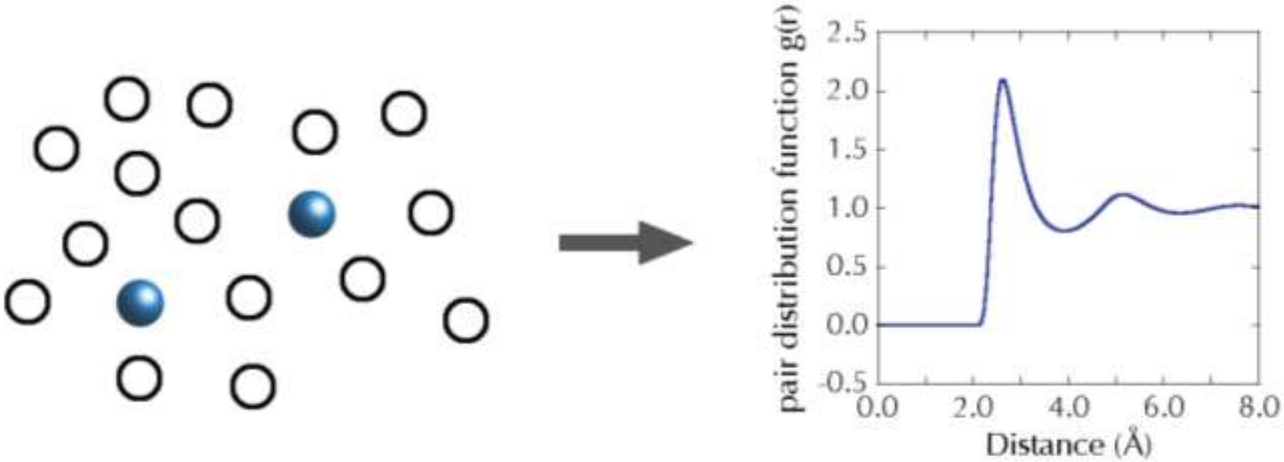


Total Correlation Function = { all simple diagrams that consist of two root circles(1 and 2), field ρ-circles, and f-bonds, and are free of articulation circles }

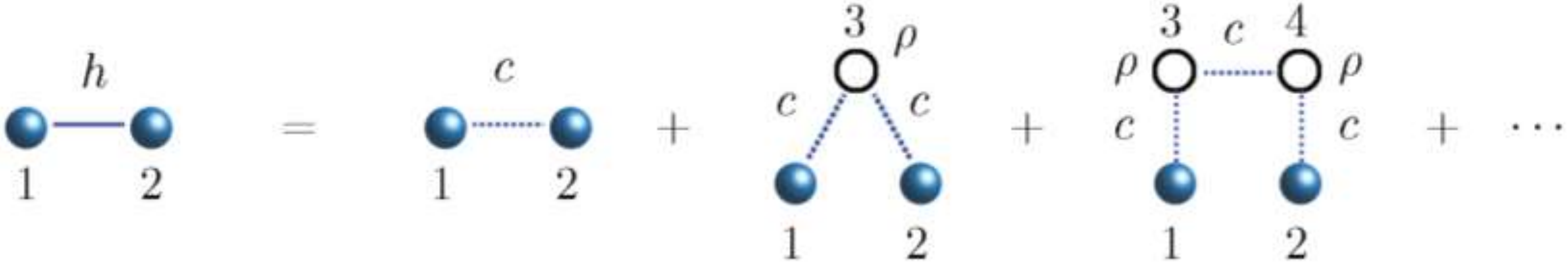


Ornstein-Zernike type integral equation theory

In liquid



$$h_{12}(r) = g_{12}(r) - 1$$



$$h_{12}(\mathbf{r}_i,\mathbf{r}_j) = c_{12}(\mathbf{r}_i,\mathbf{r}_j) + \int d\mathbf{r}_k c_{13}(\mathbf{r}_i,\mathbf{r}_k)\rho_3(\mathbf{r}_k)c_{32}(\mathbf{r}_k,\mathbf{r}_j) + \int d\mathbf{r}_k \int d\mathbf{r}_l c_{13}(\mathbf{r}_i,\mathbf{r}_k)\rho_3(\mathbf{r}_k)c_{34}(\mathbf{r}_k,\mathbf{r}_l)\rho_4(\mathbf{r}_l)c_{42}(\mathbf{r}_l,\mathbf{r}_j) + \cdots$$

Resummation of the infinite series leads to the Ornstein-Zernike(OZ) equation

$$h_{12}(\mathbf{r}_i,\mathbf{r}_j) = c_{12}(\mathbf{r}_i,\mathbf{r}_j) + \int d\mathbf{r}_k c_{13}(\mathbf{r}_i,\mathbf{r}_k)\rho_3(\mathbf{r}_k)h_{32}(\mathbf{r}_k,\mathbf{r}_j)$$

Ornstein-Zernike type integral equation theory

D. Chandler
F. Hirata
B.M. Pettitt
P. Rossky

RISM equation (generalization of OZ equation to molecular liquids)

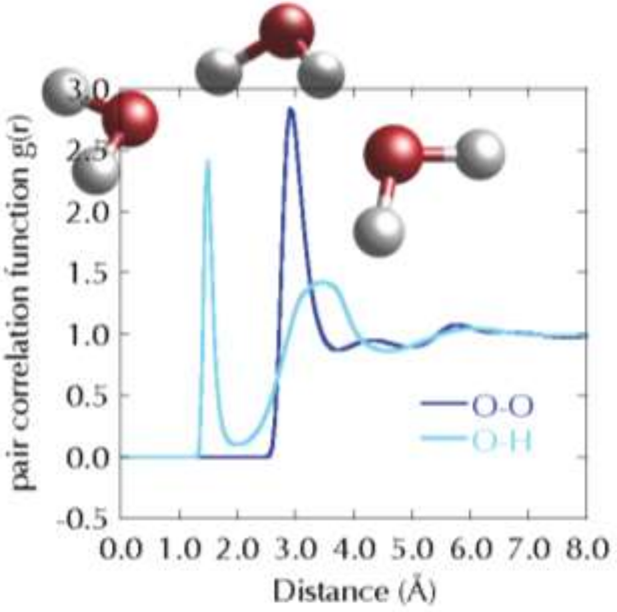
$$h_{\alpha s}(r) = \sum_{\gamma s'} \omega_{\alpha\gamma} * c_{\gamma s'} * \omega_{s' s} + \sum_{\gamma s'} \omega_{\alpha\gamma} * c_{\gamma s'} * \rho_{s'} h_{s' s}$$

$$\omega_{\alpha\gamma}(r) = \frac{\delta(r - l_{\alpha\gamma})}{4\pi l_{\alpha\gamma}^2}$$ – intramolecular matrix

Closure relation

$$g_{\alpha s}(r) = \exp(-\beta u_{\alpha s} + h_{\alpha s} - c_{\alpha s} + b_{\alpha s})$$

$$g_{\alpha s}(r) = h_{\alpha s}(r) + 1$$



ambient liquid water

RISM theory yields

- solvation structure (radial distribution function)
- solvation thermodynamics (free energy, enthalpy, entropy, partial molar volume, compressibility, etc.)

Ornstein-Zernike type integral equation theory

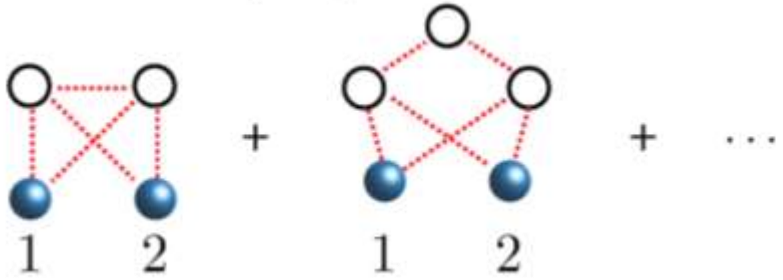
D. Chandler
F. Hirata
B.M. Pettitt
P. Rossky
G. Patey
A. Kovalenko
S. Kast

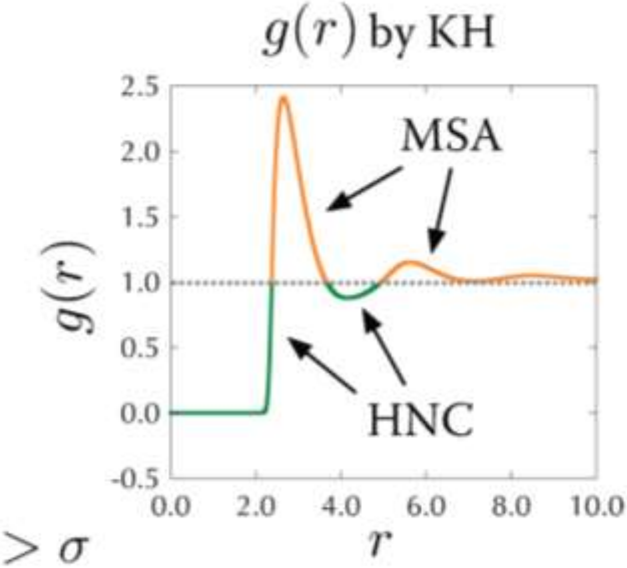
Exact closure relation

$$g_{12}(r) = \exp(-\beta u_{12}(r) + h_{12}(r) - c_{12}(r) + b_{12}(r))$$

Bridge function (bridge or elementary diagrams)

$$b_{12}(r) =$$





Hypernetted chain approximation (HNC): $b(r) = 0$

Random phase approximation (RPA): $c(r) = -\beta u(r)$

Mean spherical approximation (MSA): $c(r) = -\beta u(r), \quad r > \sigma$

$$g(r) = 0, \quad r < \sigma$$

Kovalenko-Hirata approximation (KH): $g(r) = \exp(-\beta u + h - c), \quad g(r) < 1$

$$= 1 - \beta u + h - c, \quad g(r) > 1$$

Kast approximation
partial series expansion of order n (PSE-n)

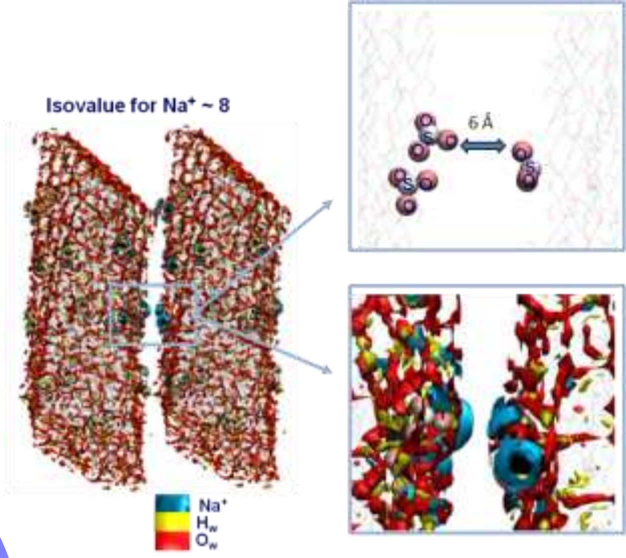
$$g(r) = \exp\left(-\beta u + h - c\right) \qquad g(r) < 1$$

$$= \sum_{i=1}^{n} \left(-\beta u + h - c\right)^{i} / i\,!, \quad g(r) < 1$$

Statistical-mechanical, 3D-RISM-KH molecular theory of solvation

3D-RISM-KH molecular theory of solvation

3D site DFT formalism: Chandler; McCoy; Singer. *JCP* 1986.
3D-RISM-HNC from 3D site DFT: Beglov; Roux. *JPCB* 1997.
3D-RISM from Molecular OZ; KH closure; SCF KS-DFT: Kovalenko; Hirata. *CPL* 1998. *JCP* 1999.
MDIIS accelerated solver: Kovalenko, Ten-no, Hirata. *JCC,* 1999. Kovalenko; Hirata. *JCP* 2000.
Analytical asymptotics: Kovalenko; Hirata. *JCP* 2000. Gusarov; Pujari; Kovalenko. *JCC* 2012.
Multiscale couplings: Kovalenko. in: *Molecular Theory of Solvation*, Kluwer, 2003; *Pure Appl.Chem.* 2013.



Water and Na⁺ counterions in 1M NaCl aqueous electrolyte solution around a pair of cellulose nanocrystals (CNCs) modified with SO_3^- surface groups

- **Three-dimensional reference interaction site model (3D-RISM) integral equation:**

$$h_\gamma(\mathbf{r}) = \sum_\alpha \int d\mathbf{r}'\, c_\alpha(\mathbf{r}-\mathbf{r}')\,\chi_{\alpha\gamma}(r')$$

where

$$g_\gamma(\mathbf{r}) = h_\gamma(\mathbf{r}) + 1$$ - distribution function / total correlation function

$$c_\gamma(\mathbf{r}) \propto -u_\gamma(\mathbf{r})/(k_\mathrm{B}T)$$ - direct correlation function
outside repulsive core: interaction potential (force field)
inside repulsive core: related to solvation free energy

$$\chi_{\alpha\gamma}(r) = \underbrace{\omega_{\alpha\gamma}(r)}_{\text{molecular geometry}} + \underbrace{\rho_\alpha}_{\text{solvent density}} \underbrace{h_{\alpha\gamma}(r)}_{\text{intermolecular correlations}}$$ - solvent susceptibility (reaction to biomolecule insertion)

$$\Delta G = \rho \int_0^{\mathrm{full}\,u} g\,\delta u$$

- **Kovalenko-Hirata (KH) closure relation**

$$g_\gamma(\mathbf{r}) = \begin{cases} \exp\left(-u_\gamma(\mathbf{r})/(k_\mathrm{B}T) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})\right) & \text{for} \quad g_\gamma(\mathbf{r}) < 1 \\ 1 - u_\gamma(\mathbf{r})/(k_\mathrm{B}T) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r}) & \text{for} \quad g_\gamma(\mathbf{r}) \geq 1 \end{cases}$$

Kast approximation: partial *n*-order series expansion (PSE-n)

$$g(r) = \exp\left(-\beta u + h - c\right) \quad g(r) < 1$$
$$= \sum_{i=1}^{n} \left(-\beta u + h - c\right)^i / i\,!, \; g(r) > 1$$

- **Solvation structure and thermodynamics in closed analytical form**

KH solvation Free Energy functional: $$\Delta G = \sum_\gamma \int_V d\mathbf{r}\, \Phi_\gamma^{\mathrm{KH}}(\mathbf{r})$$ **with KH SFE density** $$\Phi_\gamma^{\mathrm{KH}}(\mathbf{r}) = \rho_\gamma k_\mathrm{B} T \left[\frac{1}{2} h_\gamma^2(\mathbf{r}) \Theta(-h_\gamma(\mathbf{r})) - c_\gamma(\mathbf{r}) - \frac{1}{2} h_\gamma(\mathbf{r}) c_\gamma(\mathbf{r}) \right]$$

Partial Molar Volume (PMV): $$\bar{V} = k_B T \chi_T \left(1 - \sum_\gamma \rho_\gamma \int d\mathbf{r}\, c_\gamma(\mathbf{r}) \right)$$ **with solvent compressibility** $$\rho k_B T \chi_T = \left(1 - 4\pi \sum_{\alpha\gamma} \rho_\alpha \int_0^\infty r^2 dr\, c_{\alpha\gamma}(r) \right)^{-1}$$

PMV-correlated correction to SFED functional: $$\Delta \overline{G} = \Delta G + \alpha(\rho \bar{V}) + \beta$$

Enthalpy-Entropy Decomposition: $$\Delta G^{\mathrm{u}} = \Delta\varepsilon^{\mathrm{uv}} + \Delta\varepsilon^{\mathrm{vv}} - T\Delta s_V^{\mathrm{u}}\,, \qquad \Delta s_V^{\mathrm{u}} = -T^{-1}\left(\partial \Delta G^{\mathrm{u}} / \partial T\right)_{\rho^{\mathrm{v}}}$$

Potentials of mean force: **solute-solute** $$w(\mathbf{1},\mathbf{n}) = u(\mathbf{1},\mathbf{n}) + \Delta G(\mathbf{1},\mathbf{n}) - \sum_i^n \Delta G(\mathbf{i})$$ **solute-solvent site** $$w_\alpha(\mathbf{r}) = -k_\mathrm{B} T \ln\left(g_\alpha(\mathbf{r})\right)$$

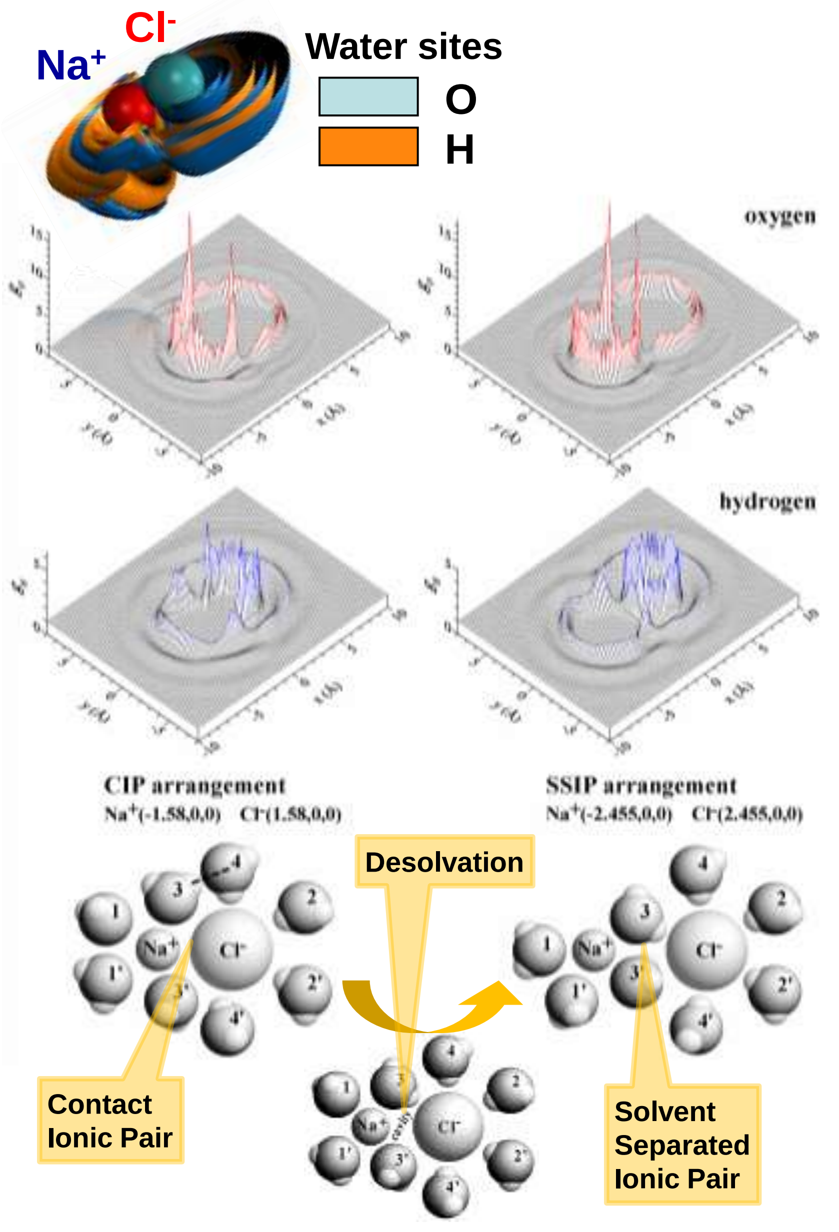
Mean Solvation Forces, Analytical Gradients:

$$\mathbf{f}(\mathbf{R}_i) \equiv -\frac{\partial \Delta G}{\partial \mathbf{R}_i} = -\sum_\gamma \rho_\gamma \int d\mathbf{r}\, g_\gamma(\mathbf{r}) \frac{\partial u_{i\gamma}(\mathbf{r}-\mathbf{R}_i)}{\partial \mathbf{R}_i}$$

Validation of 3D-RISM-KH theory:
Solvation structure and potentials of mean force of ion pairs in aqueous electrolyte solution

Potential of mean force between simple ions in aqueous solution by 3D-RISM molecular theory of solvation

Kovalenko and Hirata, JCP, **112**, 10391; 10403 (2000)

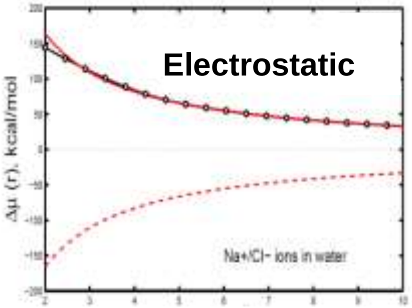


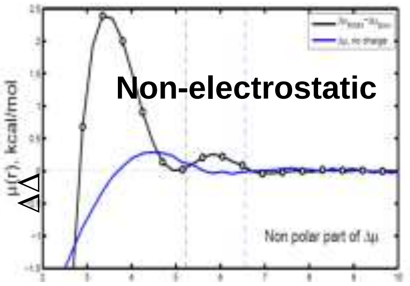
Desolvation

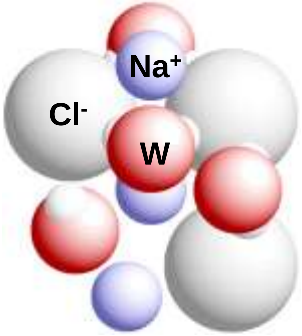
Contact Ionic Pair

Solvent Separated Ionic Pair

Water around NaCl iin the CIP and SSIP arrangements



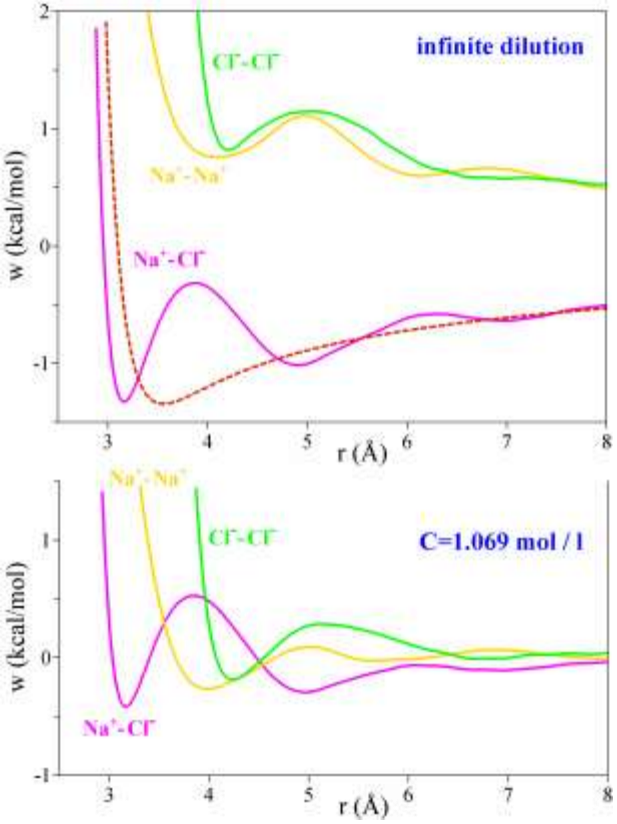




Cluster arrangement of Na^+, Cl^-, and Water in 1 mol/l solution

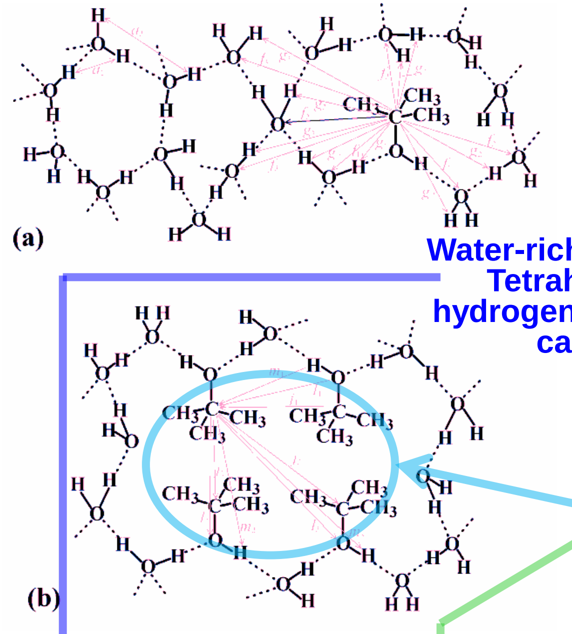
Chemical potential of NaCl pair in aqueous solution

PMFs of Na^+ and Cl^- in aqueous solution



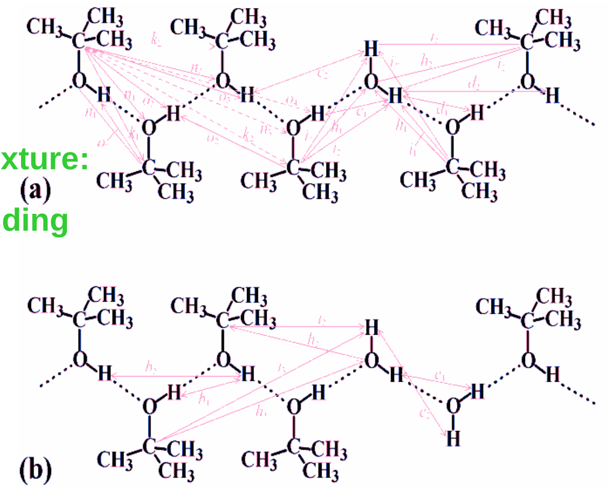
Validation of 3D-RISM-KH theory:
Solvation structure of hydrophobic and amphiphilic molecules in solvent mixtures

Structural transitions in *tert*-butanol (TBA) - water mixture by RISM-KH theory: TBA micromicelles formation and tetrahedral → zigzag hydrogen bonding

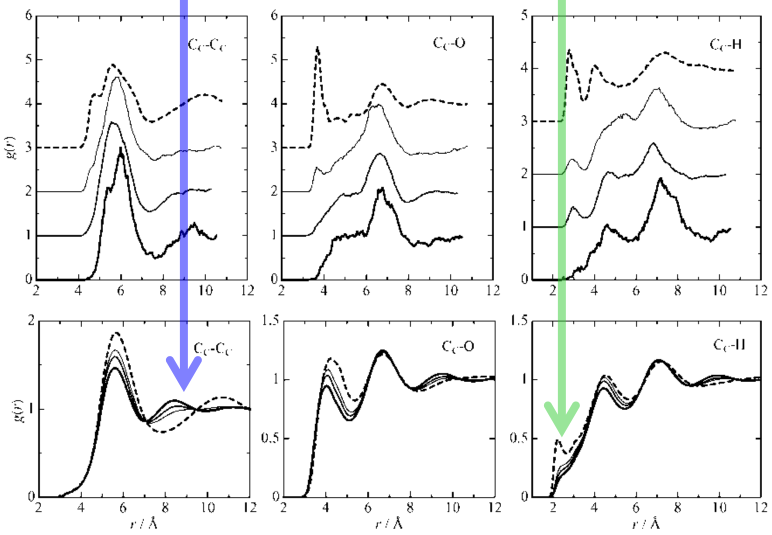


Water-rich mixture:
Tetrahedral
hydrogen bonding
cage

Alcohol-rich mixture:
Zigzag
hydrogen bonding
chains



micromicelle

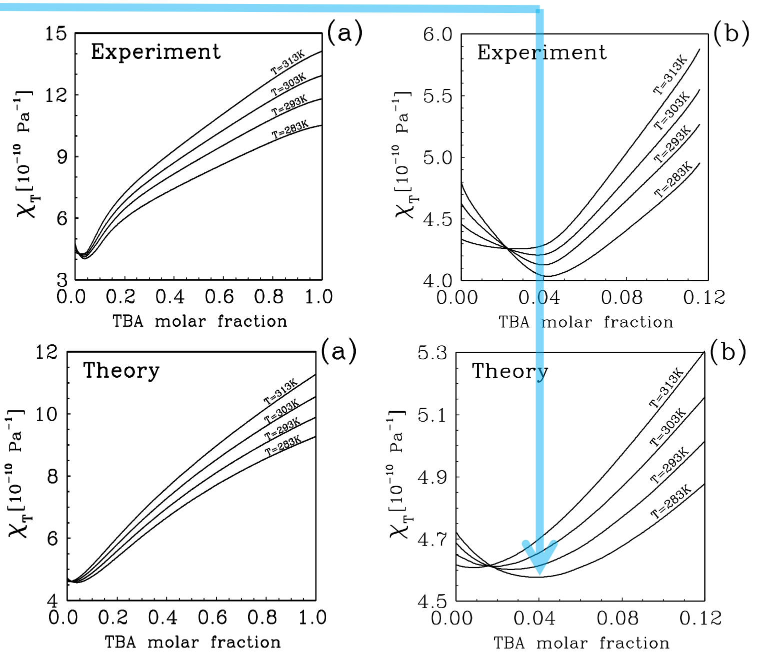


← Neutron diffraction and compressibility →

← experiment →

vs

← RISM-KH theory →



Yoshida, Yamaguchi, Kovalenko, and Hirata, JPCB (2002)

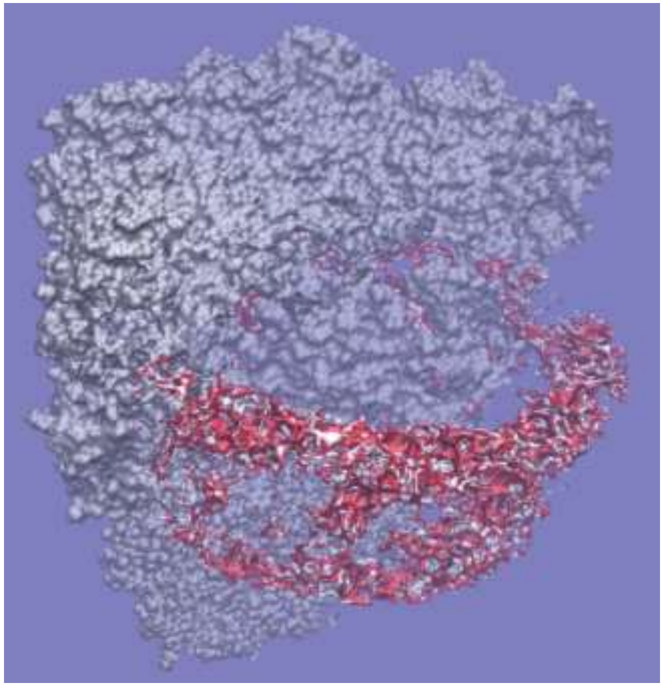
Omelyan, Kovalenko, and Hirata, JTCC (2003)

Validation of 3D-RISM-KH theory: Structural solvent in the confinement of biomolecules

Role of confined water in chaperonin functions

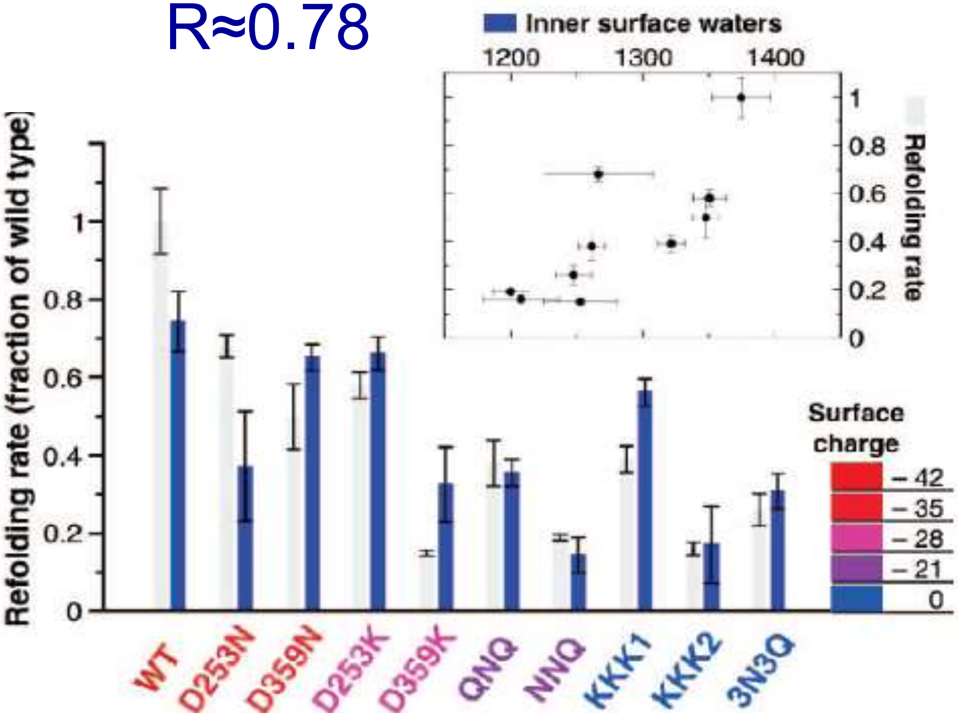
J. L. England, D. Lucent, and V. S. Pande, *JACS* ***130***, 11838 (2008)

Capacity of accumulating water at the chaperonin nanocavity surface correlates with the rate of folding catalized by the chaperonin



Solvation structure of wild type GroEL+ES.
Surface water molecules (red and white.

R≈0.78



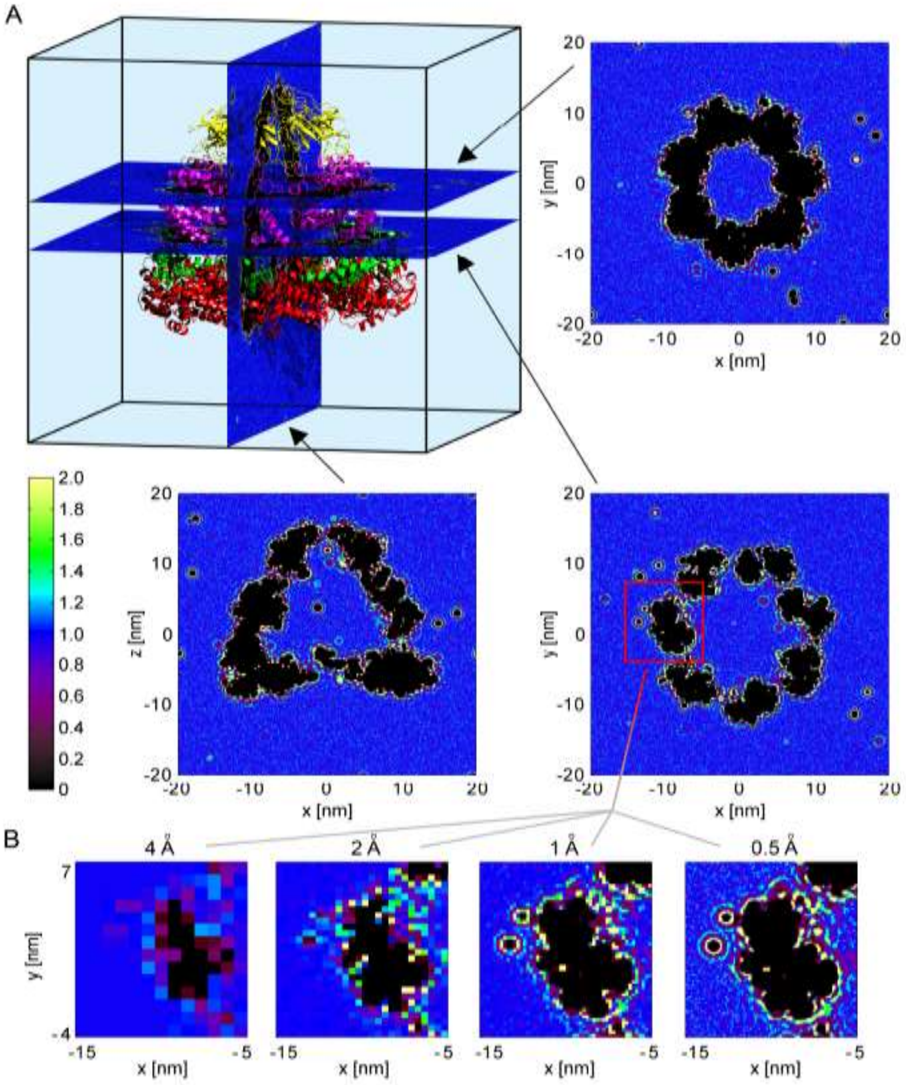
Gray bars: Experimental folding rates of DM-MBP encapsulated inside different GroEL mutants
Blue bars: Best linear fit of the number of surface waters to the measured refolding rate

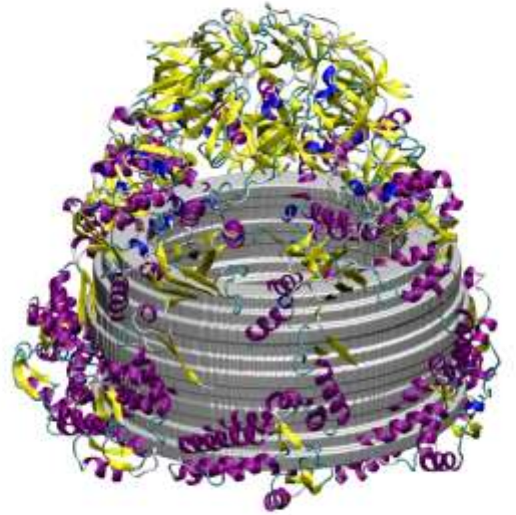
Computational challenge:

Large protein ~70,000 atoms (up to 1,000,000 atoms with solvent),
Internal solvation is essential

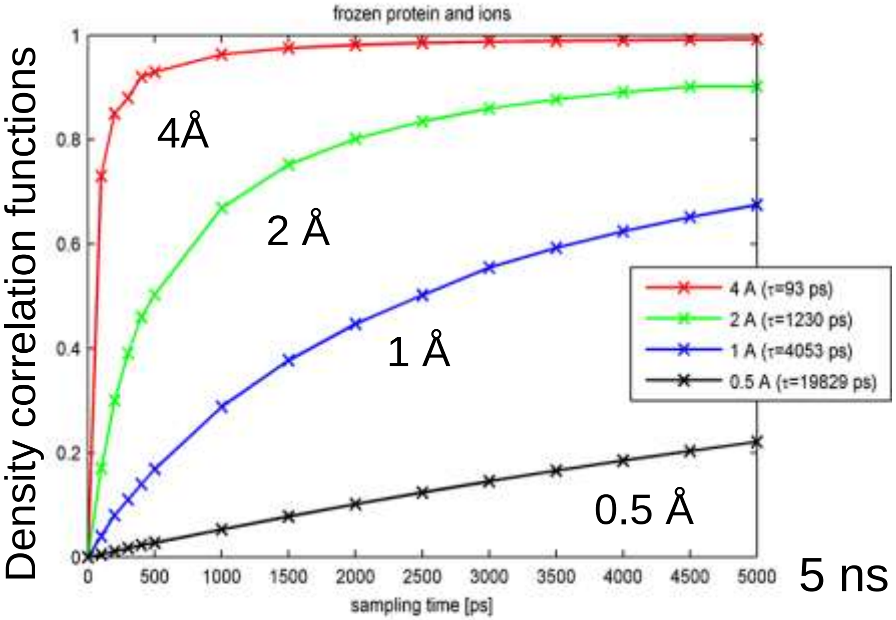
Convergence of MD water density distributions

M. Stumpe, N. Blinov, D. Wishart, A. Kovalenko, and V. Pande, *J. Chem. Phys.* B **115**, 205 (**2011**)



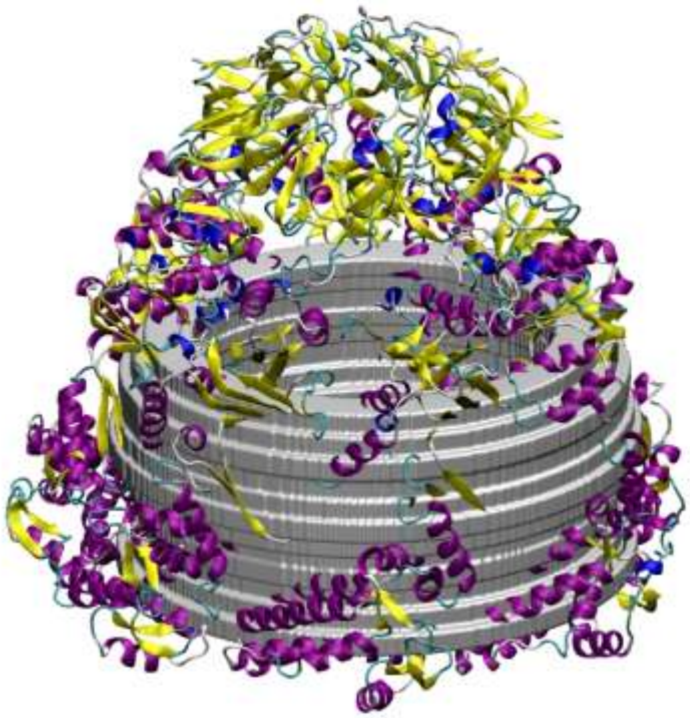


Bnning water molecules in the internal cavity of GroEL/ES

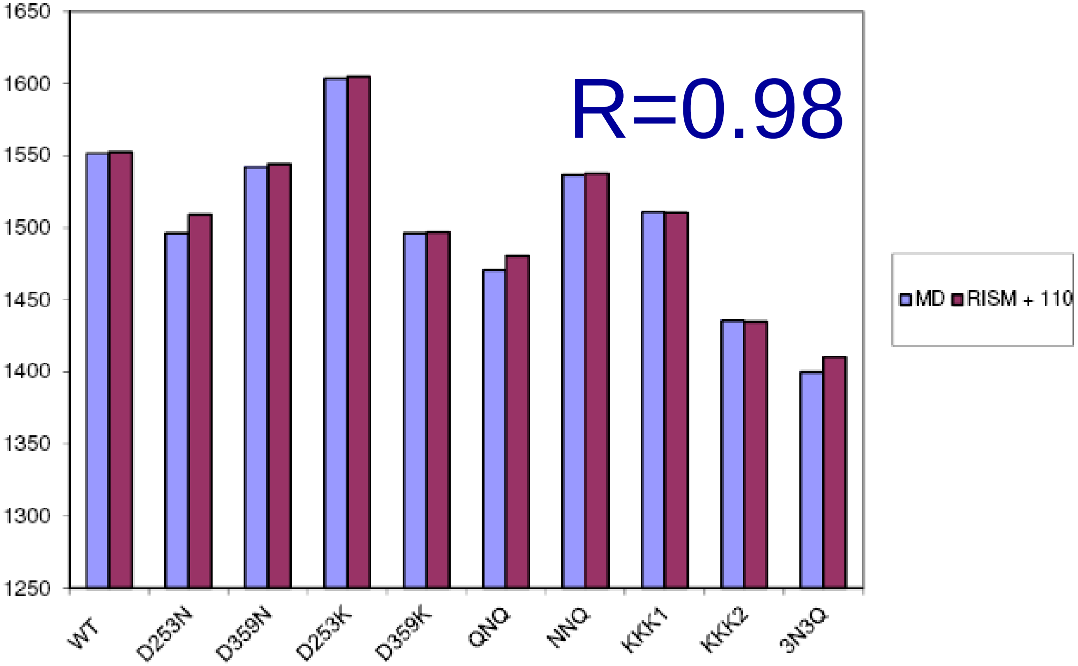


Water in internal cavity of GroRL chaperonin mutants: MD simulation vs 3D-RISM-KH

M. Stumpe, N. Blinov, D. Wishart, A. Kovalenko, and V. Pande, *J. Chem. Phys.* B **115**, 205 (**2011**)



Binning water molecules in the internal cavity of GroEL/ES

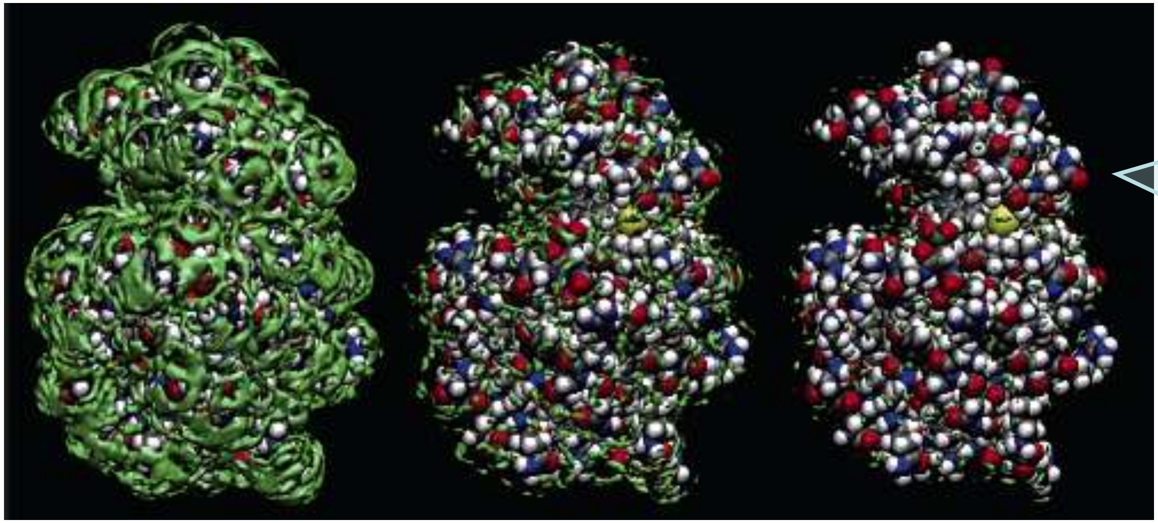


Occupation numbers of water in the internal cavity of the GroEL wild type and mutants

The 3D-RISM-KH theory speeds up calculation of the occupation numbers in the GroEL/ES internal cavity by two orders of magnitude, compared to explicit solvent MD simulation

Structural water in the Lysozyme protein cavity predicted by 3D-RISM-KH theory

T. Imai, R. Hiraoka, A. Kovalenko, and F. Hirata, JACS Communication, **127**, 15334 (2005).



Water distribution around lysozyme

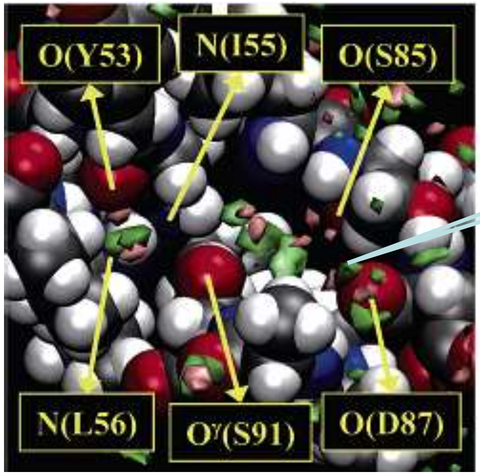
Isosurface of the 3D distribution $g_o(r)$ of water oxygen around lysozyme.
$g_o(r) > 2; 4; 8$ (left, center, right panel)
[Pot.Mean Force > 3.35; 6.7; 13.4 kcal/mol]
Results of the 3D-RISM theory

Water molecules in the cavity surrounded by residues Y53 – I58 and A82 – S91

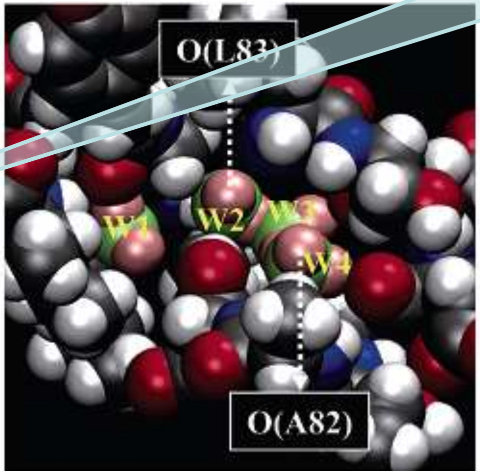
Residues A82 and L83 in the front side are not shown.

Water oxygen (green) and hydrogen (pink) distributions g(r)>8 (PMF>13.4kcal/mol)

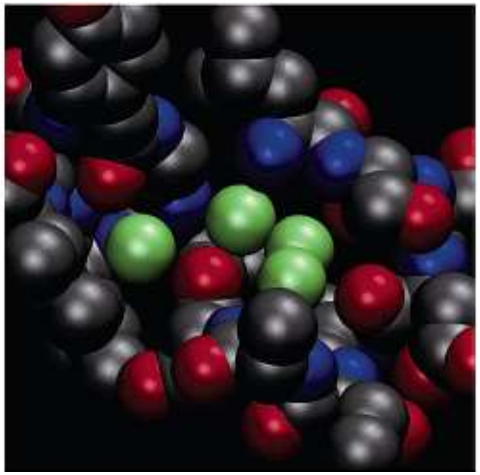
Water distribution in the cavity



Structural water model



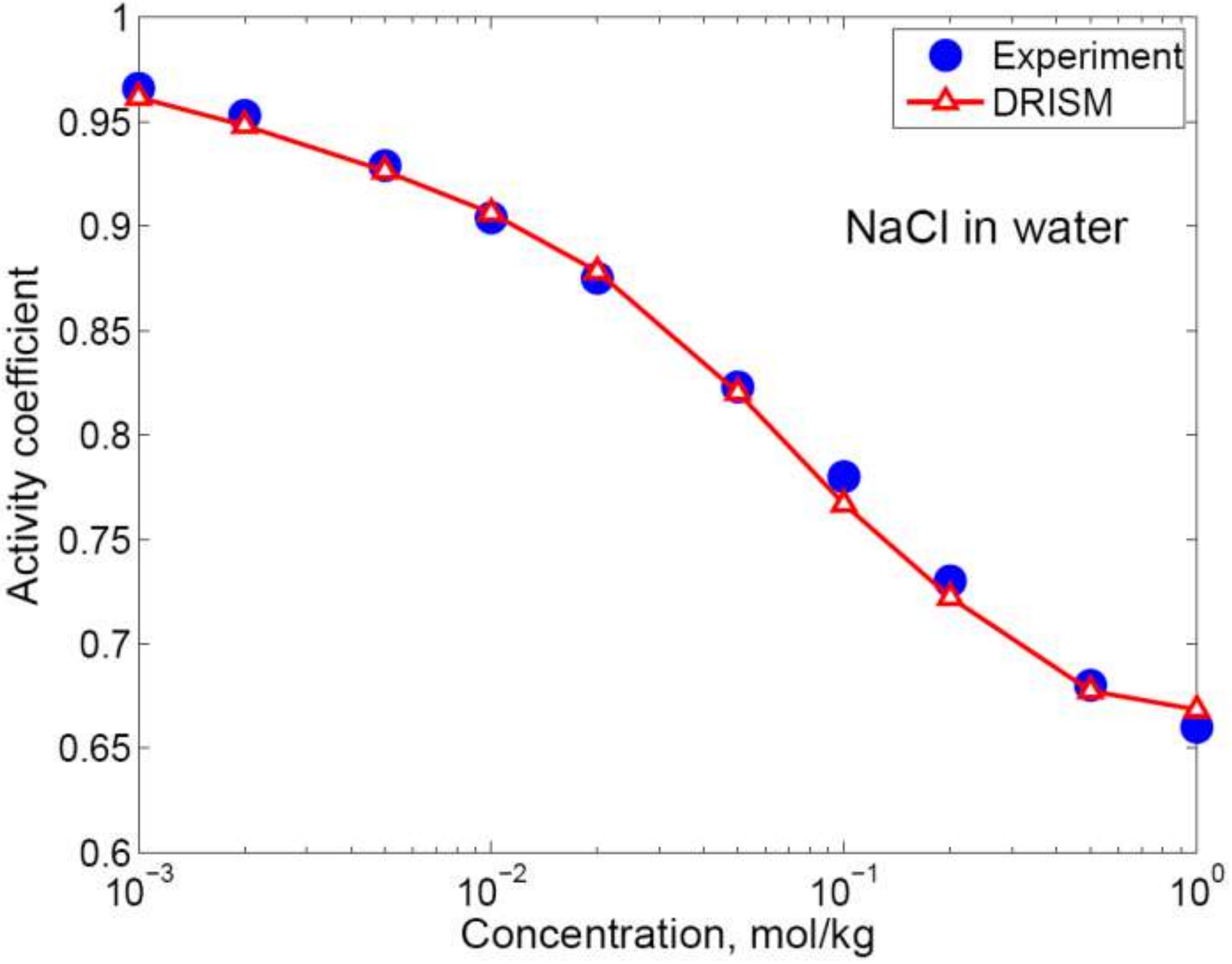
Crystallographic water sites



- **3D-RISM-KH theory detects structural water molecules confined in a small cavity of a protein**
- **The theory can access configurations peculiar to molecules confined in a small space.**

Validation of 3D-RISM-KH theory: Solvation thermodynamics

Salt concentration effects in aqueous solutions



Hydration free energy of small compounds from 3D-RISM-KH with PMV-correlated correction to the free energy functional

$$h_\gamma(\mathbf{r}) = \sum_\alpha \int d\mathbf{r}'\, c_\alpha(\mathbf{r}-\mathbf{r}')\,\chi_{\alpha\gamma}(r')$$

3D-RISM integral equation

$$g_\gamma(\mathbf{r}) = h_\gamma(\mathbf{r}) + 1$$

$$g_\gamma(\mathbf{r}) = \begin{cases} \exp\left(-u_\gamma(\mathbf{r})/(k_\mathrm{B}T) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})\right) & \text{for } g_\gamma(\mathbf{r}) < 1 \\ 1 - u_\gamma(\mathbf{r})/(k_\mathrm{B}T) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r}) & \text{for } g_\gamma(\mathbf{r}) \geq 1 \end{cases}$$

KH closure

$$\Delta G = \sum_\gamma \int_V \mathrm{d}\mathbf{r}\, \Phi_\gamma^{\mathrm{KH}}(\mathbf{r})$$

KH solvation free energy functional

$$\Phi_\gamma^{\mathrm{KH}}(\mathbf{r}) = \rho_\gamma k_\mathrm{B} T\left[\frac{1}{2}h_\gamma^2(\mathbf{r})\Theta(-h_\gamma(\mathbf{r})) - c_\gamma(\mathbf{r}) - \frac{1}{2}h_\gamma(\mathbf{r})c_\gamma(\mathbf{r})\right]$$

$$\bar{V} = k_B T \chi_T\left(1 - \sum_\gamma \rho_\gamma \int_V \mathrm{d}\mathbf{r}\, c_\gamma(\mathbf{r})\right)$$

Partial molar volume

$$\rho k_B T\,\chi_T = \left(1 - 4\pi \sum_{\alpha\gamma} \rho_\alpha \int_0^\infty r^2 c_{\alpha\gamma}(r)\,dr\right)^{-1}$$

Solvent compressibility

PMV-correlated Universal Correction (UC) to the KH solvation free energy

Palmer, D. S.; Frolov, A. I.; Ratkova, E. L.; Fedorov, M. V.
J. Phys. Condens. Matter **2010**, *22*, 492101

$$\Delta\bar{G} = \Delta G + \alpha(\rho\bar{V}) + \beta$$

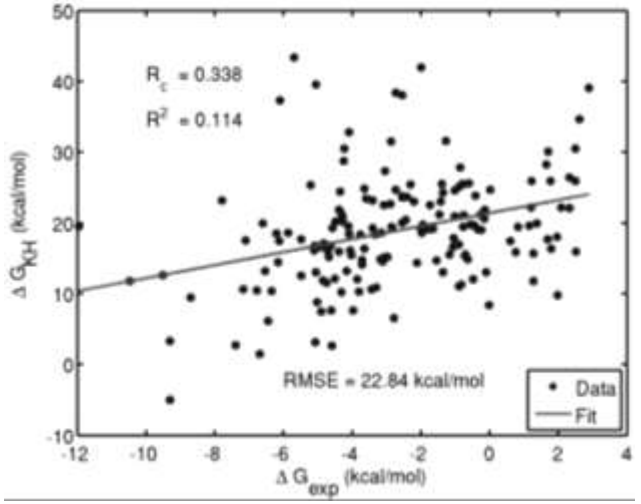
MAX-PLANCK-GESELLSCHAFT

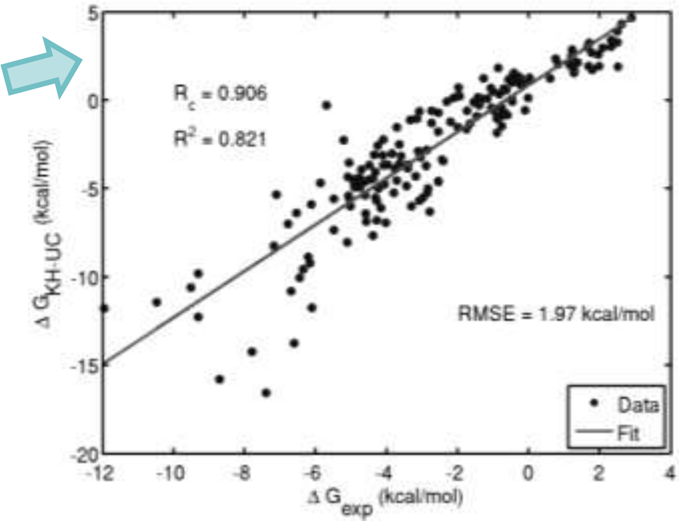
Cavity Correction (CC) to the KH solvation free energy

Truchon, J.-F.; Pettitt, B.M.; Labute, P.
J. Chem. Theory Comput. **2014**, *10*, 934

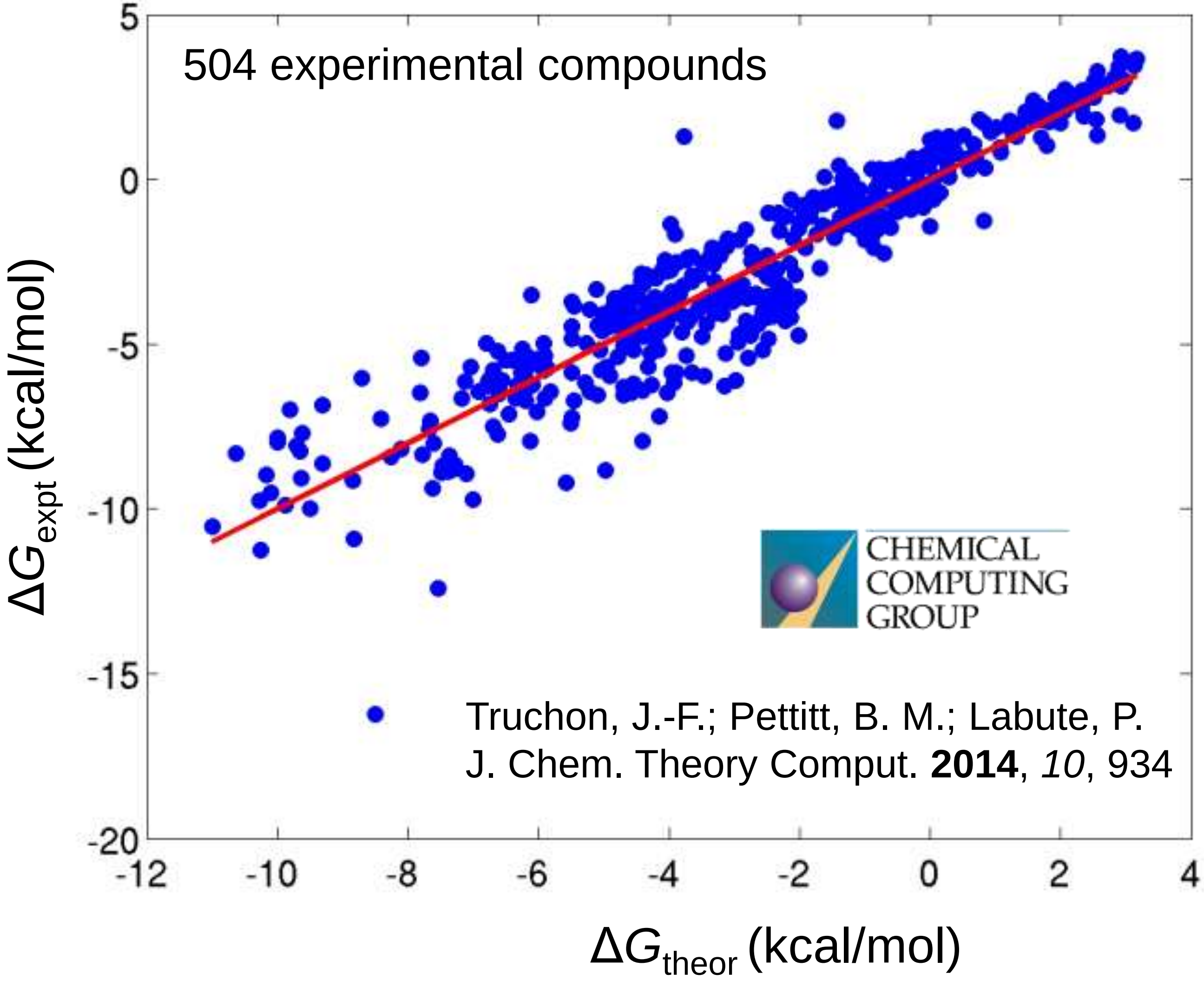
$$\Delta\bar{G} = \Delta G + \tfrac{1}{2}\rho_0 k_\mathrm{B} T(1-\gamma)\int_V \mathrm{d}\mathbf{r}\, c_0^{\mathrm{np}}(\mathbf{r})$$

CHEMICAL
COMPUTING
GROUP





Solvation free energy: experiment vs 3D-RISM-KH / KH-CC



Truchon, J.-F.; Pettitt, B. M.; Labute, P.
J. Chem. Theory Comput. **2014**, *10*, 934

3D-RISM-KH with the Cavity Correction (CC) to the KH solvation free energy
Truchon, J.-F.; Pettitt, B. M.; Labute, P. *J. Chem. Theory Comput.* **2014**, *10*, 934

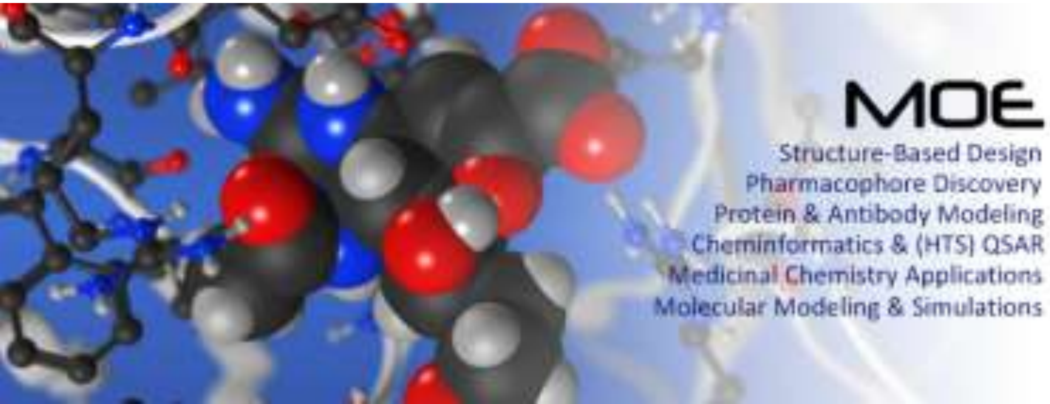
Detection and placement of structural solvent by 3D-RISM-KH molecular theory of solvation

Solvent analysis with 3D-RISM-KH molecular theory of solvation In Molecular Operating Environment (MOE) integrated drug discovery software package

- Calculate water density and binding desolvation penalty maps using 3D-RISM-KH theory
- Detect non-obvious hydrophobic regions of binding sites created by correlation and cavitation effects to prioritize ligand modifications.
- Orders of magnitude faster than alternative methods based on Molecular Dynamics or Monte Carlo simulations

Virtual screening of large libraries of fragment-/drug-like compounds

- Selection of receptors models
- Preparing library of compounds
- *In silico* screening experiments



Fragment-based drug design, interpretation of the experimental data, drug optimization for improving efficacy

- Analysis of conformational population of receptors (MD and clustering analysis)
- Prediction of structural solvent for docking experiments
- Flexible docking
- Pharmacophore Discovery
- Optimization of compounds for efficient delivery

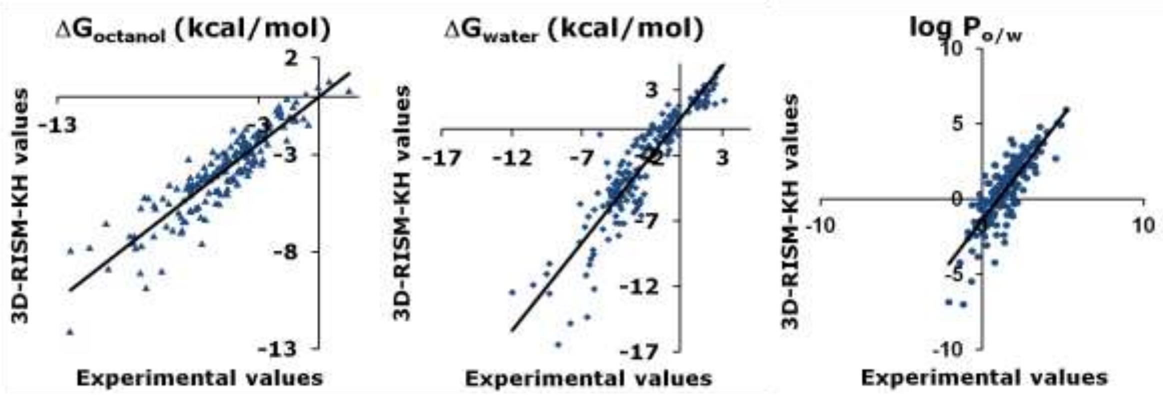
CHEMICAL
COMPUTING
GROUP

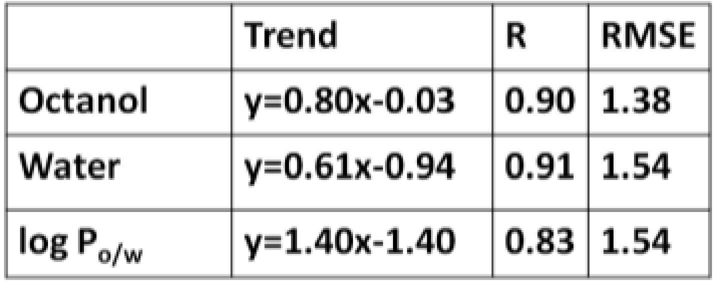
3D-RISM-KH / KH-UC approach to partitioning at liquid interfaces and biomembranes

Octanol–Water partition coefficient for compound translocation across biomembranes

Huang, W.-J.; Blinov, N.; Kovalenko, A. *J. Phys. Chem. B* **2015**, *119,* 5588.

Water–Octanol partitioning: 3D-RISM-KH / KH-UC theory vs experiment



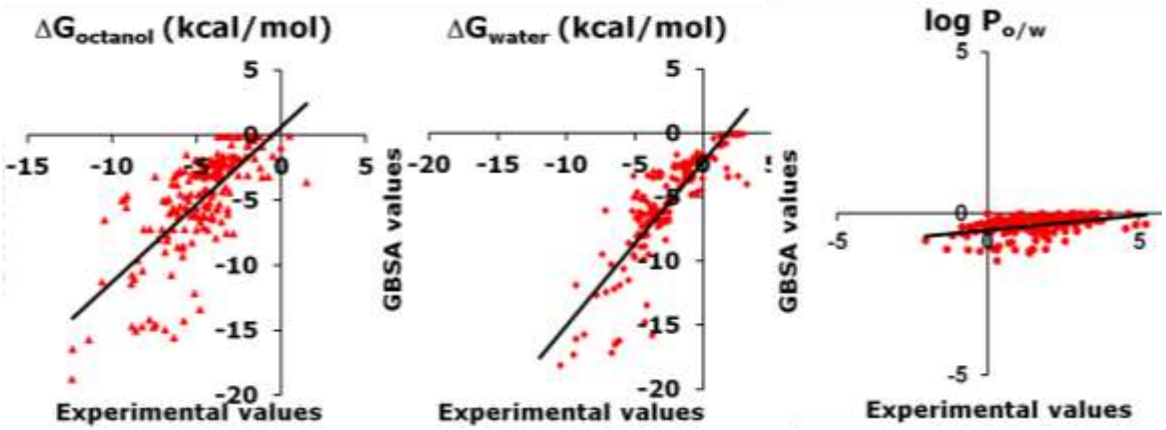


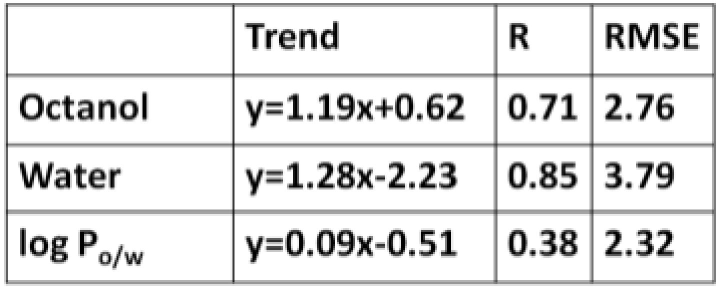
	Trend	R	RMSE
Octanol	y=0.80x-0.03	0.90	1.38
Water	y=0.61x-0.94	0.91	1.54
log $P_{o/w}$	y=1.40x-1.40	0.83	1.54

Trend: Equation; R: Correlation coefficients;
RMSE: Root mean square errors

$$\log P_{\mathrm{O/W}} = \left(\Delta G_{\mathrm{W}} - \Delta G_{\mathrm{O}}\right)/\left(k_{\mathrm{B}}T\right)$$

Water–Octanol partitioning: GBSA continuum solvation model vs experiment





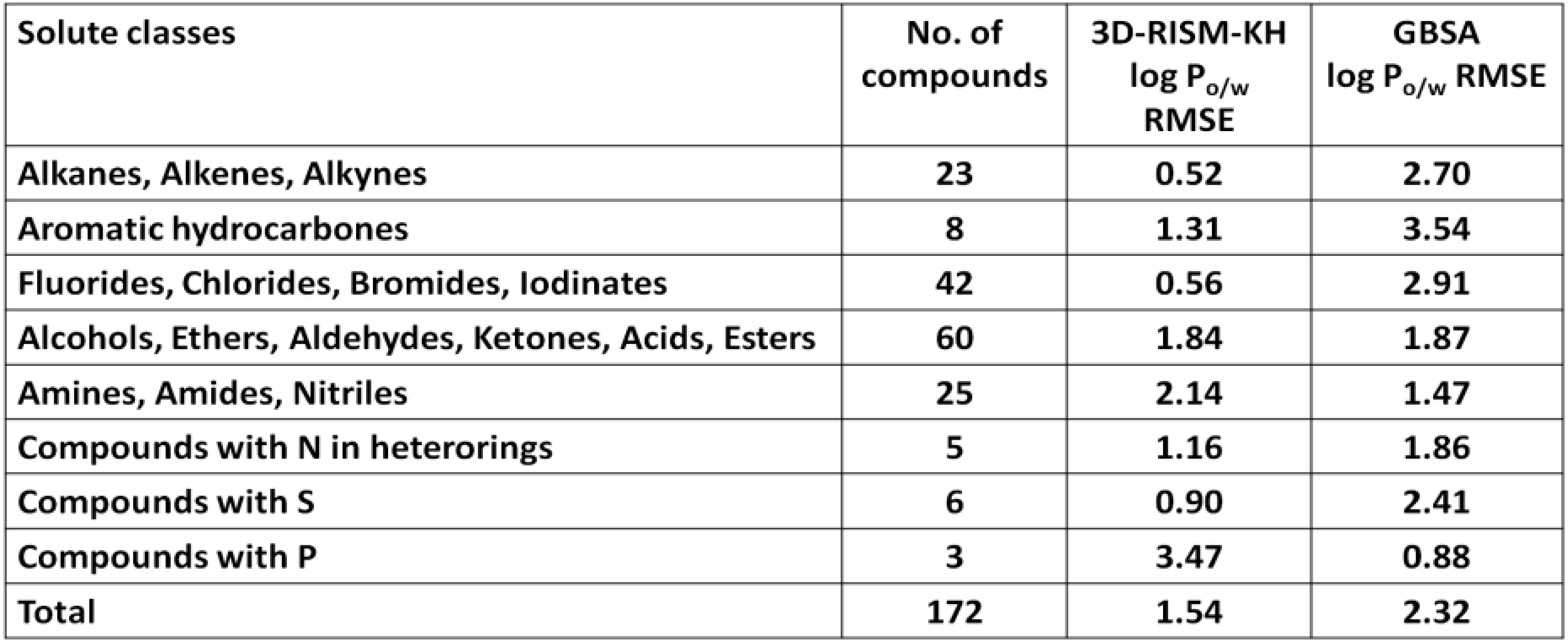
	Trend	R	RMSE
Octanol	y=1.19x+0.62	0.71	2.76
Water	y=1.28x-2.23	0.85	3.79
log $P_{o/w}$	y=0.09x-0.51	0.38	2.32

Octanol–Water partition coefficient for compound translocation across biomembranes

Huang, W.-J.; Blinov, N.; Kovalenko, A. *J. Phys. Chem. B* **2015**, *119,* 5588.

3D-RISM-KH / KH-UC theory:

- Accurate: approaching accuracy of empirical multi-parameter models
- Transferable to a given solvent system and solute molecule – as distinct from empirical models
- Twice as accurate compared to the GBSA implicit solvation model widely used in drug design

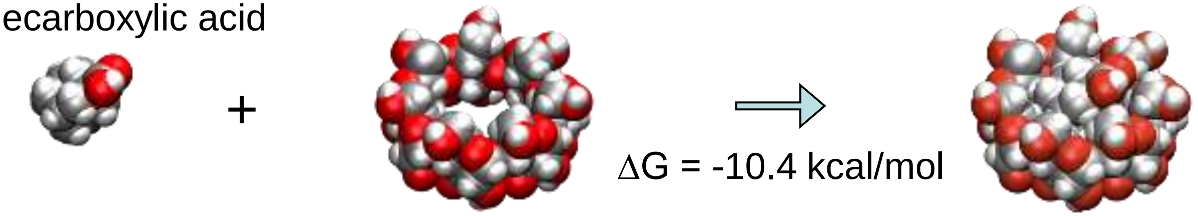


Solute classes	No. of compounds	3D-RISM-KH log $P_{o/w}$ RMSE	GBSA log $P_{o/w}$ RMSE
Alkanes, Alkenes, Alkynes	23	0.52	2.70
Aromatic hydrocarbones	8	1.31	3.54
Fluorides, Chlorides, Bromides, Iodinates	42	0.56	2.91
Alcohols, Ethers, Aldehydes, Ketones, Acids, Esters	60	1.84	1.87
Amines, Amides, Nitriles	25	2.14	1.47
Compounds with N in heterorings	5	1.16	1.86
Compounds with S	6	0.90	2.41
Compounds with P	3	3.47	0.88
Total	172	1.54	2.32

Spatial Decomposition Analysis (SDA) of solvation free energy, based on 3D-RISM-KH molecular theory of solvation

T. Yamazaki and A. Kovalenko, *J. Chem. Theory. Comput.*, **5**, 1723 (2009).
T. Yamazaki and A. Kovalenko,, *J. Phys. Chem. B,* **115**, 310-318 (2011).

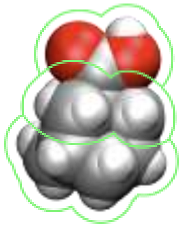
SDA of association thermodynamics of a cyclodextrin complex


1-adamantanecarboxylic acid

\+

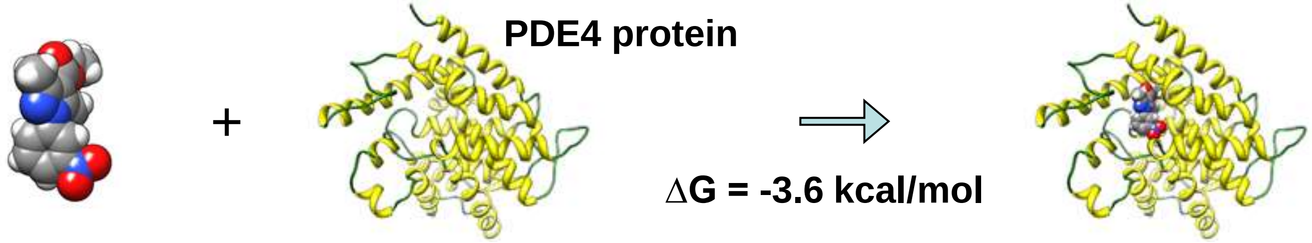
β-cyclodextrin

ΔG = -10.4 kcal/mol


ΔG_A = 0.9 kcal/mol
ΔG_B = -4.3 kcal/mol
ΔG_C = -4.9 kcal/mol

} adamantyl group contributes most to the association process.

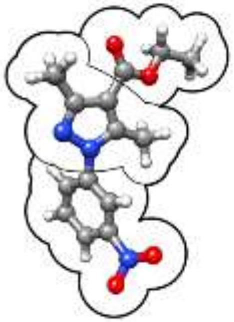
SDA of a protein-ligand complex


ligand

\+

PDE4 protein

ΔG = -3.6 kcal/mol


ΔG_A = –44 kcal/mol
ΔG_B = +76 kcal/mol
ΔG_C = –28 kcal/mol

→ fragment B is suggested for replacement

Writing
new equations?
It's cool !

Integral Equation Theory of Molecular Liquids:
Liquid-Liquid Interface

Lovett - Mou - Buff - Wertheim equation for atomic fluids

R. Lovett, C.Y.Mou, and F. P. Buff, J. Chem. Phys. **65**, 570 (1976)
M. S. Wertheim, J. Chem. Phys. **65**, 2377 (1976)

External field: $\gamma(\mathbf{r}) = \beta[\mu - u(\mathbf{r})] - 3\ln\Lambda$ where $\beta = 1/k_\mathrm{B}T$ and $\Lambda = h(2\pi m k_\mathrm{B}T)^{-1/2}$

Fluid density: $\rho(\mathbf{r}) = \rho(\mathbf{r},[\gamma])]$ and conversely $\gamma(\mathbf{r}) = \gamma(\mathbf{r},[\rho])]$

Functional derivatives:

$$\delta\rho(\mathbf{r}_1)/\delta\gamma(\mathbf{r}_2) = \rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2) - \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) + \rho(\mathbf{r}_1)\delta(\mathbf{r}_2 - \mathbf{r}_1)$$

$$\delta\gamma(\mathbf{r}_1)/\delta\rho(\mathbf{r}_2) = \delta(\mathbf{r}_2 - \mathbf{r}_1)/\rho(\mathbf{r}_1) - c(\mathbf{r}_1,\mathbf{r}_2)$$

Functional expansion for a small variation of the external field:

$$\rho(\mathbf{r}_1,[\gamma(\mathbf{r}+\delta)]) - \rho(\mathbf{r}_1,[\gamma(\mathbf{r})]) = \int \mathrm{d}\mathbf{r}_2 \delta\rho(\mathbf{r}_1)/\delta\gamma(\mathbf{r}_2)[\gamma(\mathbf{r}_2+\delta) - \gamma(\mathbf{r}_2)] + O(\delta\gamma)^2$$

In the limit of $\delta \to 0$:

$$\nabla_1\rho(\mathbf{r}_1) = \int \mathrm{d}\mathbf{r}_2 \frac{\delta\rho(\mathbf{r}_1)}{\delta\gamma(\mathbf{r}_2)} \nabla_2\gamma(\mathbf{r}_2) \quad \text{and} \quad \nabla_1\gamma(\mathbf{r}_1) = \int \mathrm{d}\mathbf{r}_2 \frac{\delta\gamma(\mathbf{r}_1)}{\delta\rho(\mathbf{r}_2)} \nabla_2\rho(\mathbf{r}_2)$$

Integral Equation Theory of Molecular Liquids: Liquid-Liquid Interface

Substitution of the functional derivatives $\delta\rho/\delta\gamma$ and $\delta\gamma/\delta\rho$:

$$\nabla_1\rho(\mathbf{r}_1) = -\beta\rho(\mathbf{r}_1)\nabla_1 u(\mathbf{r}_1) + \beta\int d\mathbf{r}_2[\rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2)-\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)]\nabla_2\gamma(\mathbf{r}_2)$$

Similar to BGY: $\nabla_1\rho(\mathbf{r}_1) = -\beta\rho(\mathbf{r}_1)\nabla_1 u(\mathbf{r}_1) - \beta\int d\mathbf{r}_2\rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2)\nabla_1\phi(r_{12})$

$$\nabla_1 \ln\rho(\mathbf{r}_1) = -\beta\nabla_1 u(\mathbf{r}_1) + \int d\mathbf{r}_2 c(\mathbf{r}_1,\mathbf{r}_2)\nabla_2\rho(\mathbf{r}_2)$$

– **LMBW equation**

$$\nabla_1 \ln\rho_a(\mathbf{r}_1) = -\beta\nabla_1 u_a(\mathbf{r}_1) + \int d\mathbf{r}_2 c_{ab}(\mathbf{r}_1,\mathbf{r}_2)\nabla_2\rho_b(\mathbf{r}_2)$$

– obvious generalization to a multicomponent system

Approximations for the *inhomogeneous two-body* direct correlation function:

If $c(\mathbf{r}_1,\mathbf{r}_2) = c(|\mathbf{r}_2-\mathbf{r}_1|)$ then LMBW turns into OZ/HNC:

$$\nabla_1 \ln\rho(\mathbf{r}_1) = -\beta\nabla_1 u(\mathbf{r}_1) + \int d\mathbf{r}_2 c(\mathbf{r}_1,\mathbf{r}_2)\nabla_2\rho(\mathbf{r}_2)$$

Construct: $c_{ab}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{2}\left[c_{ab}(|\mathbf{r}_2-\mathbf{r}_1|;\{\rho_c(\mathbf{r}_1)\}) + c_{ab}(|\mathbf{r}_2-\mathbf{r}_1|;\{\rho_c(\mathbf{r}_2)\})\right]$

S. Iatsevitch and F. Forstmann, J. Chem. Phys. **107**, 6925 (1997)

Site-site generalization of LMBW equation to molecular fluids

A. Kovalenko and F. Hirata, Phys. Chem. Chem. Phys. **7**, 1785 (2005)

Variation $\delta\gamma_\alpha / \delta\rho_\gamma$ similarly to atomic liquids gives :

$$\nabla_1\rho_\alpha(\mathbf{r}_1) = -\rho_\alpha(\mathbf{r}_1)\nabla_1\beta u(\mathbf{r}_1) + \rho_\alpha(\mathbf{r}_1)\sum_\gamma \int d\mathbf{r}_2 \tilde{c}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2;[\rho])\nabla_2\rho_\gamma(\mathbf{r}_2)$$

where *inhomogeneous* site-site DCF is defined in terms of susceptibility/density-density CF

$$\chi^{-1}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) = \rho^{-1}_\alpha(\mathbf{r}_1)\delta_{\alpha\gamma}\delta(\mathbf{r}_1-\mathbf{r}_2) - \tilde{c}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)$$

D. Chandler, J. D. McCoy, and S. J. Singer J. Chem. Phys. **85**, 5971; 5977 (1986)

The susceptibility can be subdivided into the *intra*- and *inter*molecular terms

$$\chi_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) = \underbrace{\chi^{(0)}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)}_{\text{intramolecular}} + \rho_\alpha(\mathbf{r}_1)\rho_\gamma(\mathbf{r}_2)h_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)$$

Approximation for the intramolecular part:

$$\chi^{(0)}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) = \rho_\alpha(\mathbf{r}_1)\omega_{\alpha\gamma}(|\mathbf{r}_1-\mathbf{r}_2|)$$

with the intramolecular matrix $\omega_{\alpha\gamma}(|\mathbf{r}_1-\mathbf{r}_2|) = \dfrac{1}{4\pi l^2_{\alpha\gamma}}\delta(|\mathbf{r}_1-\mathbf{r}_2| - l_{\alpha\gamma})$

SS-LMBW / RISM-KHM theory
for interfaces of molecular fluids

Kovalenko and Hirata, PCCP, 7, 1785 (2005).

Minimization of the Helmholtz free energy with respect to variations $\delta\rho_\alpha$ gives

$$\boldsymbol{\nabla}_1\,\rho_\alpha(\mathbf{r}_1) = -\rho_\alpha(\mathbf{r}_1)\boldsymbol{\nabla}_1\,\beta u_\alpha^{\mathrm{ext}}(\mathbf{r}_1)+\rho_\alpha(\mathbf{r}_1)\sum_\gamma\int \mathrm{d}\mathbf{r}_2\,\tilde{c}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2;[\rho])\boldsymbol{\nabla}_2\,\rho_\gamma(\mathbf{r}_2),$$

with the inhomogeneous site-site DCF $\tilde{c}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)$ defined in terms of the susceptibility

$$\chi^{-1}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) = \rho_\alpha^{-1}(\mathbf{r}_1)\delta_{\alpha\gamma}\delta(\mathbf{r}_1-\mathbf{r}_2) - \tilde{c}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2).$$

The *intramolecular* ideal gas part and the *intermolecular* term of the susceptibility,

$$\chi_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) = \chi^{(0)}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) + \rho_\alpha(\mathbf{r}_1)\rho_\gamma(\mathbf{r}_2)h_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2).$$

SS-LMBW / RISM-KHM theory

In the limit of a uniform fluid, the intra part is expressed through the intramolecular matrix,

$$\chi^{(0)}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)\Big|_{\text{uniform}} = \rho_\alpha \omega_{\alpha\gamma}(|\mathbf{r}_1-\mathbf{r}_2|).$$

Defining the *intramolecular* site-site DCF as

$$c^{(0)}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) = \rho_\alpha^{-1}(\mathbf{r}_1)\delta_{\alpha\gamma}\delta(\mathbf{r}_1-\mathbf{r}_2) - \left[\chi^{(0)}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)\right]^{-1},$$

the whole direct correlation function can be subdivided as

$$\tilde{c}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) = c^{(0)}_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) + c_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2),$$

where the latter term is identified as the *intermolecular* site-site direct correlation function. For a uniform fluid, $c_{\alpha\gamma}(|\mathbf{r}_1-\mathbf{r}_2|)$ satisfies the SSOZ equation of Chandler and Andersen.

SS-LMBW / RISM-KHM theory

The full SS-DCF $\tilde{c}_{\alpha\gamma}(\mathbf{r}_1, \mathbf{r}_2) = \delta_{\alpha\gamma}(\mathbf{r}_1, \mathbf{r}_2)\, \rho_\gamma^{-1}(\mathbf{r}_1) - \omega_{\alpha\gamma}^{-1}(\mathbf{r}_1, \mathbf{r}_2)\, \rho_\gamma^{-1}(\mathbf{r}_1) + c_{\alpha\gamma}(\mathbf{r}_1, \mathbf{r}_2)$ leads to

$$\boldsymbol{\nabla}_1\, \rho_\alpha(\mathbf{r}_1) = \omega_{\alpha\gamma}(r_{12}) * \Big(-\rho_\gamma(\mathbf{r}_2)\, \boldsymbol{\nabla}_2\, \beta u_\gamma^{\mathrm{ext}}(\mathbf{r}_2) + \rho_\gamma(\mathbf{r}_2)\, c_{\gamma\mu}(\mathbf{r}_2, \mathbf{r}_3) * \boldsymbol{\nabla}_3\, \rho_\mu(\mathbf{r}_3) \Big).$$

Artifact of uniform $\chi_{\alpha\gamma}^{(0)}(\mathbf{r}_1, \mathbf{r}_2)$: nonphysical regions of $\rho_\alpha(\mathbf{r}) < 0$ at the gas side.

Approximation to ensure consistency:

$$\boldsymbol{\nabla}_1\, \ln \rho_\alpha(\mathbf{r}_1) = \omega_{\alpha\gamma}(r_{12}) * \Big(-\boldsymbol{\nabla}_2\, \beta u_\gamma^{\mathrm{ext}}(\mathbf{r}_2) + c_{\gamma\mu}(\mathbf{r}_2, \mathbf{r}_3) * \boldsymbol{\nabla}_3\, \rho_\mu(\mathbf{r}_3) \Big).$$

This is our *new SS-LMBW equation.*

For comparison, the inhomogeneous RISM approach by Chandler, Singer, and McCoy

$$\rho_\alpha(\mathbf{r}_1) = \prod_\gamma \omega_{\alpha\gamma}(r_{12}) * \exp\Big(-\beta u_\gamma^{\mathrm{ext}}(\mathbf{r}_2) + \sum_\gamma c_{\gamma\mu}^{\mathrm{blk}}(r_{23}) * (\rho_\mu(\mathbf{r}_3) - \rho_\mu^{\mathrm{blk}}) \Big).$$

SS-LMBW / RISM-KHM theory

Constructing the inhomogeneous site-site DCFs:

$$c_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)=\frac{1}{2}\Big[\tilde{c}_{\alpha\gamma}\big(r_{12};\overline{\boldsymbol{\rho}}(\mathbf{r}_1)\big)+\tilde{c}_{\alpha\gamma}\big(r_{12};\overline{\boldsymbol{\rho}}(\mathbf{r}_2)\big)\Big]$$

where the effective homogeneous DCFs $\tilde{c}_{\alpha\gamma}(r;\overline{\boldsymbol{\rho}})$ are **nonlinearly** interpolated between the bulk DCFs of liquid phases 1 and 2, and a gas phase 0 with all species at a vanishing density,

$$\tilde{c}_{\alpha\gamma}(r;\boldsymbol{\rho})=\sum_{i=0}^{2}\tilde{w}_i(\boldsymbol{\rho})\,c_{\alpha\gamma}^{\mathrm{blk}}(r;\boldsymbol{\rho}^{(i)}),$$

with the interpolation weights obtained by projection and distortion

$$w_i(\boldsymbol{\rho})=\frac{(\boldsymbol{\rho}-\boldsymbol{\rho}^{(0)})\cdot(\boldsymbol{\rho}^{(i)}-\boldsymbol{\rho}^{(0)})}{(\boldsymbol{\rho}^{(i)}-\boldsymbol{\rho}^{(0)})\cdot(\boldsymbol{\rho}^{(i)}-\boldsymbol{\rho}^{(0)})}\qquad\text{for}\quad i=1,2$$

$$\tilde{w}_i=w_i\big(f_i w_i(w_i-1)^2+1\big)\quad\text{for}\quad i=1,2$$

$$\tilde{w}_0=1-\tilde{w}_1-\tilde{w}_2.$$

The distortion $\tilde{w}_i-w_i$ quadratically vanishes with $w_i\to 1$ or $w_i\to 0$ in the bulk phases.

SS-LMBW / RISM-KHM theory

The "distorting" factors f_i for $i = 1, 2$ are determined from the boundary conditions

$$\boldsymbol{\rho}_{\mathrm{I}}(\mathbf{r}) \to \boldsymbol{\rho}_{\mathrm{I}}^{\mathrm{rich}} \ \text{ as } \ \mathbf{r} \to -\infty \,,$$
$$\boldsymbol{\rho}_{\mathrm{II}}(\mathbf{r}) \to \boldsymbol{\rho}_{\mathrm{II}}^{\mathrm{rich}} \ \text{ as } \ \mathbf{r} \to +\infty \,.$$

Eigenvalues of the SS-LMBW problem:

$$\begin{aligned} \boldsymbol{\rho}_{\mathrm{I}}(+\infty) \ &: \qquad \boldsymbol{\rho}_{\mathrm{I}}(\mathbf{r}) \to \boldsymbol{\rho}_{\mathrm{I}}^{\mathrm{poor}} \ \text{ as } \ \mathbf{r} \to +\infty \,, \\ \boldsymbol{\rho}_{\mathrm{II}}(-\infty) \ &: \qquad \boldsymbol{\rho}_{\mathrm{II}}(\mathbf{r}) \to \boldsymbol{\rho}_{\mathrm{II}}^{\mathrm{poor}} \ \text{ as } \ \mathbf{r} \to -\infty \,. \end{aligned}$$

Density normalization constant:

$$\rho_\alpha^{\mathrm{av}} = V^{-1} \int_V \mathrm{d}\mathbf{r}\, \rho_\alpha(\mathbf{r})$$

The coarse-grained local density

$$\overline{\rho}_\alpha(\mathbf{r}) = \frac{3}{4\pi (r_\alpha^{\mathrm{w}})^3}\, \Theta\Big(r_\alpha^{\mathrm{w}} - |\mathbf{r} - \mathbf{r}'| \Big) * \rho_\alpha(\mathbf{r}') \,.$$

SS-LMBW / RISM-KHM theory

The SS-LMBW equation for a planar interface:

$$\frac{\mathrm{d}\ln\rho_\alpha(z_1)}{\mathrm{d}z_1} = \sum_\gamma \int_{-\infty}^{\infty}\mathrm{d}z_2\,\omega_{\alpha\gamma}(r_{12})\left(-\beta\frac{\mathrm{d}u_\gamma^{\mathrm{ext}}(z_2)}{\mathrm{d}z_2}+\frac{1}{2}\sum_{i=0}^{2}\right.$$
$$\times\left[\tilde{w}_i(\overline{\boldsymbol{\rho}}(z_2))\int_{-\infty}^{\infty}\mathrm{d}z_3\,c_{\alpha\gamma}^{\mathrm{blk}}\big(p=0,|z_2-z_3|;\boldsymbol{\rho}^{(i)}\big)\frac{\mathrm{d}\rho(z_3)}{\mathrm{d}z_3}\right.$$
$$\left.\left.+\int_{-\infty}^{\infty}\mathrm{d}z_3\,c_{\alpha\gamma}^{\mathrm{blk}}\big(p=0,|z_2-z_3|;\boldsymbol{\rho}^{(i)}\big)\,\tilde{w}_i(\overline{\boldsymbol{\rho}}(z_3))\frac{\mathrm{d}\rho(z_3)}{\mathrm{d}z_3}\right]\right)$$

with the zeroth planar harmonics of the DCFs

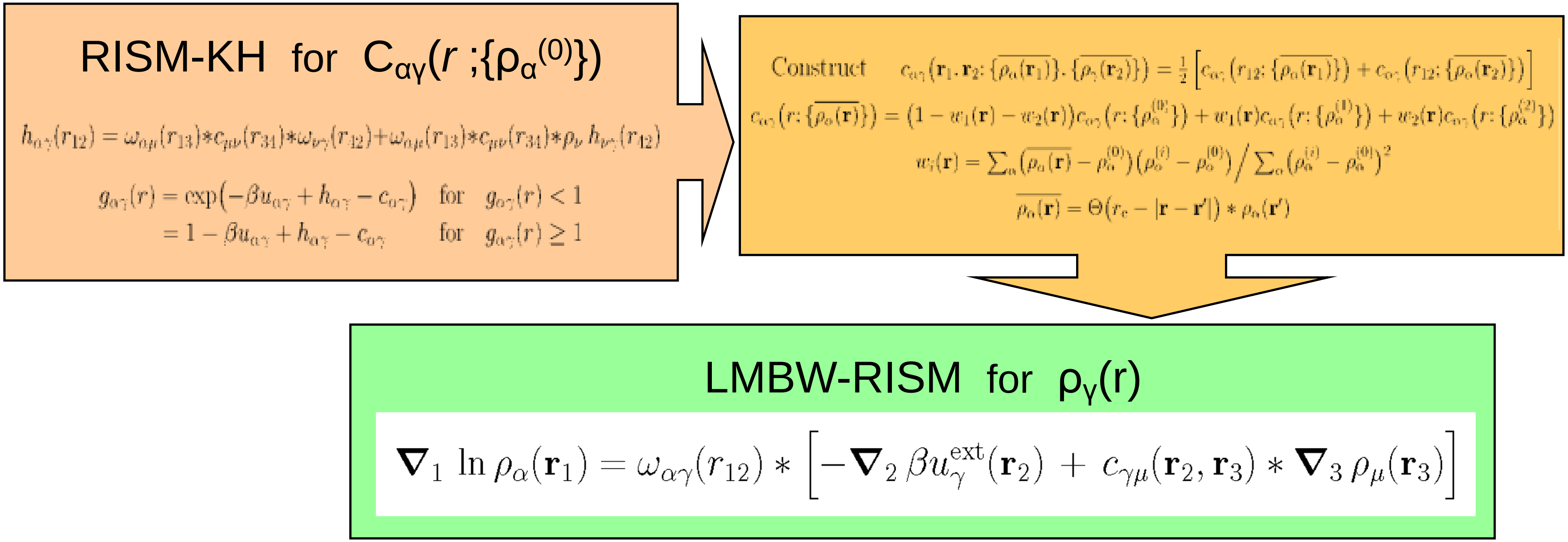
$$c_{\alpha\gamma}^{\mathrm{blk}}(p=0,z_1,z_2;\boldsymbol{\rho}^{(i)}) = 2\pi\int_0^\infty s\mathrm{d}s\,c_{\alpha\gamma}^{\mathrm{blk}}(s,z_1,z_2;\boldsymbol{\rho}^{(i)}).$$

SS-LMBW / RISM-KHM theory
for interfaces of molecular fluids

Kovalenko and Hirata, PCCP, 7, 1785 (2005).

- RISM integral equation with the KH closure
for the correlation functions in the coexisting bulk phases

- Site-site generalization of the Lovett-Mou-Buff-Wertheim equation
for the site density distributions at the interface



RISM-KH for $C_{\alpha\gamma}(r\,;\{\rho_\alpha^{(0)}\})$

$$h_{\alpha\gamma}(r_{12}) = \omega_{\alpha\mu}(r_{13}) * c_{\mu\nu}(r_{34}) * \omega_{\nu\gamma}(r_{42}) + \omega_{\alpha\mu}(r_{13}) * c_{\mu\nu}(r_{34}) * \rho_\nu\, h_{\nu\gamma}(r_{42})$$

$$\begin{aligned} g_{\alpha\gamma}(r) &= \exp\bigl(-\beta u_{\alpha\gamma} + h_{\alpha\gamma} - c_{\alpha\gamma}\bigr) && \text{for} \quad g_{\alpha\gamma}(r) < 1 \\ &= 1 - \beta u_{\alpha\gamma} + h_{\alpha\gamma} - c_{\alpha\gamma} && \text{for} \quad g_{\alpha\gamma}(r) \geq 1 \end{aligned}$$

Construct $c_{\alpha\gamma}(\mathbf{r}_1, \mathbf{r}_2; \{\overline{\rho_\alpha(\mathbf{r}_1)}\}, \{\overline{\rho_\gamma(\mathbf{r}_2)}\}) = \frac{1}{2}\left[c_{\alpha\gamma}(r_{12}; \{\overline{\rho_\alpha(\mathbf{r}_1)}\}) + c_{\alpha\gamma}(r_{12}; \{\overline{\rho_\alpha(\mathbf{r}_2)}\})\right]$

$$c_{\alpha\gamma}(r; \{\overline{\rho_\alpha(\mathbf{r})}\}) = (1 - w_1(\mathbf{r}) - w_2(\mathbf{r}))c_{\alpha\gamma}(r; \{\rho_\alpha^{(0)}\}) + w_1(\mathbf{r})c_{\alpha\gamma}(r; \{\rho_\alpha^{(1)}\}) + w_2(\mathbf{r})c_{\alpha\gamma}(r; \{\rho_\alpha^{(2)}\})$$

$$w_i(\mathbf{r}) = \textstyle\sum_\alpha \bigl(\overline{\rho_\alpha(\mathbf{r})} - \rho_\alpha^{(0)}\bigr)\bigl(\rho_\alpha^{(i)} - \rho_\alpha^{(0)}\bigr) \Big/ \sum_\alpha \bigl(\rho_\alpha^{(i)} - \rho_\alpha^{(0)}\bigr)^2$$

$$\overline{\rho_\alpha(\mathbf{r})} = \Theta(r_c - |\mathbf{r} - \mathbf{r}'|) * \rho_\alpha(\mathbf{r}')$$

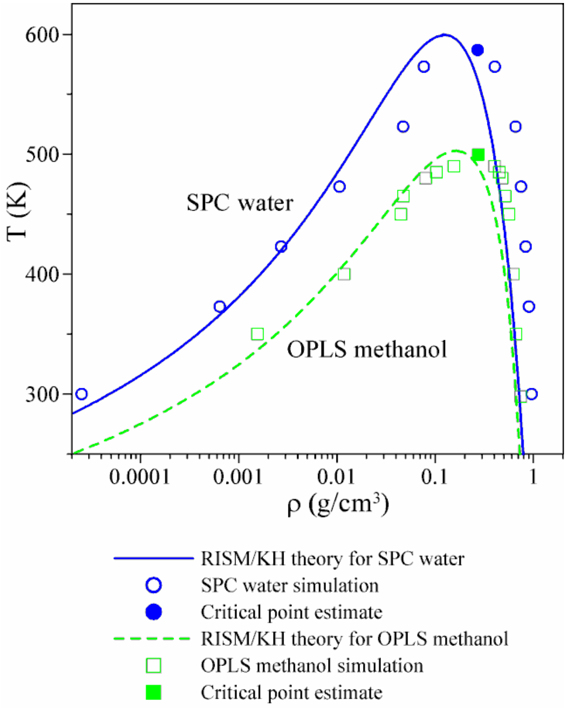
LMBW-RISM for $\rho_\gamma(r)$

$$\boldsymbol{\nabla}_1 \ln \rho_\alpha(\mathbf{r}_1) = \omega_{\alpha\gamma}(r_{12}) * \left[-\boldsymbol{\nabla}_2\, \beta u_\gamma^{\text{ext}}(\mathbf{r}_2) + c_{\gamma\mu}(\mathbf{r}_2, \mathbf{r}_3) * \boldsymbol{\nabla}_3\, \rho_\mu(\mathbf{r}_3)\right]$$

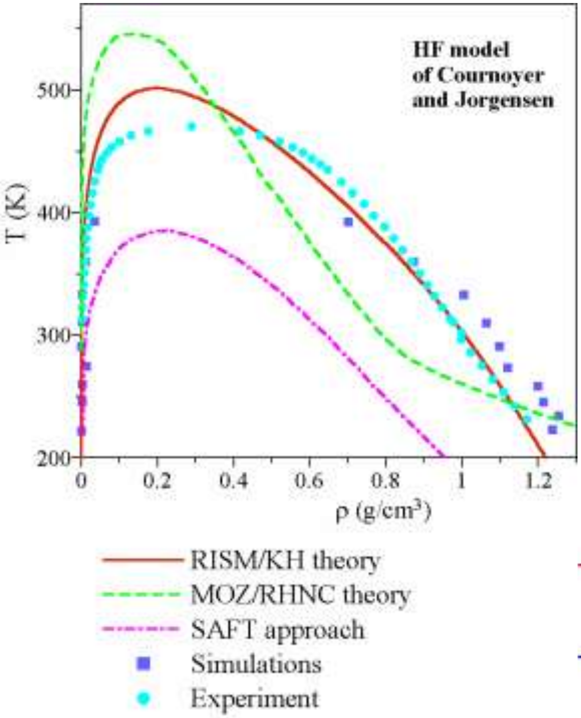
Liquid-vapor coexistence for water, methanol, hydrogen fluoride, and DMSO by the RISM-KH

Kovalenko and Hirata, JTCC, 1, 381 (2001)

Water vs Methanol

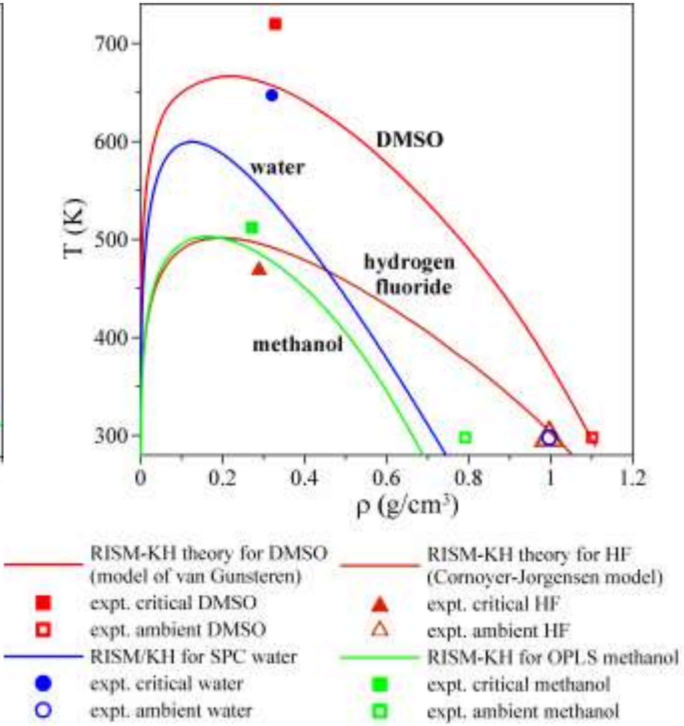


Hydrogen Fluoride



Theory vs Experiment

(critical and ambient conditions)



Liquid interfaces: SS-LMBW/RISM-KHM theory

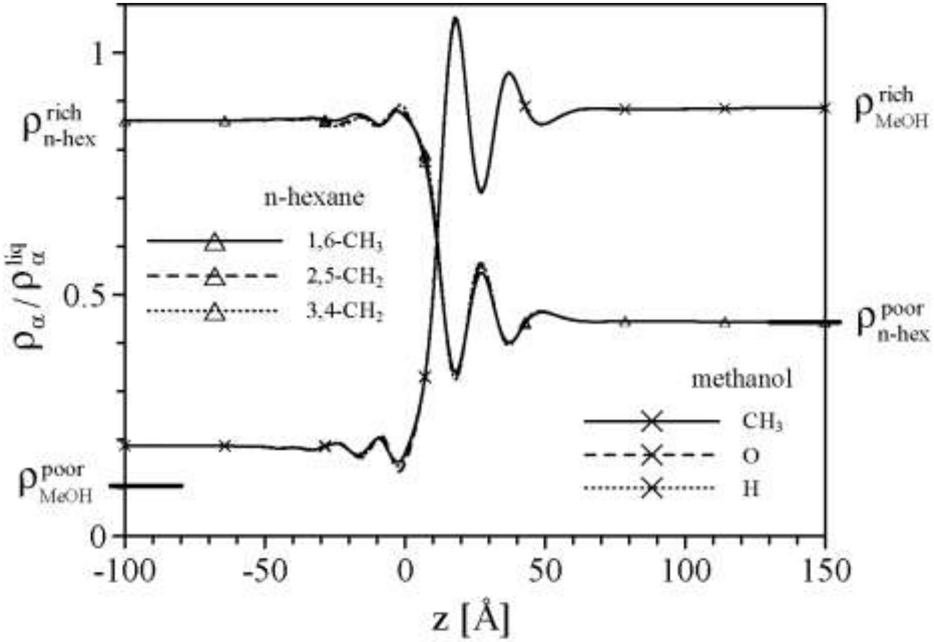
Kovalenko and Hirata, PCCP, 7, 1785 (2005).

RISM-KHM for $C_{\alpha\gamma}(r\,;\{\rho_\alpha^{(0)}\})$

construct $C_{\alpha\gamma}(r_1, r_2\,;\{\overline{\rho_\alpha(r_1)}\},\{\overline{\rho_\alpha(r_1)}\})$

SS-LMBW for $\rho_\gamma(r)$

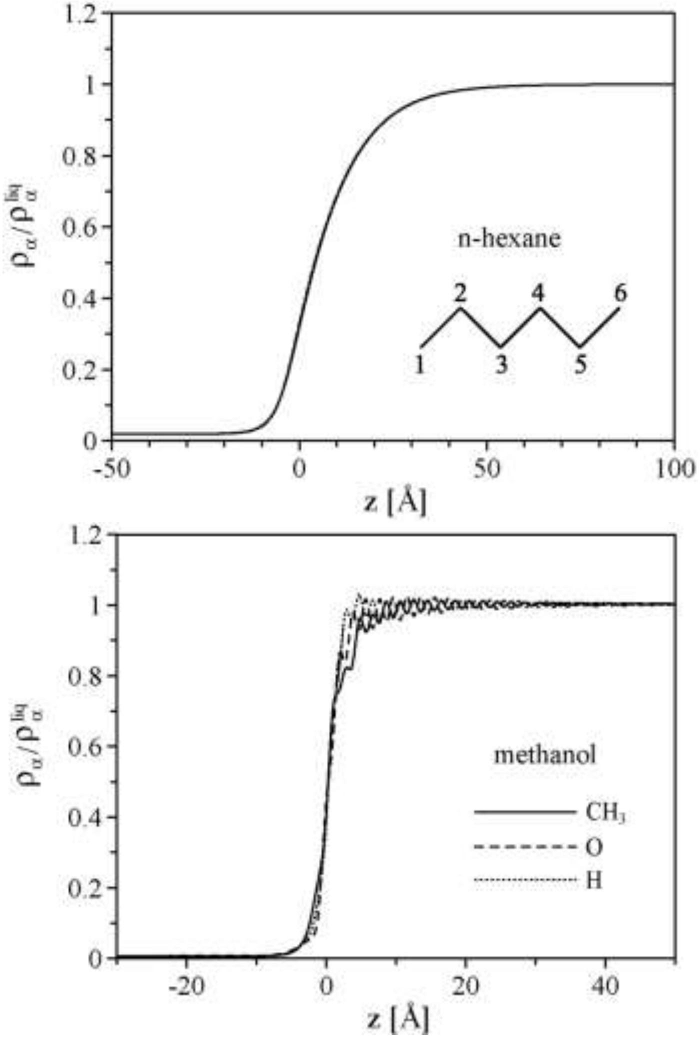
n-hexane rich phase



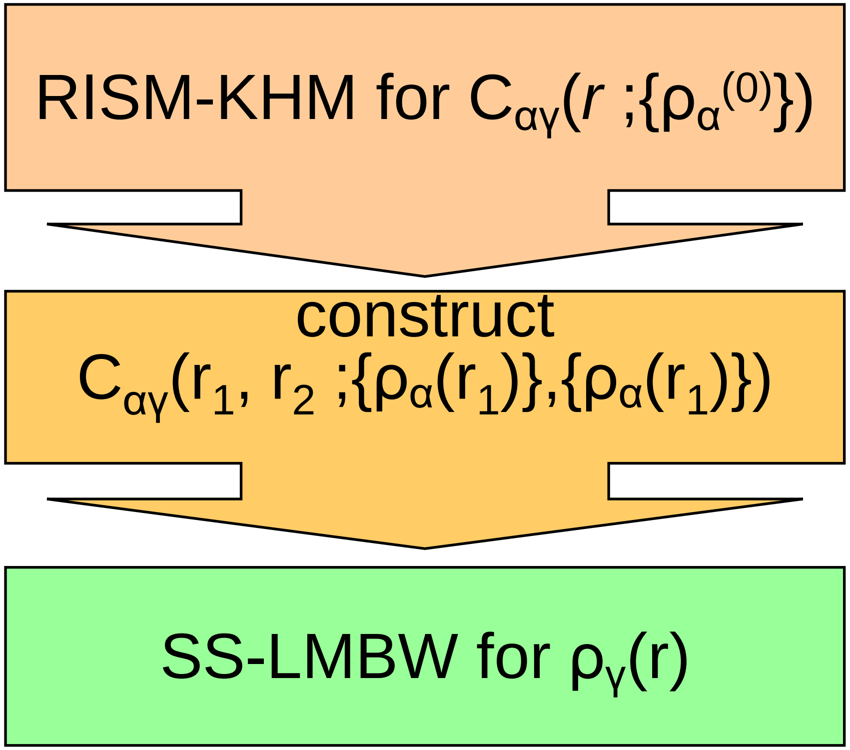
methanol rich phase

Liquid-liquid interface in the ambient solution of n-hexane and methanol (partially miscible)

Liquid-vapor interfaces



Liquid interfaces and soft matter structures

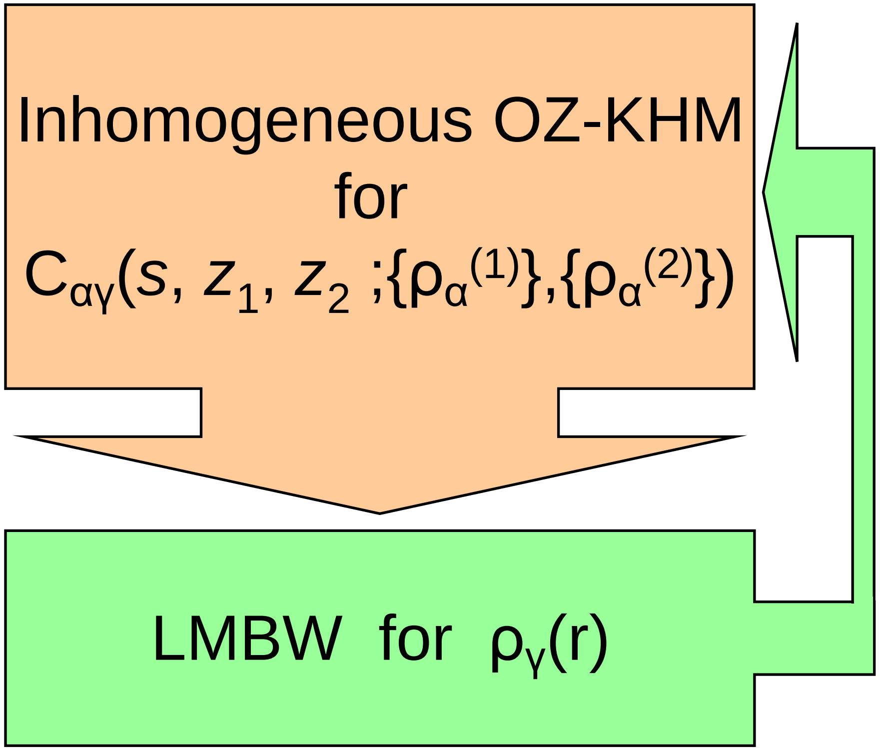


RISM-KHM for $C_{\alpha\gamma}(r\ ;\{\rho_\alpha^{(0)}\})$

construct
$C_{\alpha\gamma}(r_1, r_2\ ;\{\rho_\alpha(r_1)\},\{\rho_\alpha(r_1)\})$

SS-LMBW for $\rho_\gamma(r)$

Approximating
the inhomogeneous two-body
direct correlation functions



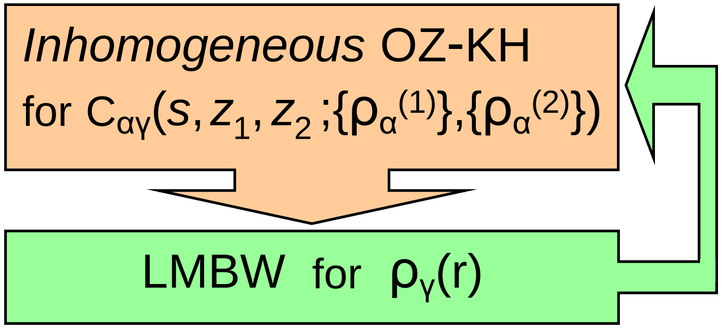
Inhomogeneous OZ-KHM
for
$C_{\alpha\gamma}(s, z_1, z_2\ ;\{\rho_\alpha^{(1)}\},\{\rho_\alpha^{(2)}\})$

LMBW for $\rho_\gamma(r)$

Deriving
the inhomogeneous two-body
direct correlation functions

Inhomogeneous integral equation theory of liquid interfaces

- **Self-consistent approach**



Inhomogeneous OZ-KH for $C_{\alpha\gamma}(s, z_1, z_2; \{\rho_\alpha^{(1)}\},\{\rho_\alpha^{(2)}\})$

LMBW for $\rho_\gamma(r)$

  - ***Inhomogeneous* OZ-KH equation for the three-particle correlations**

$$h_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) = c_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) + \int d\mathbf{r}_3\, c_{\alpha\mu}(\mathbf{r}_1,\mathbf{r}_3)\,\rho_\mu(\mathbf{r}_3)\,h_{\mu\gamma}(\mathbf{r}_3,\mathbf{r}_2)$$

$$\begin{aligned} g_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) &= \exp\big(-\beta u_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)+h_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)-c_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)\big) \quad \text{for} \quad g_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) < 1 \\ &= 1-\beta u_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)+h_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2)-c_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) \quad \text{for} \quad g_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) \geq 1 \end{aligned}$$

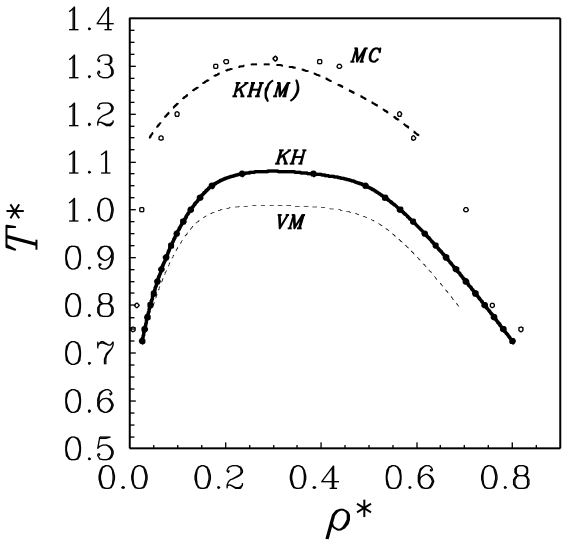
& **LMBW for the single-particle density distributions**

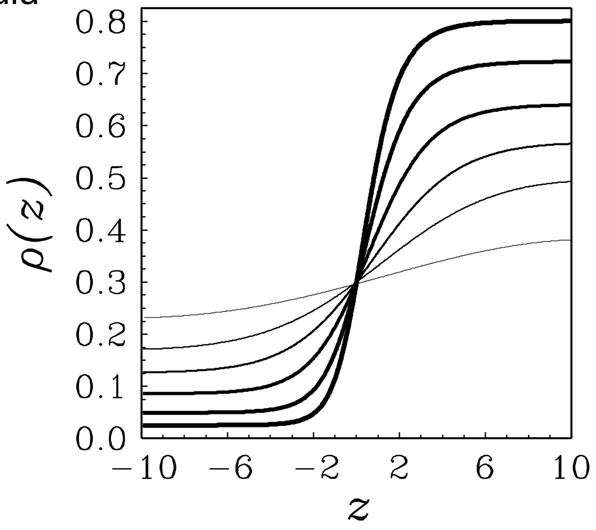
$$\boldsymbol{\nabla}_1\,\rho_\alpha(\mathbf{r}_1) = -\beta\boldsymbol{\nabla}_1\,u_\alpha(\mathbf{r}_1) + \rho_\alpha(\mathbf{r}_1)\,c_{\alpha\gamma}(\mathbf{r}_1,\mathbf{r}_2) * \boldsymbol{\nabla}_2\,\rho_\gamma(\mathbf{r}_2)$$

Liquid-Vapor interface of simple fluid by the Inhomogeneous OZ-KH/LMBW theory

Omelyan, Kovalenko, and Hirata, CPL 397, 368 (2004).

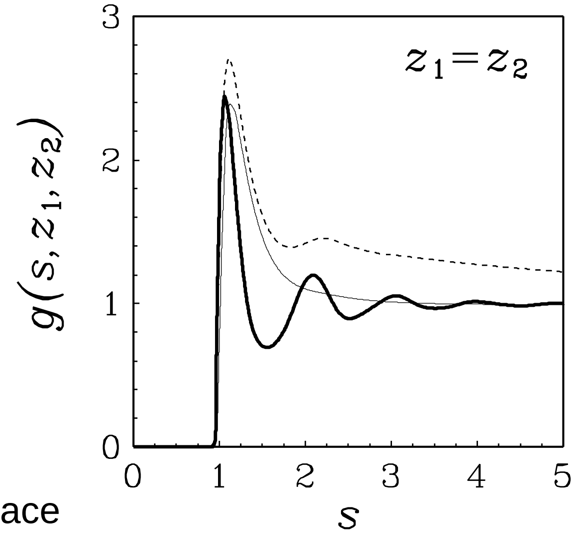
Vapor-liquid coexistence curve of LJ fluid





Density profiles at the vapor-liquid interface

Section of two-particle distributions along the vapor-liquid interface

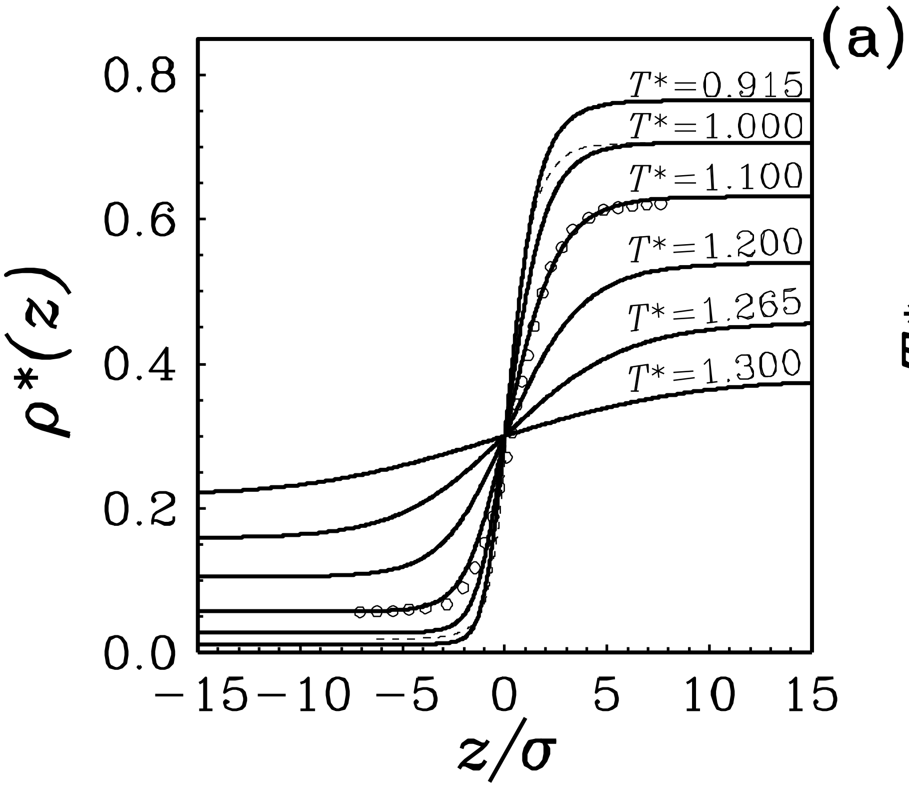


- Inhomogeneous Ornstein-Zernike integral equation for the two-particle correlations at an interface
- Kovalenko-Hirata closure (modified)
- Lovett-Mou-Buff-Wertheim equation for the density distributions at the interface

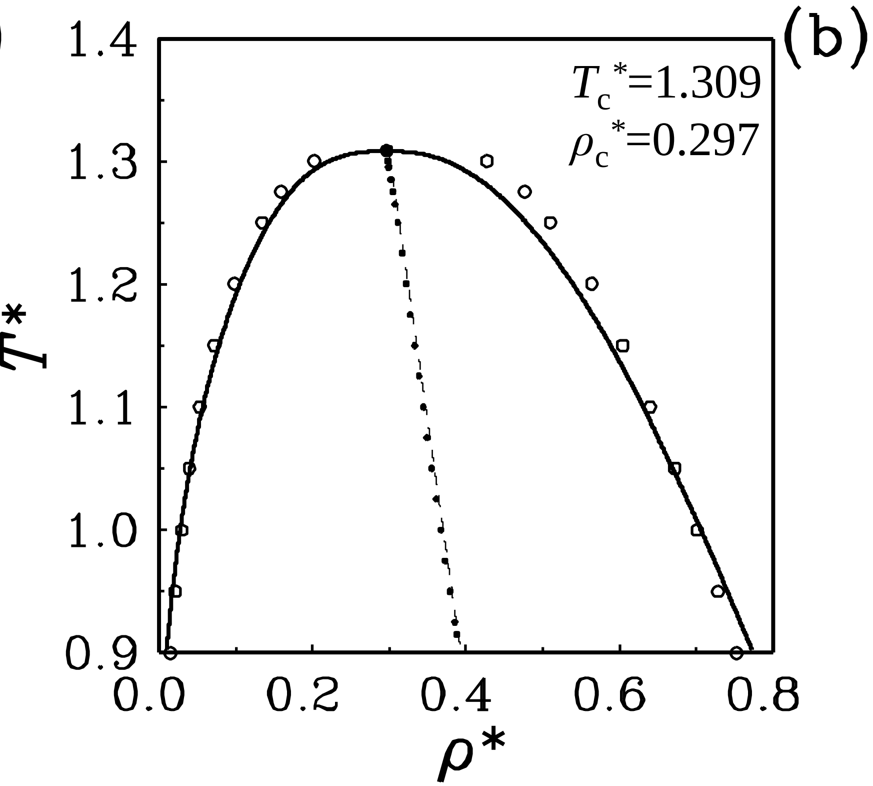
IOZ-KH/LMBW theory analytically predicts long-range lateral correlations at a liquid interface, similar to critical-regime fluctuations

Liquid-vapor interface of LJ fluid by IOZ-DH / LMBW

Omelyan, Kovalenko, and Hirata, PCCP



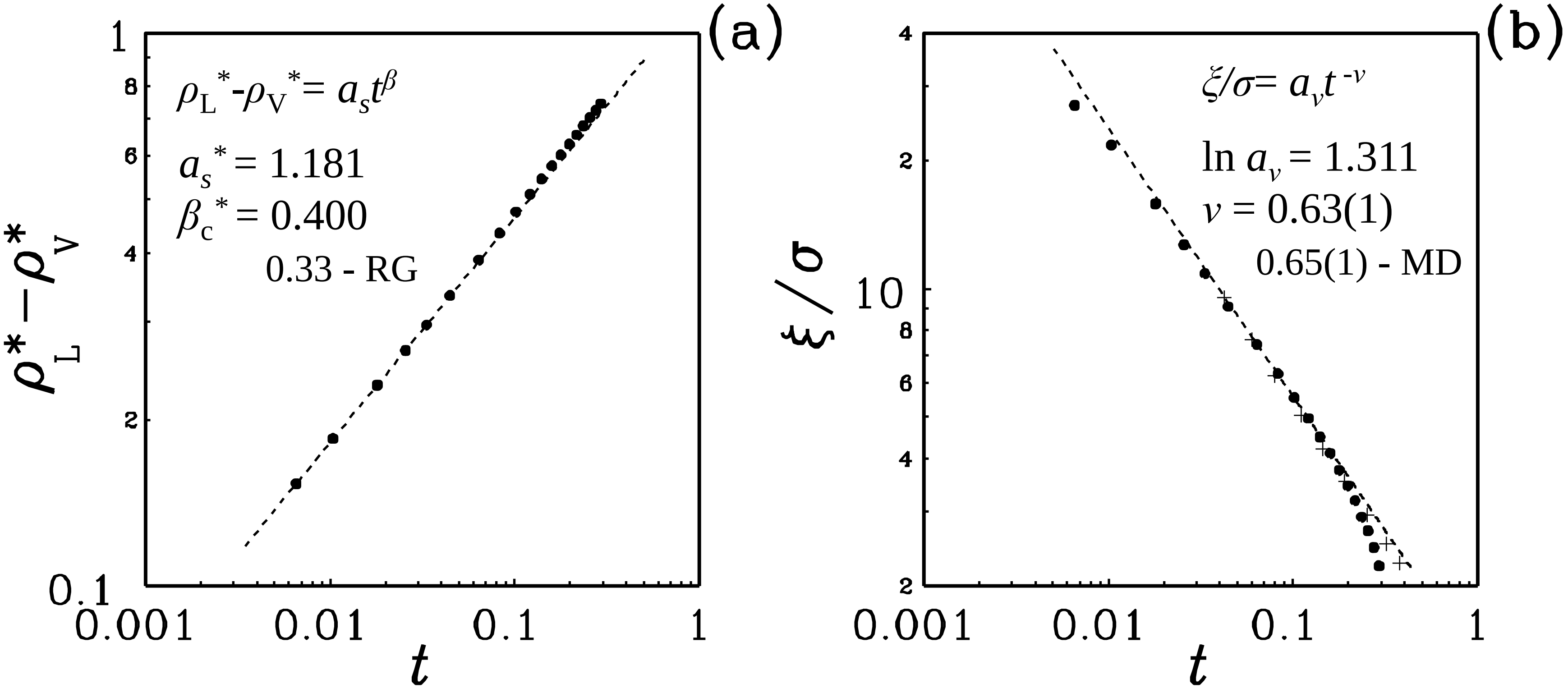
Intrinsic density profile



Binodal curve

Liquid-vapor interface of LJ fluid by IOZ-DH / LMBW

Omelyan, Kovalenko, and Hirata, PCCP 2007



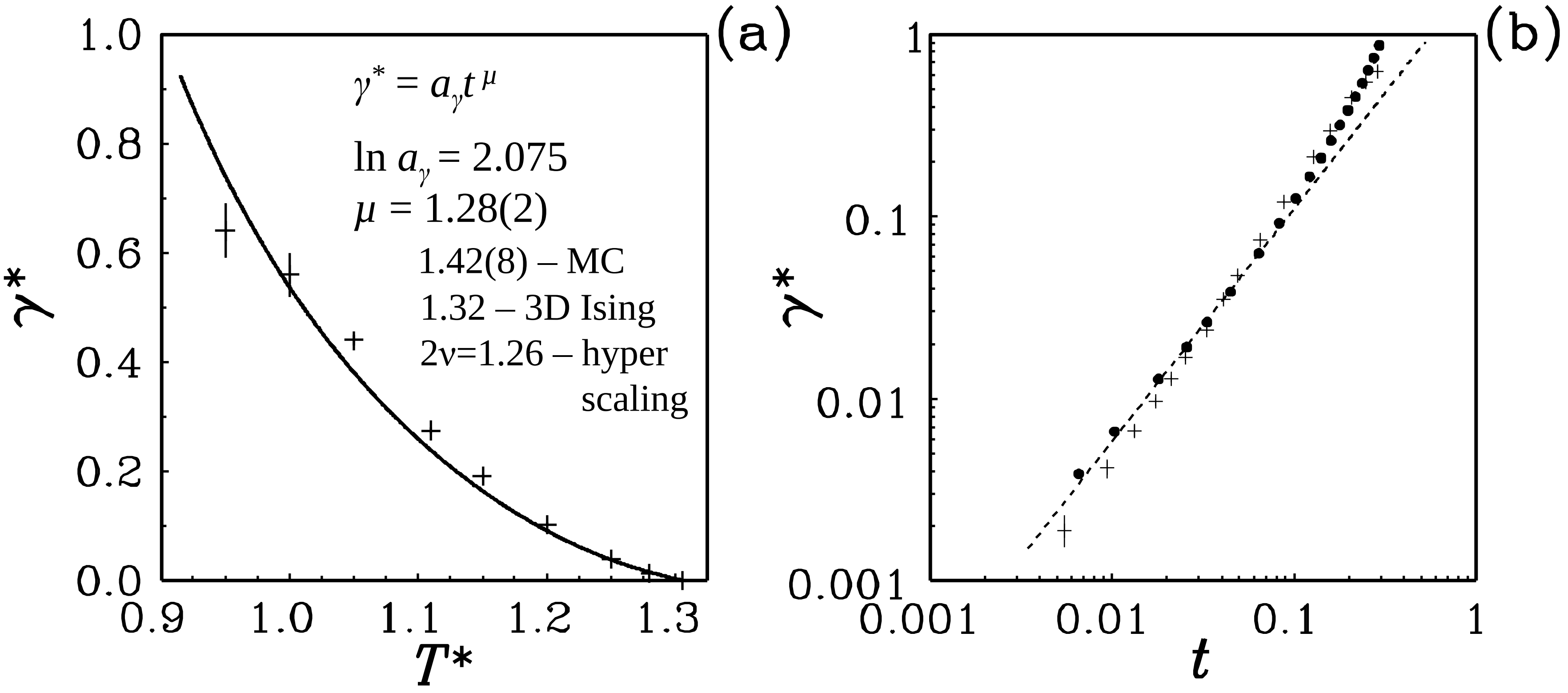
Binodal

Intrinsic interfacial width

Critical behavior (dashed lines) derived from the IOZ-DH / LMBW theory (dots). Simulation data (crosses).

Liquid-vapor interface of LJ fluid by IOZ-DH / LMBW

Omelyan, Kovalenko, and Hirata, PCCP 2007

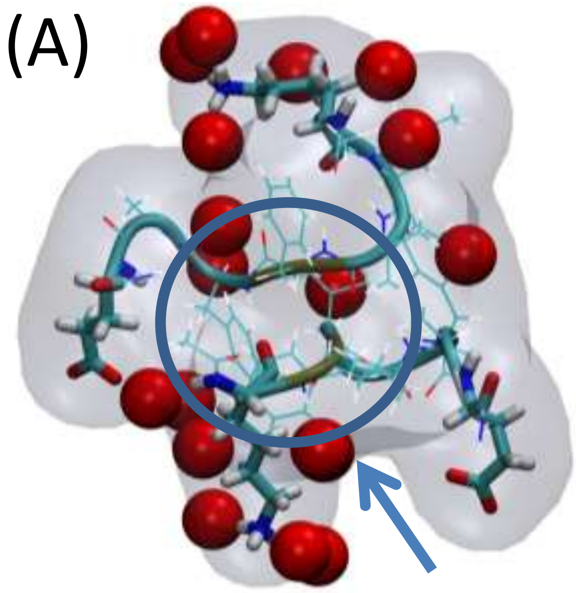


Critical behavior of the intrinsic surface tension (lines) derived from the IOZ-DH / LMBW theory (dots). Simulation data (crosses).

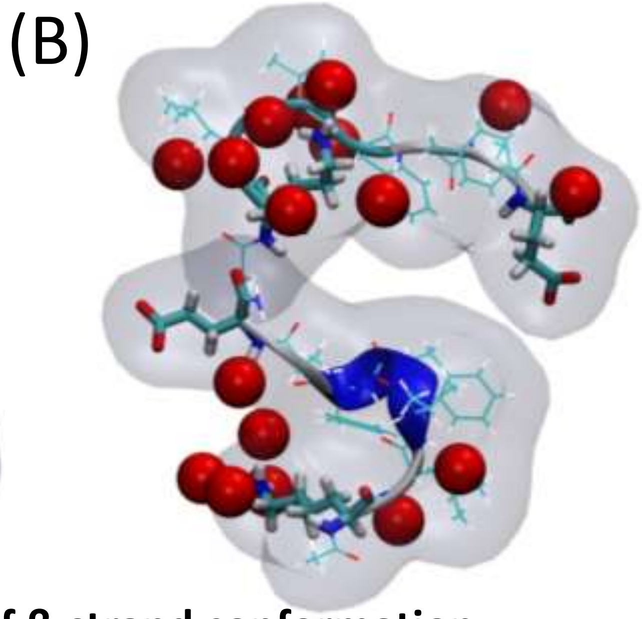
MM / 3D-RISM-KH post-processing replaces MM / GB(PB)SA

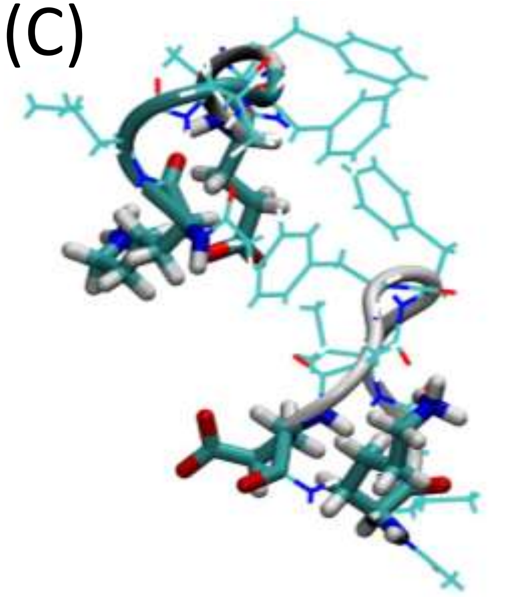
Structural characteristics of small Aβ oligomers

Dimeric intermediates in oligomerization pathways



Inception of β-strand conformation





(A) Hydrophobic and electrostatic (persistent) contacts. Dry interface between hydrophobic residues. Polar residues are partially solvated.

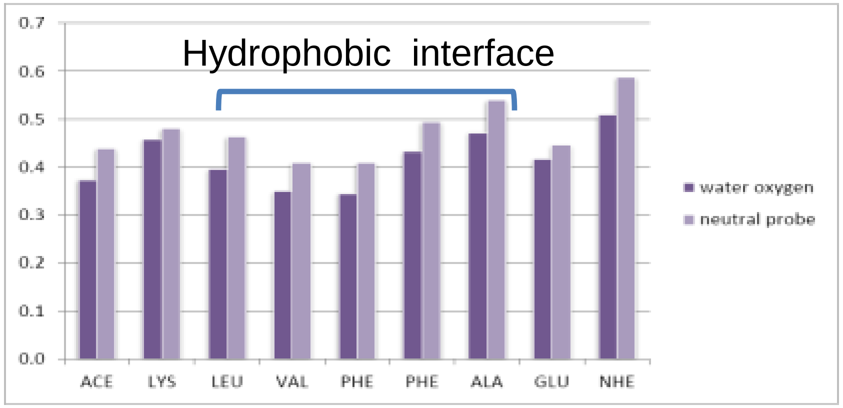
(B) Electrostatic (non-persistent) contacts. Hydrophobic and polar residues are partially solvated.

(C) Non-polar hydrophobic (non-persistent) contacts. Fast desolvation of hydrophobic residues, polar residues are solvated.

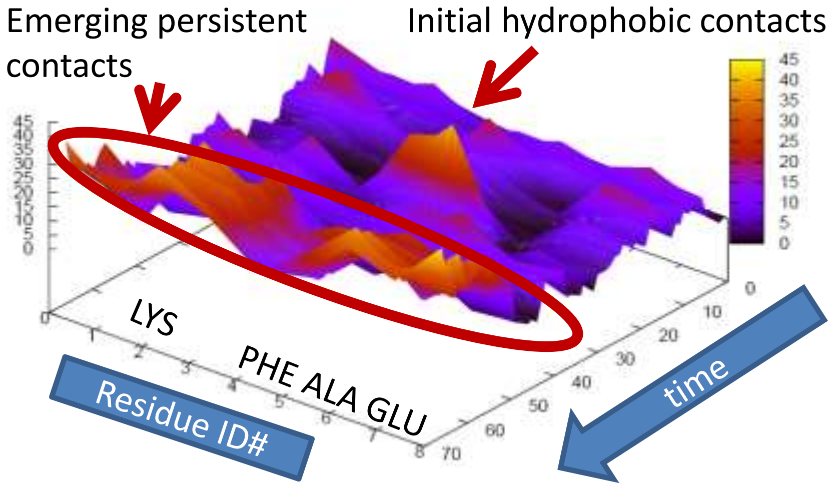
Initial contacts between Aβ peptides are hydrophobic, whereas persistent contacts include both hydrophobic and electrostatic interactions

3D-RISM-KH based descriptors for propensity for aggregation

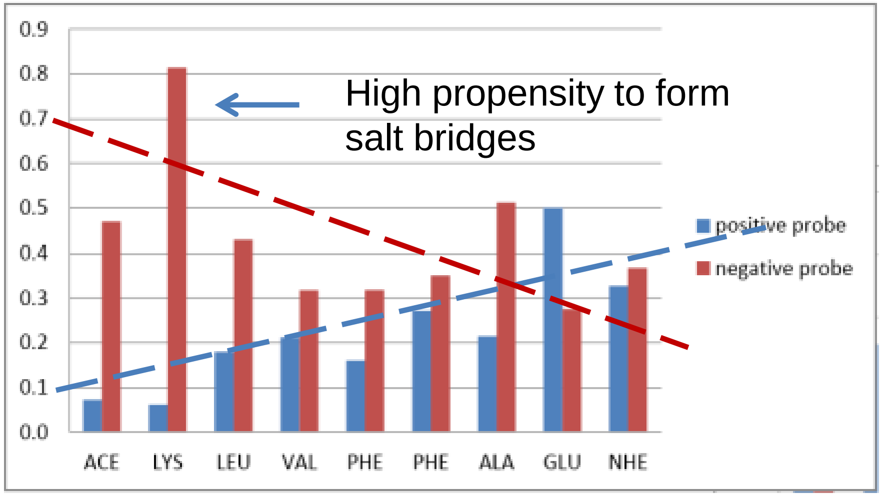
New protocol for prediction of propensity for aggregation of amyloidogenic peptides uses a probe particle number density and solvation free energy density as descriptors



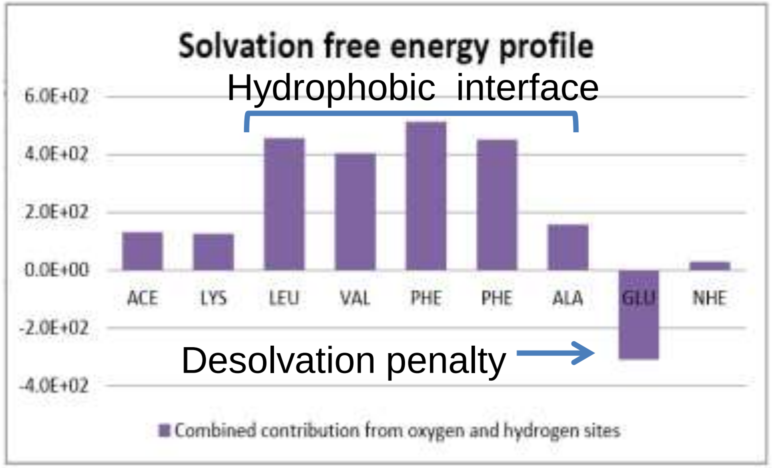
Neutral probe particles density profiles



Evolution of inter-peptide contacts in oligomerization



Charged probe particles density profiles



Solvation free energy density profiles

Initial contacts between Aβ 16-22 peptides are hydrophobic, while persistent contacts include both hydrophobic and electrostatic interactions

3D-RISM-KH based descriptors of propensity for aggregation: 3D density maps of solvation free energy and solvation entropy

Kovalenko-Hirata (KH) functional of the solvation free energy

$$\Delta G = \int_V d\mathbf{r}\, \Phi^{\mathrm{KH}}(\mathbf{r})$$

with the solvation free energy density

$$\Phi^{\mathrm{KH}}(\mathbf{r}) = \sum_\gamma \Phi_\gamma^{\mathrm{KH}}(\mathbf{r})$$

comprising solvent site contributions

$$\Phi_\gamma^{\mathrm{KH}}(\mathbf{r}) = \rho_\gamma k_{\mathrm{B}} T \left[\frac{1}{2} h_\gamma^2(\mathbf{r}) \Theta(-h_\gamma(\mathbf{r})) - c_\gamma(\mathbf{r}) - \frac{1}{2} h_\gamma(\mathbf{r}) c_\gamma(\mathbf{r}) \right]$$

Solvation entropy

$$\Delta S_V = -T^{-1} \left(\partial \Delta G / \partial T \right)_\rho = \int_V d\mathbf{r}\, \Xi_V^{\mathrm{KH}}(\mathbf{r})$$

with the solvation entropy density

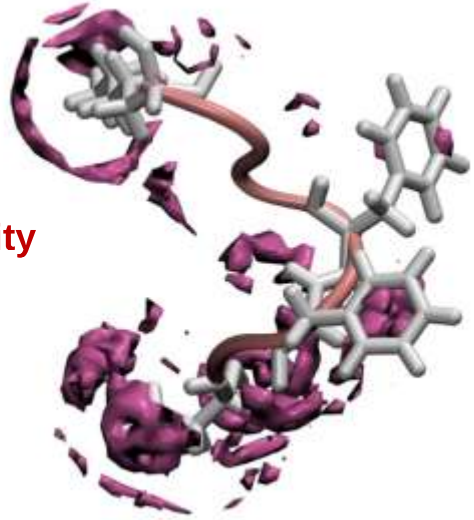
$$\Xi_V^{\mathrm{KH}}(\mathbf{r}) = -T^{-1} \left(\partial \Phi^{\mathrm{KH}}(\mathbf{r}) / \partial T \right)_\rho = -\sum_\gamma T^{-1} \left(\partial \Phi_\gamma^{\mathrm{KH}}(\mathbf{r}) / \partial T \right)_\rho$$

solvation free energy density

$$\Phi^{\mathrm{KH}}(\mathbf{r})$$

solvation entropy density

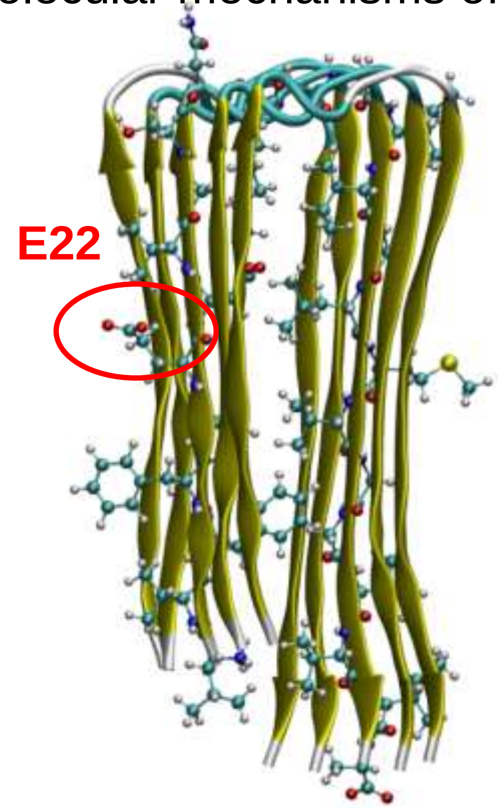
$$\Sigma^{\mathrm{KH}}(\mathbf{r})$$



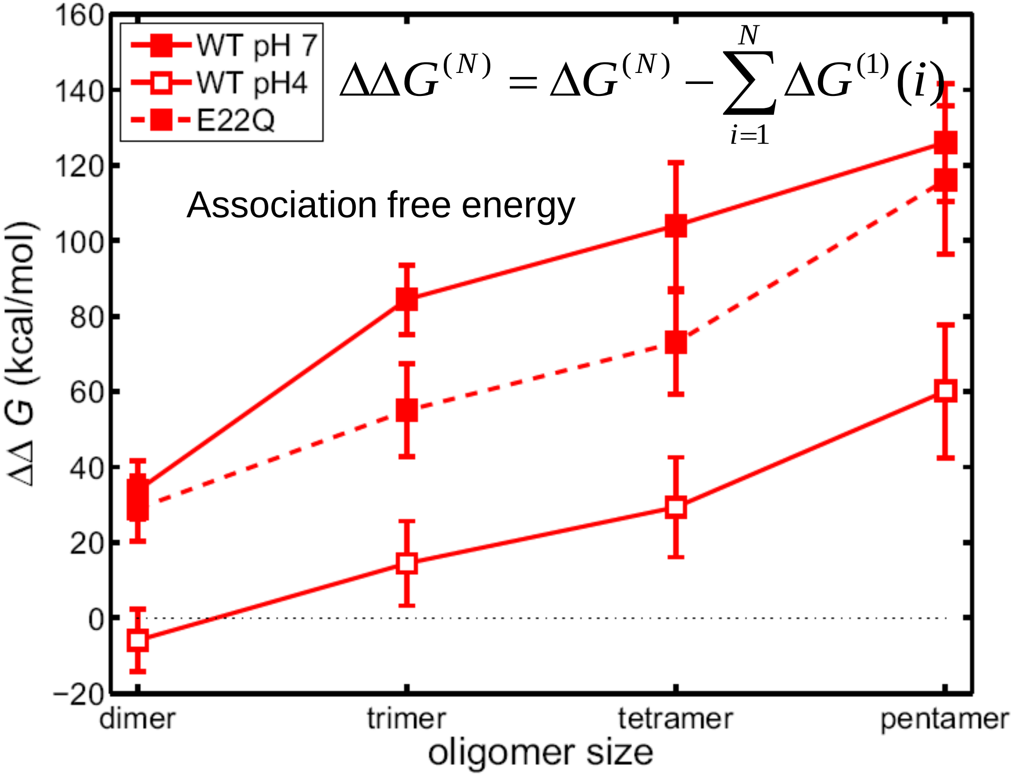
3D maps of the solvation free energy density and the solvation entropy density around the Aβ peptides. Both are related to the (local) energy penalty/gain due to formation of the dry interface between peptides.

MM / 3D-RISM-KH: Effect of mutations on propensity of oligomerization of Alzheimer’s Aβ peptides

Dutch (E22Q) mutants: High oligomerization propensity and fibril elongation rate. Molecular mechanisms of enhanced propensity for oligomerization are unknown



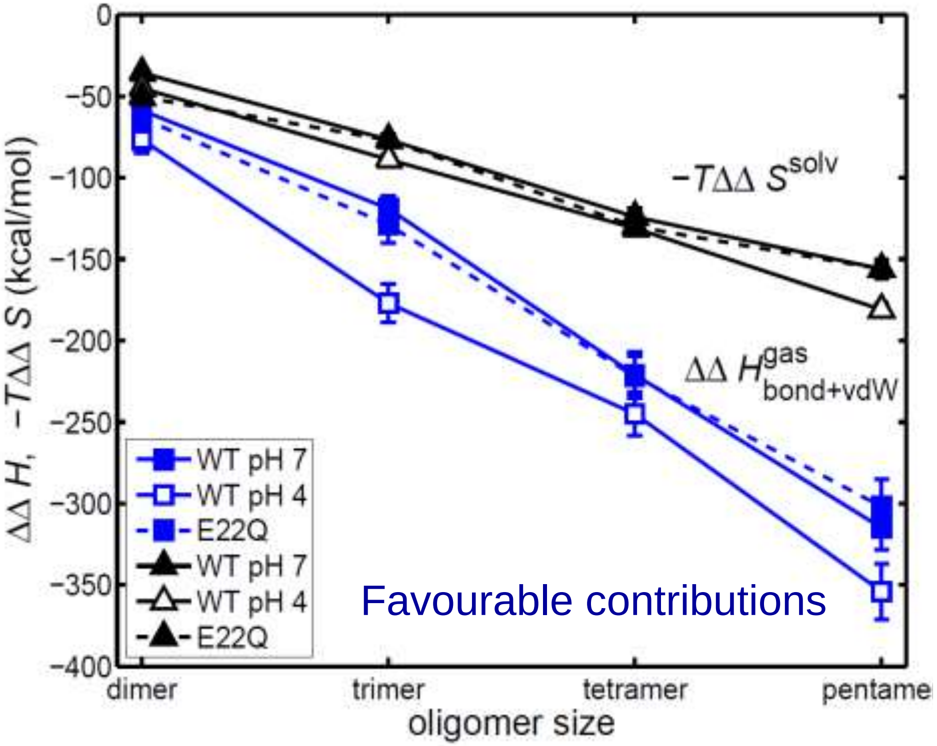
Experimental structure of a fragment of Aβ fibril (PDB ID Code 2BEG) T. Lührs *et. al*, *PNAS* ***102*****,** 17342 (2005)



$$\Delta\Delta G^{(N)} = \Delta G^{(N)} - \sum_{i=1}^{N} \Delta G^{(1)}(i)$$

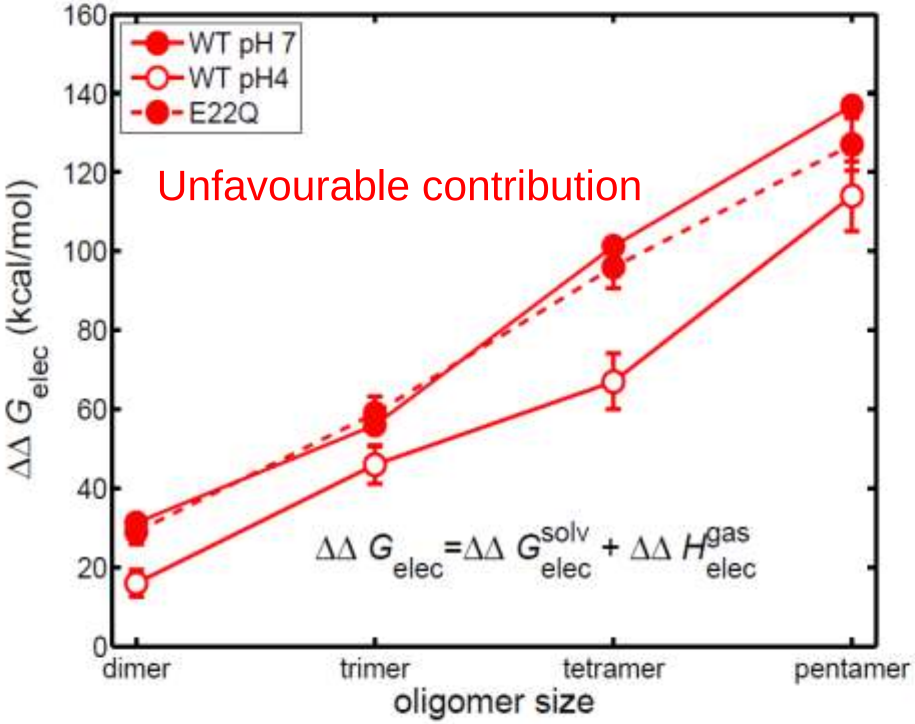
Blinov, N.; Dorosh, L.; Wishart, D.; Kovalenko, A. *Biophys. J.* **2010,** *98*, 282*.*

MM / 3D-RISM-KH: Components of the association free energy of β-sheet Aβ oligomers



Favourable contributions

Solvent entropic (hydrophobic) and gas-phase non-polar contributions to association free energy of Aβ-oligomers

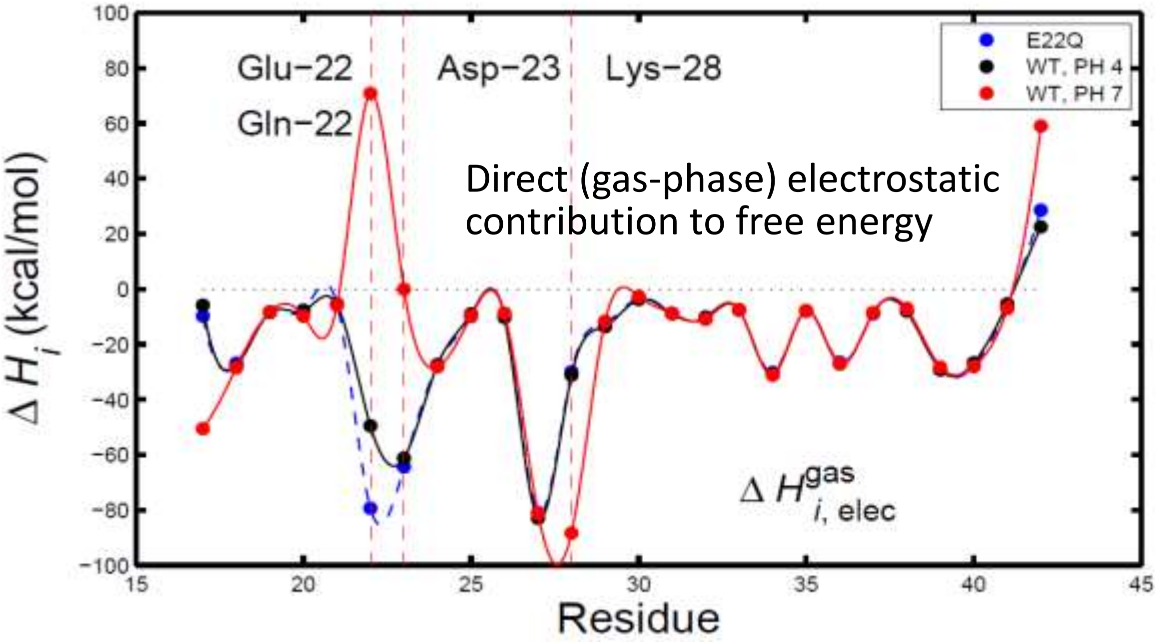


Unfavourable contribution

Combined electrostatic (solvation and gas-phase) contribution to association free energy of Aβ-oligomers

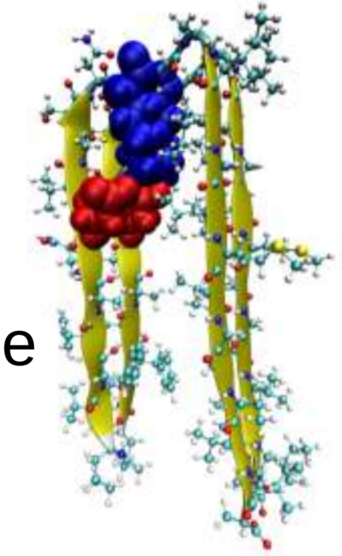
Blinov, N.; Dorosh, L.; Wishart, D.; Kovalenko, A. *Biophys. J.* **2010,** *98*, 282.

MM / 3D-RISM-KH: Local electrostatic contributions to the free energy of β-sheet oligomers



Direct (gas-phase) electrostatic contribution to free energy

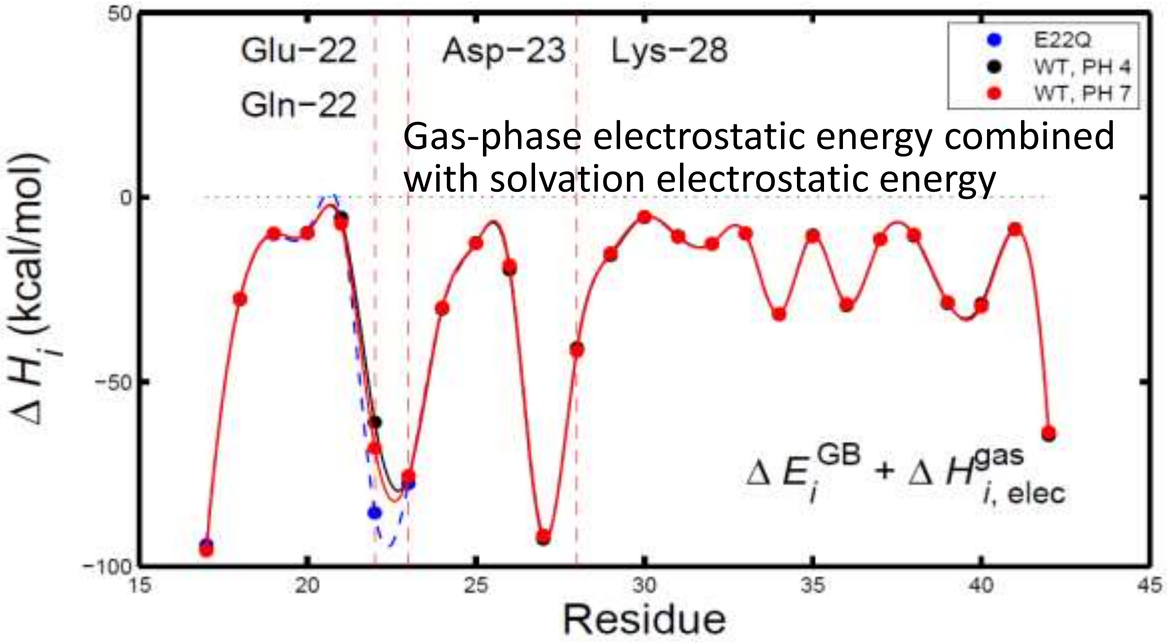
Electrostatic gas-phase contribution



LYS28

ASP23

Salt bridge

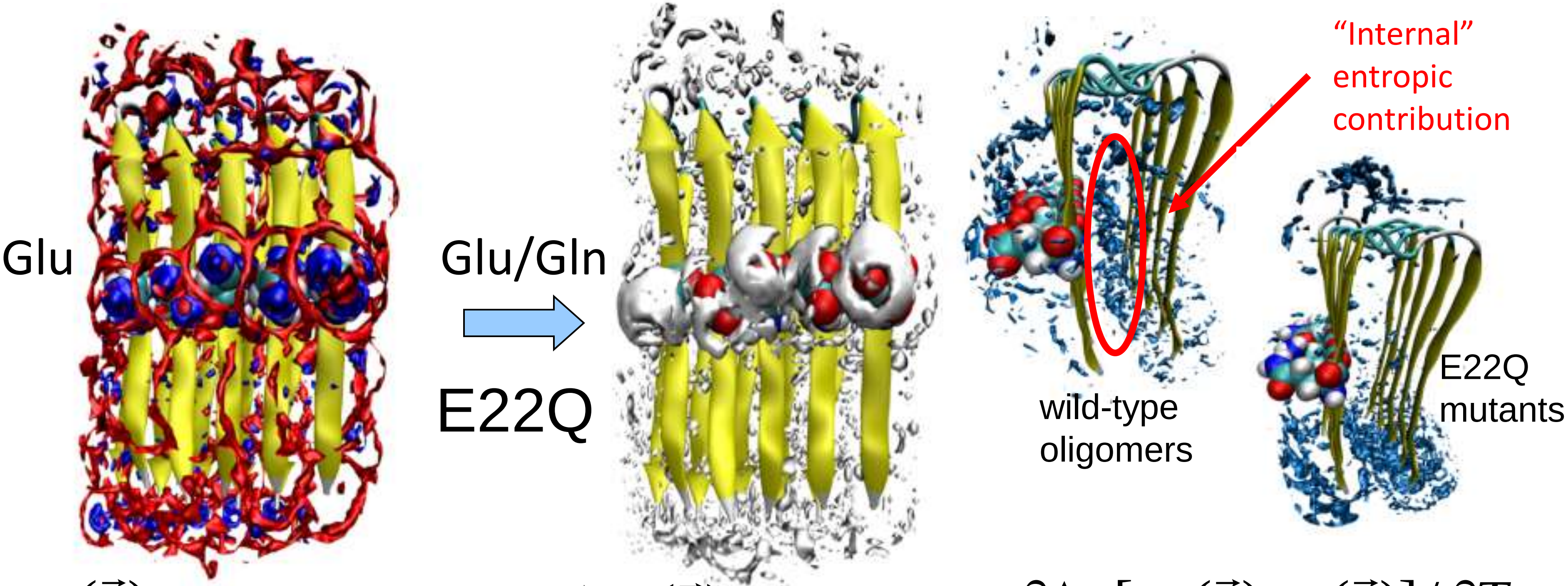


Gas-phase electrostatic energy combined with solvation electrostatic energy

Combined electrostatic (gas-phase and solvation) contributions

Blinov, N.; Dorosh, L.; Wishart, D.; Kovalenko, A. *Biophys. J.* **2010,** *98*, 282.

MM / 3D-RISM-KH: Effect of mutation on microscopic solvation structure of β-sheet oligomers



$$g_\gamma(\vec{r})$$

3D solvation map: isosurface representation of water oxygen (red colour) and hydrogen (blue colour) densities ($\rho=3\rho_{bulk}$)

$$\Delta g_\gamma(\vec{r})$$

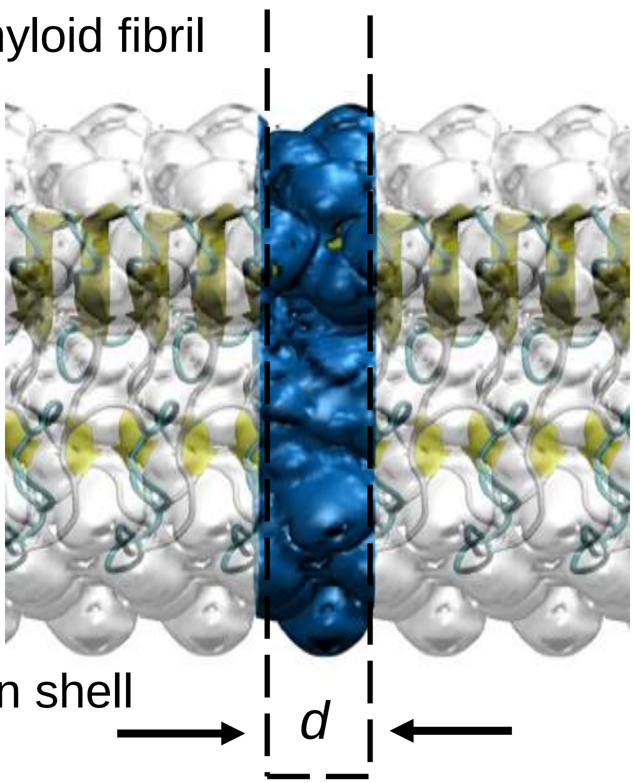
Differential solvation densities: reconstruction of solvation structure upon point mutation

$$-\partial\Delta\mu[g_\gamma(\vec{r}),c_\gamma(\vec{r})]/\partial T$$

Effect of mutation on local entropic contribution to solvation free energy

Blinov, N.; Dorosh, L.; Wishart, D.; Kovalenko, A. *Biophys. J.* **2010,** *98*, 282.

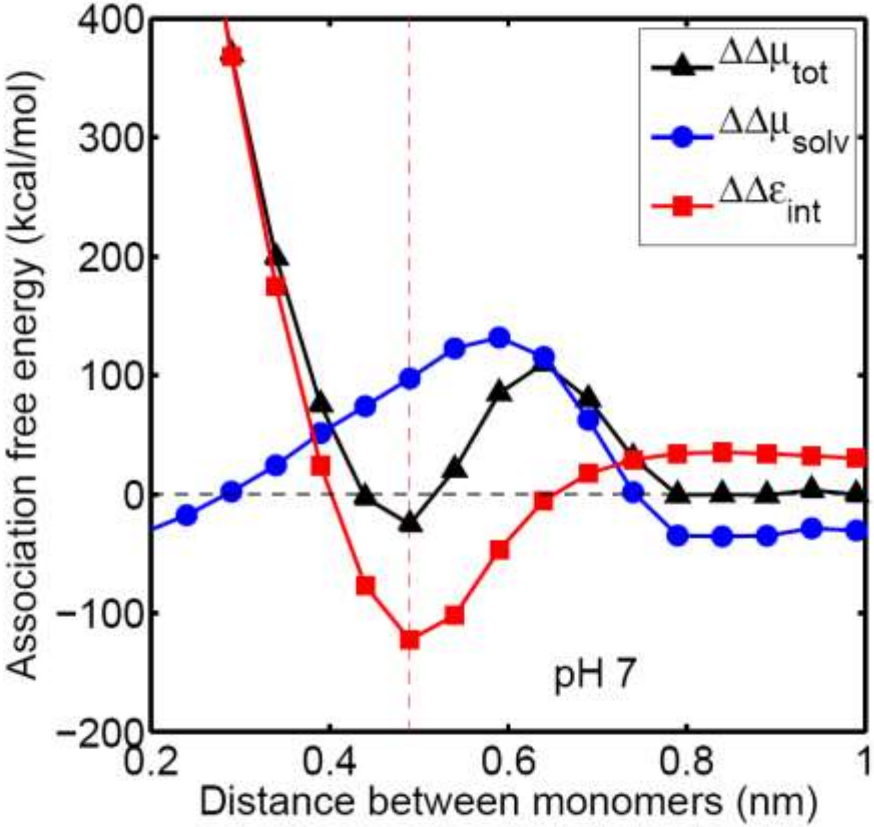
MM / 3D-RISM-KH: Thermodynamic stability and solvation structure of Aβ amyloid fibril



Periodic structural unit of amyloid fibril

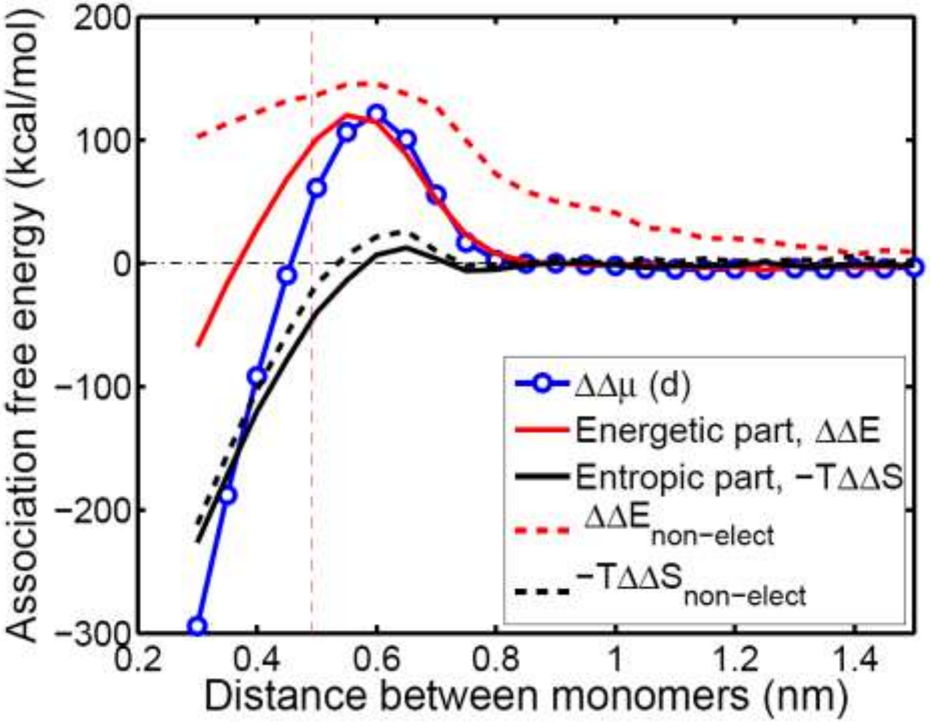
Solvation shell

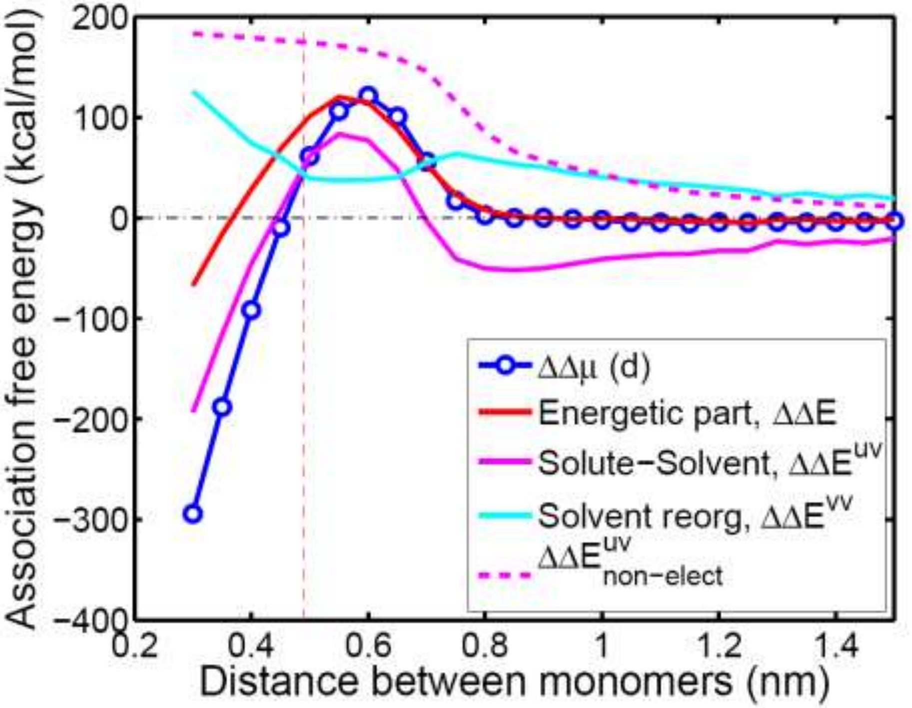
distance between monomers, *d*



Blinov, N.; Dorosh, L.; Wishart, D.; Kovalenko, A. *Molec. Simul*. **2011,** *37*, 718.

MM / 3D-RISM-KH: Decomposition of the association free energy of Aβ amyloid fibril into physical components

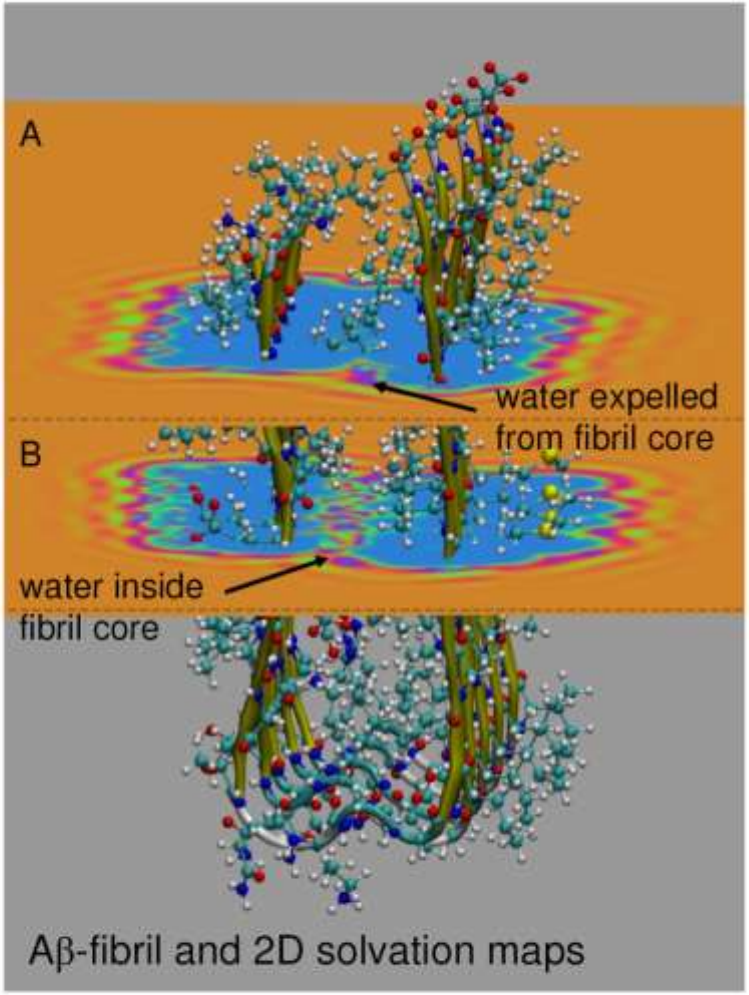




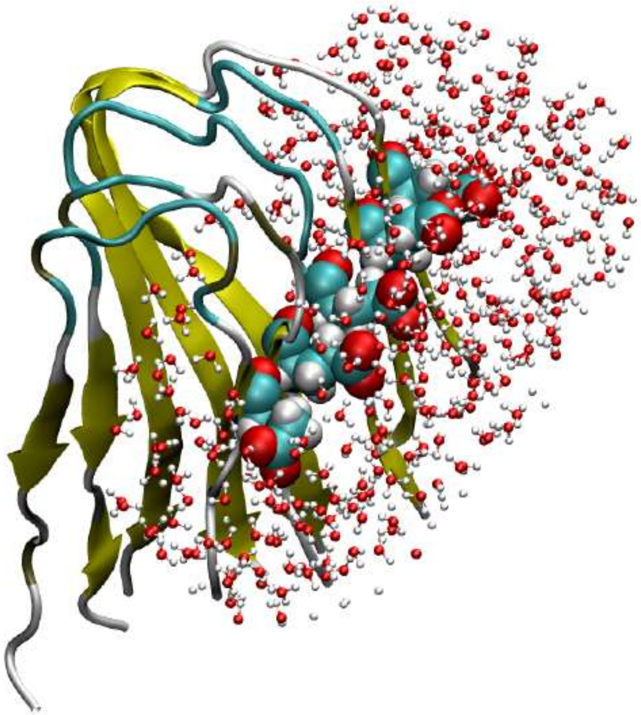
- Desolvation barrier is defined by enthalpic effects
- Solvation energy is mostly defined by electrostatic effects
- Solvation entropy is defined mostly by solute-solvent non-polar interactions

Blinov, N.; Dorosh, L.; Wishart, D.; Kovalenko, A. *Molec. Simul*. **2011,** *37*, 718.

MM / 3D-RISM-KH: Water and ion channels in Aβ amyloid fibril



- Water and solvent ions can be locked inside the fibril core
- Water bridges might form at Aβ surface at pH7
- Water and ion channel in the fibril core? [Zheng *et al*., *Biophys. J*. **93,** 3046 (**2007**)]



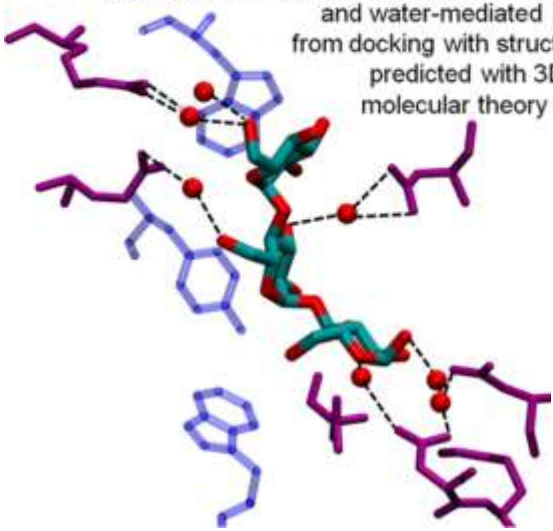
Blinov, N.; Dorosh, L.; Wishart, D.; Kovalenko, A. *Molec. Simul*. **2011,** *37*, 718.

Docking Protocol with Water Detection and Placement using 3D-RISM-KH Molecular Theory of Solvation:

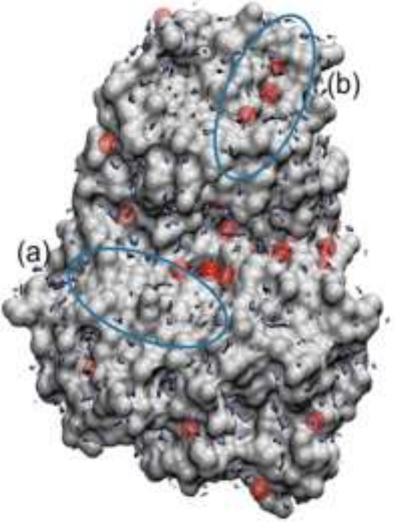
Role of Water in Ligand Binding to Maltose-Binding Protein

Role of Water in Ligand Binding to Maltose-Binding Protein

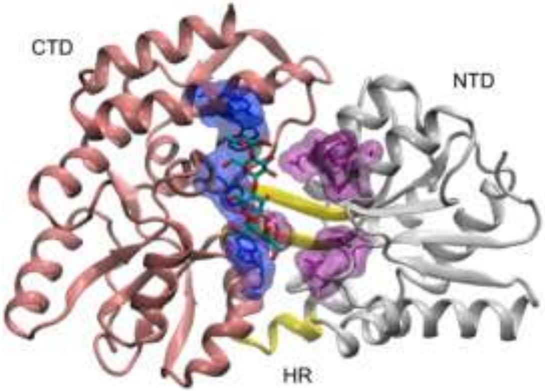
Huang, W.-J. ; Blinov, N.; Wishart, D. S.; Kovalenko, A.
J. Chem. Inf. Model. **2015**, *55*,317.



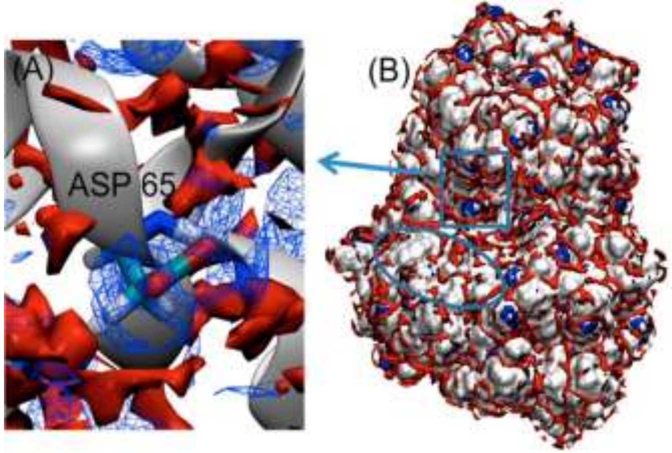
Maltotriose binding mode at maltose-binding protein
and water-mediated interactions
from docking with structural waters
predicted with 3D-RISM-KH
molecular theory of solvation



High density peaks of 3D water maps around the apo-open state of MBP predicted with the 3D-RISM-KH theory. Water oxygen density isosurface of 5 times the bulk water density (blue color). X-ray waters from the apo-open state of MBP (re shown with red vdW spheres of water oxygens. Solvent exposed surface of the protein is in gray. Region (a) corresponds to the experimental binding site of MBP. Area (b) with three X-ray water molecules overlapping with the high density peaks in water distribution.



Holo-open structure of MBP with bound maltotetraitol (PDB accession code 1EZ9)



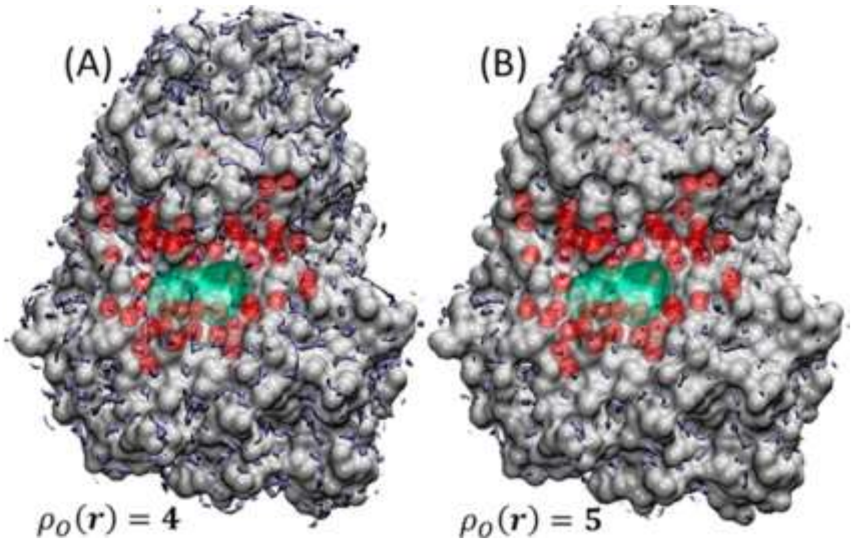
Water oxygen and hydrogen density distributions around the apo-open state of MBP predicted with the 3D-RISM-KH theory. Density isosurfaces of 4 for oxygen (red) and 3 for hydrogen (blue) times the bulk water density. Polarization of water is clearly seen around the polar residues.

(A) Zoomed view around Asp65 where a hydrogen bond forming water molecule was identified in X-ray experiments.

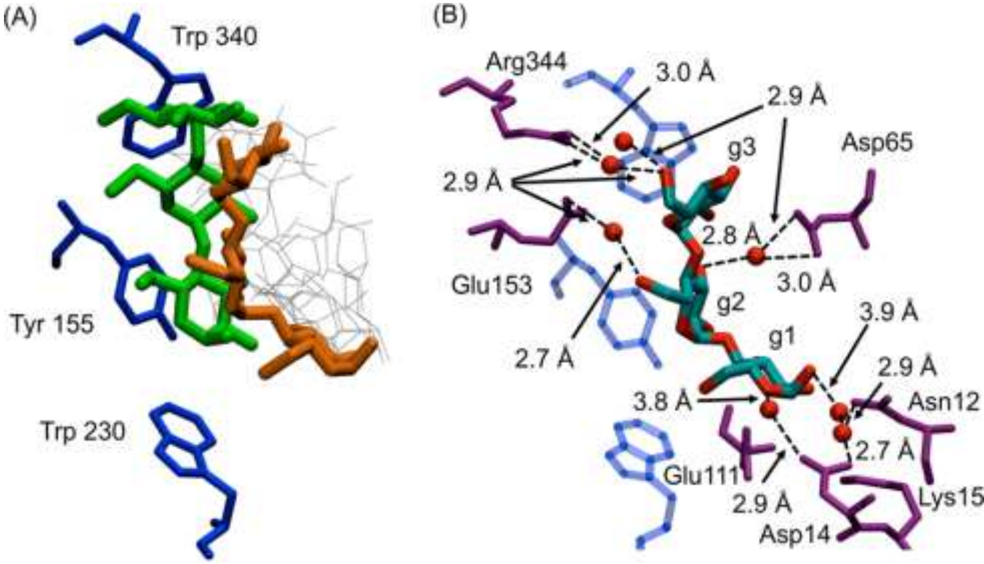
(B) Water density polarization indicative of strong water-protein bonding is seen in the Nterminal domain above the maltotriose putative binding site (marked with the blue ellipse).

Role of Water in Ligand Binding to Maltose-Binding Protein

Huang, W.-J. ; Blinov, N.; Wishart, D. S.; Kovalenko, A. *J. Chem. Inf. Model.* **2015**, *55*,317.



Water oxygen density distributions (shown in blue) around the apo-open state of MBP obtained with the 3D-RISM-KH theory. Structural water molecules predicted by the 3D-RISM-KH theory are shown with red van der Waals spheres. Solvent exposed surface of the protein is in gray. The density maps are shown for isolevel of 4 (panel A) and 5 (panel B) in excess of the water bulk density. The ligand binding site is shown by the green surface.



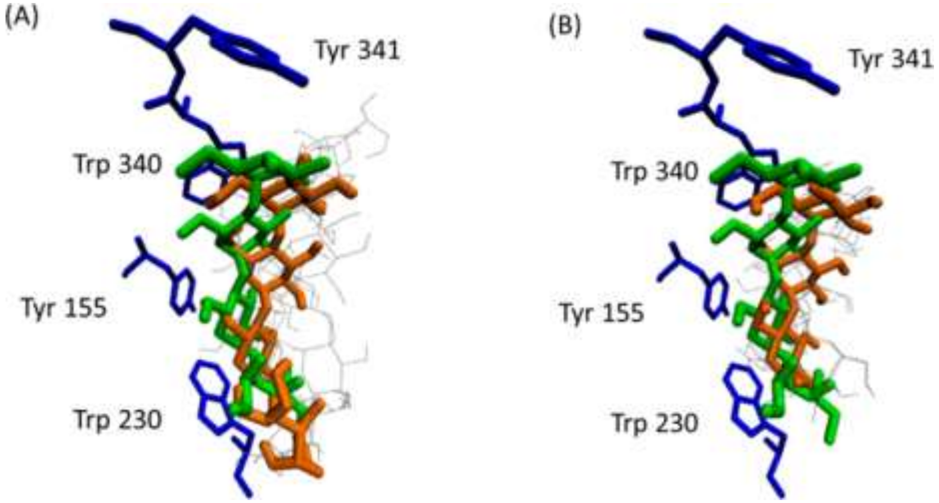
Binding modes of maltotriose at the open state of MBP from docking experiments with explicit water molecules predicted by the 3D-RISM-KH molecular theory of solvation.

(A) Top-ranked maltotriose conformation from docking experiments (orange) compared to the conformation of maltotriose from the superimposed holo-closed X-ray structure (green).
The next four low-ranked binding modes are shown in gray.

(B) Predicted water-mediated hydrogen bonds (dashed lines) between the top-ranked binding mode and the MBP polar residues. The polar and aromatic residues are shown in purple and blue. Water oxygens are represented by red van der Waals spheres.

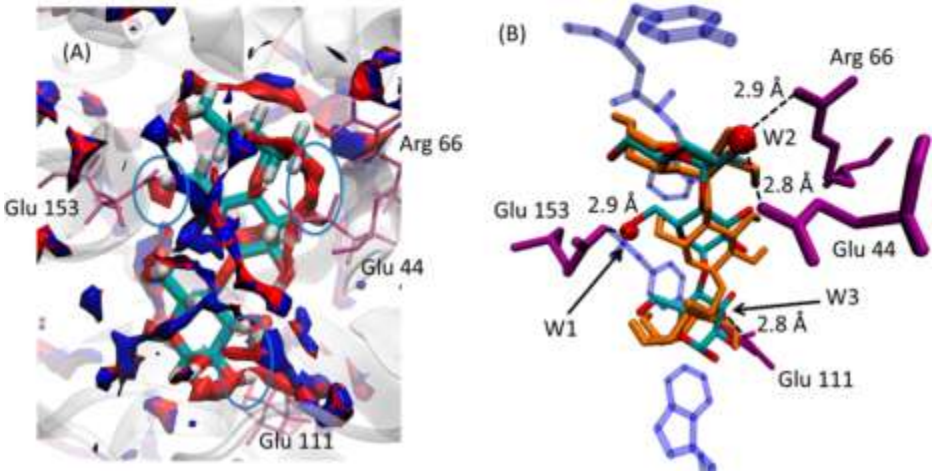
Role of Water in Ligand Binding to Maltose-Binding Protein

Huang, W.-J. ; Blinov, N.; Wishart, D. S.; Kovalenko, A. *J. Chem. Inf. Model.* **2015**, *55*,317.



(A) Binding modes of maltotetraitol at the open state of MBP from docking experiments with explicit water molecules predicted by the 3D-RISM-KH molecular theory of solvation. The top-ranked ligand conformation from docking experiments (orange) is compared to the experimental structure (PDB accession code 1EZ9) (green).
The next four low-ranked binding modes are shown in gray.

(B) Binding modes of maltotriose at the open state of MBP (receptor structure for docking was modeled after the X-ray holo-open structure of MBP, PDB accession code 1EZ9).
The notations are the same as in panel A.

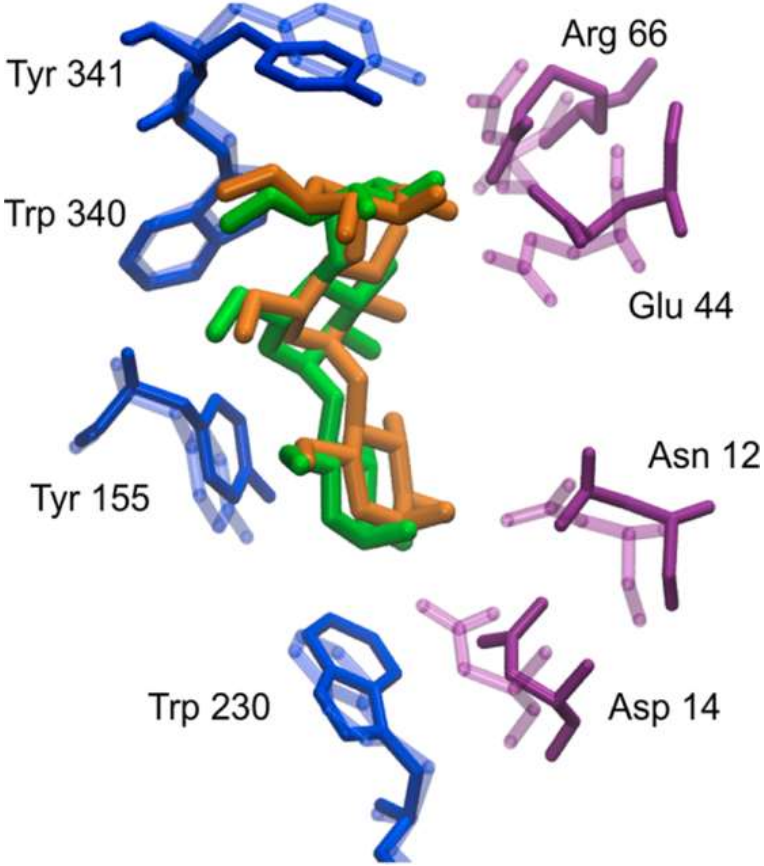


(A) Water oxygen density distributions around the maltotriose binding site in holo-closed (blue) and the apo-closed (red) states of MBP, as predicted with the 3D-RISM-KH theory. Density isosurfaces at isovalue 4.
(Some −OH groups of maltotriose in the binding site overlap (in blue ellipse areas) with the high peaks of the water oxygen density distribution in the MBP unliganded state.)

(B) X-ray structure of the holo-closed state of MBP.
Water molecules blocking the binding site as predicted with the water placement protocol for the apo-closed structure are shown as red van der Waals spheres. Polar and aromatic residues are shown in purple and blue, respectively.

Role of Water in Ligand Binding to Maltose-Binding Protein

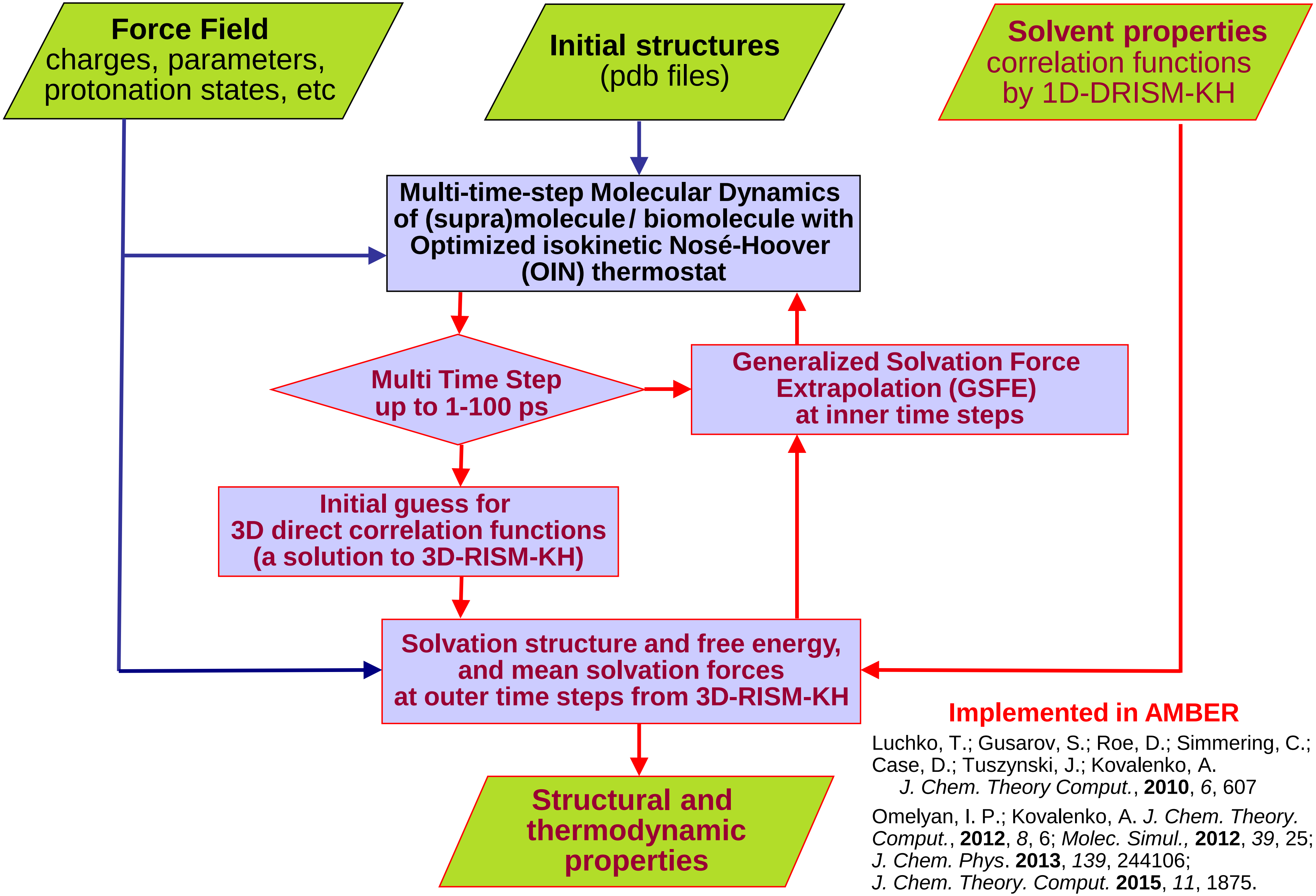
Huang, W.-J. ; Blinov, N.; Wishart, D. S.; Kovalenko, A. *J. Chem. Inf. Model.* **2015**, *55*,317.



Snapshot of MD simulation started from the top-ranked maltotriose conformation from docking experiments with explicit water (orange) compared to the conformation of maltotriose in the holo-closed X-ray structure (green). Aromatic residues in the binding site are shown in blue, and polar residues are shown in purple.

MTS-MD of biomolecule steered with 3D-RISM-KH mean solvation forces: OIN / GSFE / 3D-RISM-KH integrator

MTS-MD / 3D-RISM-KH in Amber package



Force Field
charges, parameters,
protonation states, etc

Initial structures
(pdb files)

Solvent properties
correlation functions
by 1D-DRISM-KH

Multi-time-step Molecular Dynamics of (supra)molecule / biomolecule with Optimized isokinetic Nosé-Hoover (OIN) thermostat

Multi Time Step up to 1-100 ps

Generalized Solvation Force Extrapolation (GSFE) at inner time steps

Initial guess for 3D direct correlation functions (a solution to 3D-RISM-KH)

Solvation structure and free energy, and mean solvation forces at outer time steps from 3D-RISM-KH

Structural and thermodynamic properties

Implemented in AMBER

Luchko, T.; Gusarov, S.; Roe, D.; Simmering, C.; Case, D.; Tuszynski, J.; Kovalenko, A. *J. Chem. Theory Comput.*, **2010**, *6*, 607

Omelyan, I. P.; Kovalenko, A. *J. Chem. Theory. Comput.*, **2012**, *8*, 6; *Molec. Simul.*, **2012**, *39*, 25; *J. Chem. Phys*. **2013**, *139*, 244106; *J. Chem. Theory. Comput.* **2015**, *11*, 1875.

Generalized solvation force extrapolation

- At long outer time steps:
  Mean solvation forces are obtained by converging 3D-RISM-KH

  Mean solvation forces change smoothly in space and time !
- At successive short inner time steps:
  Extrapolate with the basis of previously saved force-coordinate pairs
- Basis of previous saved pairs of solvation force – biomolecule atomic coordinates
  - Save many previous pairs
  - Choose the (best) closest in coordinated
- Make non-Eckart like transformation of coordinates – to make solvation forces in transformed coordinates (with weights) changing even smoother
- Linear squares minimization for coefficients of expansion in the basis of previous solvation force – biomolecule atomic coordinates
  - Balancing the normal equations

Omelyan, I. P.; Kovalenko, A.. *Chem. Theory. Comput.* **2015**, *11*, 1875.

Multi time step MD simulation of a biomolecule steered with 3D-RISM-KH effective solvation forces

First MD/3D-RISM coupling: T. Miyata and F. Hirata, *J. Comput. Chem.*, (2008).

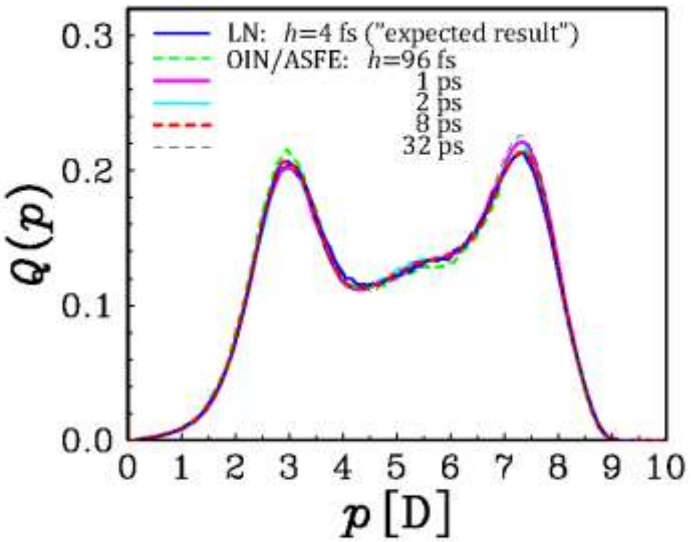
Multi time step LN / GSF / 3D-RISM-KH implemented in AMBER – outer time steps up to 20 fs
Luchko, T.; Gusarov, S.; Roe, D.; Simmering, C.; Case, D.; Tuszynski, J.; Kovalenko, A.
J. Chem. Theory Comput., **2010**, *6*, 607.

Generalized isokinetic ensemble MD, optimized isokinetic Nosé-Hoover chain (OIN) algorithm
Multi time step MD: OIN / GSFE / 3D-RISM-KH integrator (implementation in AMBER)

- **Outer time steps up to 1-4 ps – up to 20 faster (direct comparison) than explicit solvent MD**
- **Protein folding – up to 1000 times faster than with explicit solvent MD**

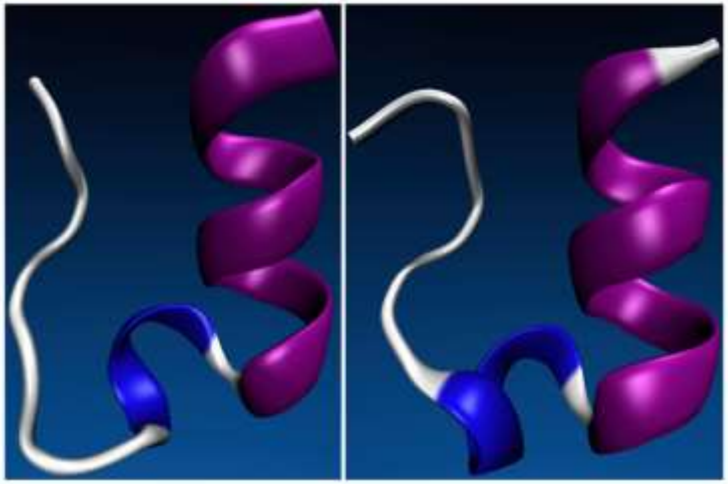
Omelyan, I. P.; Kovalenko, A. *J. Chem. Theory. Comput.* **2012**, *8*, 6; *Mol. Simul.* **2012**, *39*, 25;
J. Chem. Phys. **2013**, *139*, 244106; *J. Chem. Theory. Comput.* **2015**, *11*, 1875.

- Solvation structure and thermodynamics of large biomolecules on a very long-time scale: the method properly captures statistics of slow and rare solvation events (exchange of solvent molecules, counter ions, ligands, etc with inner spaces and pockets of solute)
- Solvent and host species in inner spaces (e.g. pockets and nanochannels)
- Simulation of large biomolecules (e.g. microtubules) in solution with co-solvent, buffer, drug molecules
- Accelerated calculation of protein-drug and protein-protein effective interactions in solution
- *MM / 3D-RISM-KH post-processing replaces MM / GB(PB)SA*

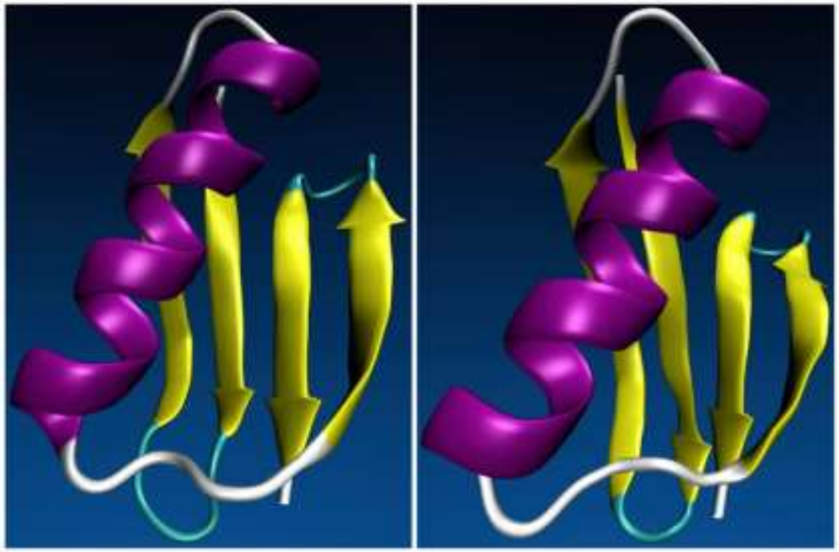


Hydrated alanine dipeptide:
Dipole moment distribution
by OIN/ASFE/3D-RISM-KH
versus the “expected “result by LN/3D-RISM-KH

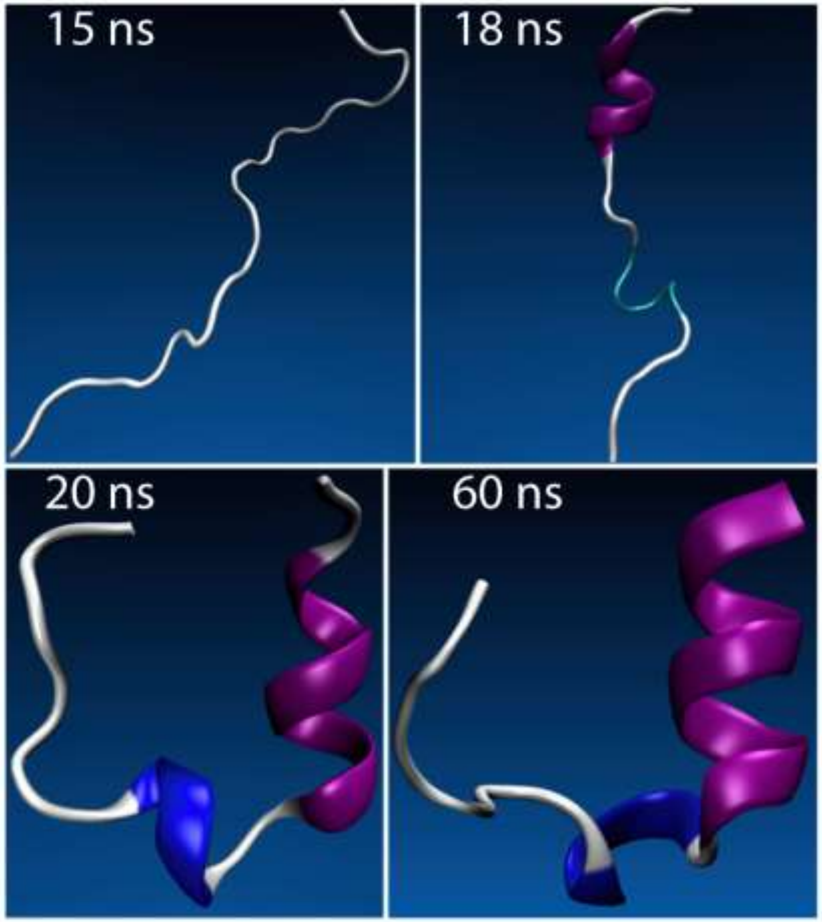
MTS-MD steered with 3D-RISM-KH mean solvation forces extrapolated with GSFE and stabilized with OIN thermostat



Tertiary structures of miniprotein 1L2Y
NMR experiment vs OIN / GESF / 3D-RISM-KH
length 25ns, outer step 1ps



Tertiary structures of protein G
NMR experiment vs OIN / GSFE / 3D-RISM-KH



Folding miniprotein 1L2Y

- **In 60 nanoseconds by OIN/GSFE/3D-RISM-KH**
- **in 4-9 microseconds of physical time in experiment**

100-fold acceleration

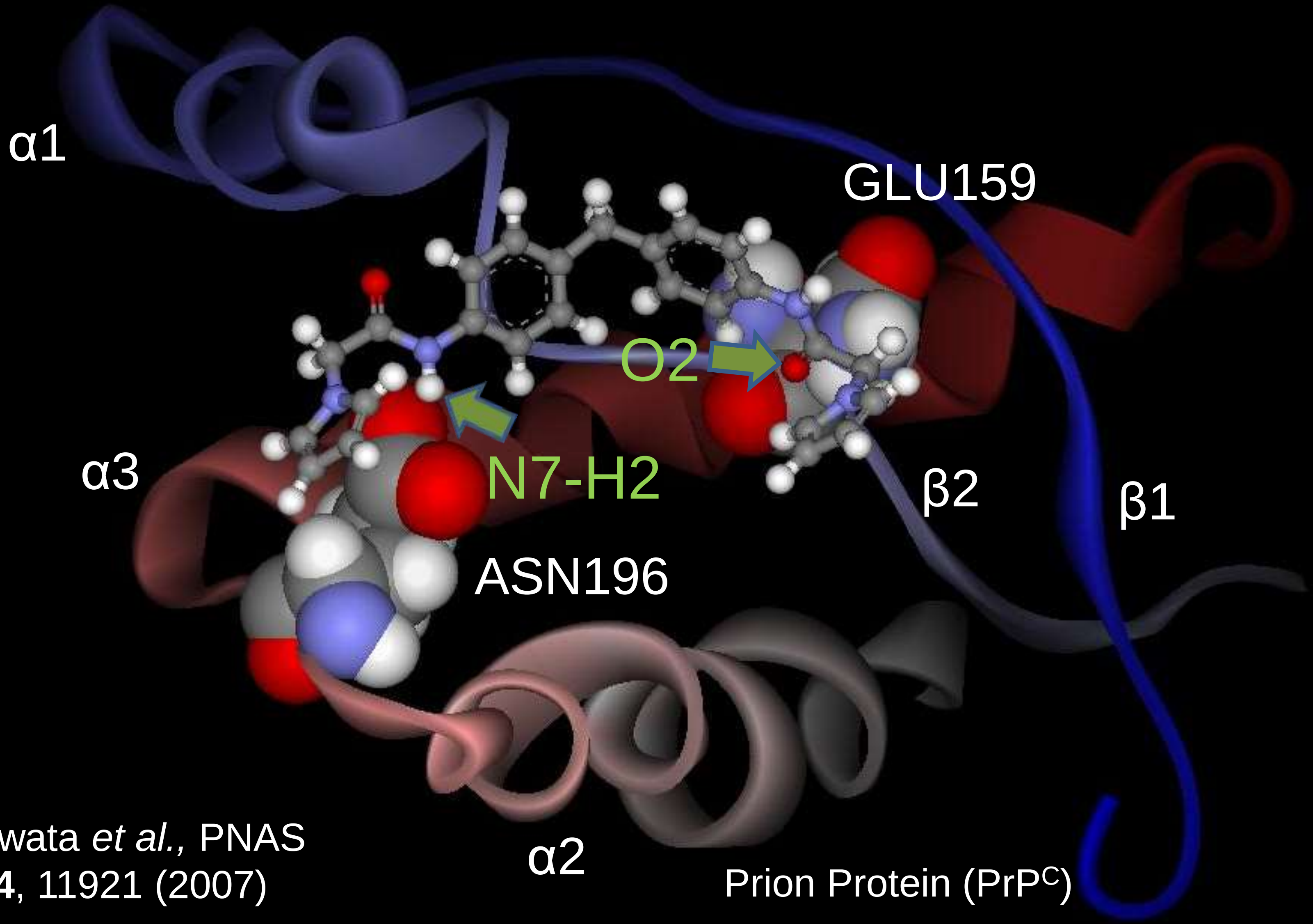
Omelyan, I. P.; Kovalenko, A. *J.Chem.Theory Comput.* **2015**, *11*, 1875

3D-RISM-Dock approach for protein-ligand mapping and fragment-based drug design

Ligands (ligand fragments) as distinct solvent species:
3D maps of ligand binding affinity

Binding of GN8 antiprion compound at PrP^C

2-pyrrolidin-1-yl-*N*-[4-[4-(2-pyrrolidin-1-yl-acetylamino)-benzyl]-phenyl]-acetamide

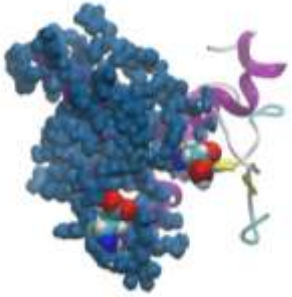


Kuwata *et al.*, PNAS
104, 11921 (2007)

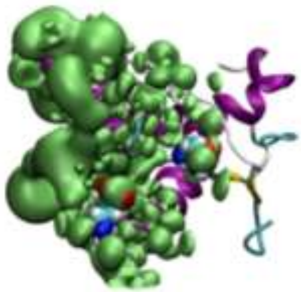
Prion Protein (PrP^C)

3D-RISM-Dock: New methodology for docking

Conventional docking: 3D affinity maps (AutoDock)



Ligand site affinity map

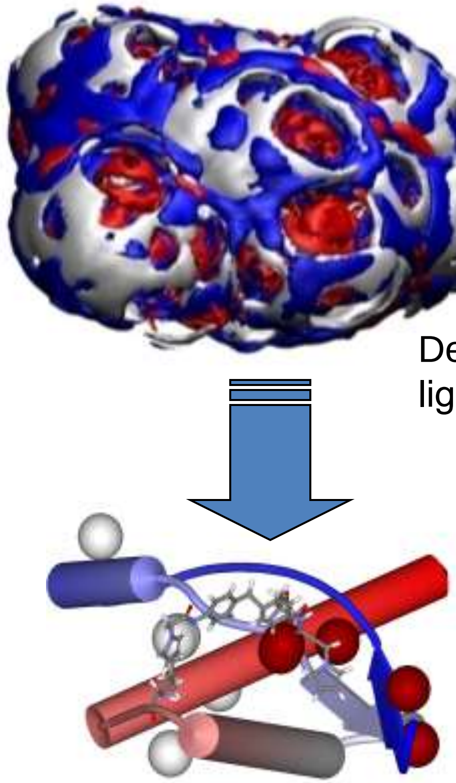


Electrostatic map

Desolvation map

Affinity maps from 3D-RISM-KH

[T. Imai *et al.*, *JACS* **131,** 12430 (2009)]



Densities of ligand's sites

🟥 O2
🟦 N7
▨ H2

GN8 antiprion compound bound to prion protein, Kuwata *et al.*, PNAS **104**, 11921 (2007)

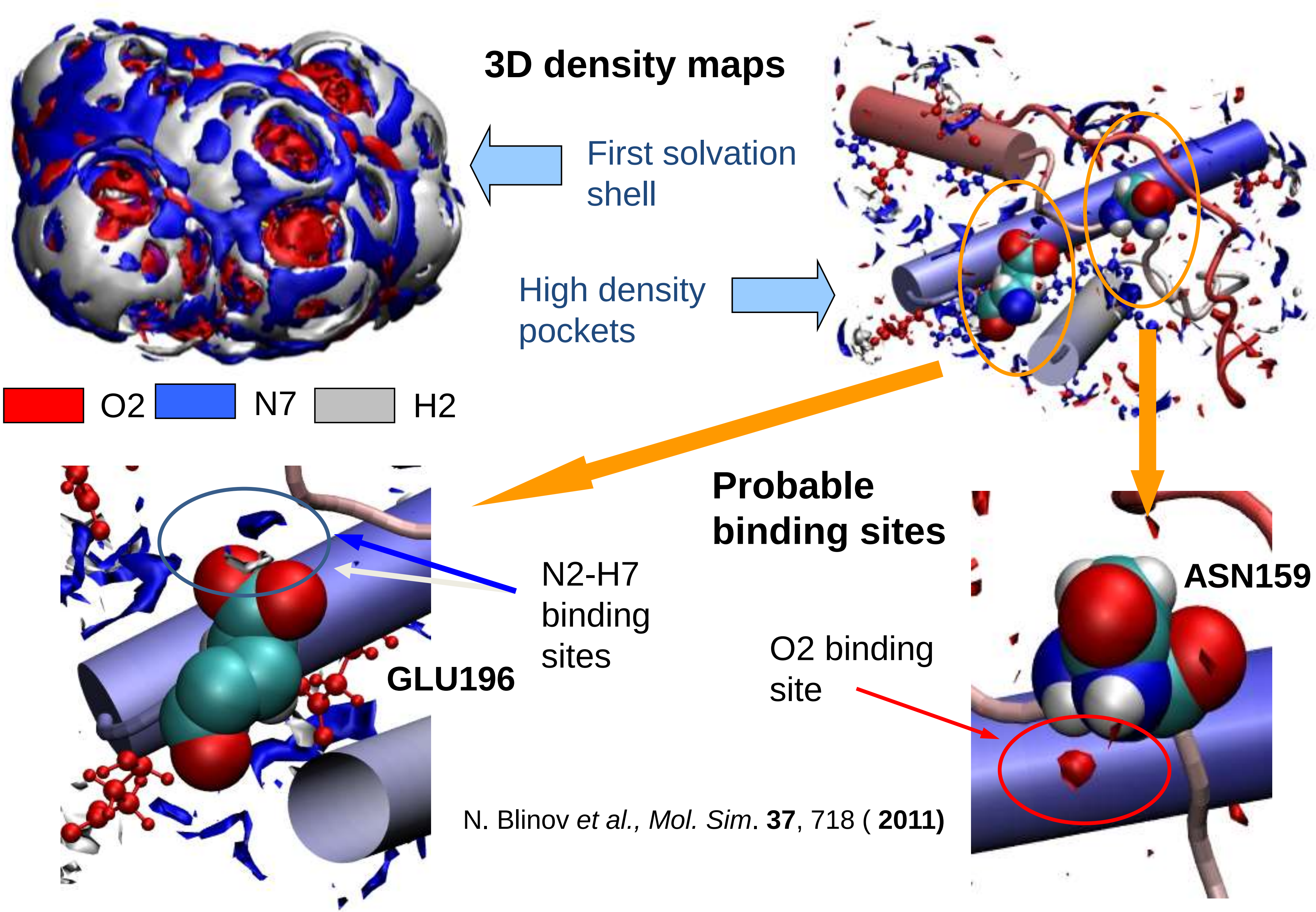
Advantages of 3D-RISM-DOCK algorithm

- No phenomenological input required
- Rigorous treatment of solvation effects
- Concentration/temperature effects on docking
- Composition effects: Competition of (multiple) ligands, water, co-solvent, and ions for binding sites

Implemented in AutoDock

Nikolić, D.; Blinov, N.; Wishart, D. S.; Kovalenko, A. *J. Chem. Theory Comput.* **2012**, *8*, 3356.

Mapping GN8 antiprion compound at PrP^C



3D density maps

First solvation shell

High density pockets

🟥 O2 🟦 N7 ⬜ H2

Probable binding sites

N2-H7 binding sites

GLU196

O2 binding site

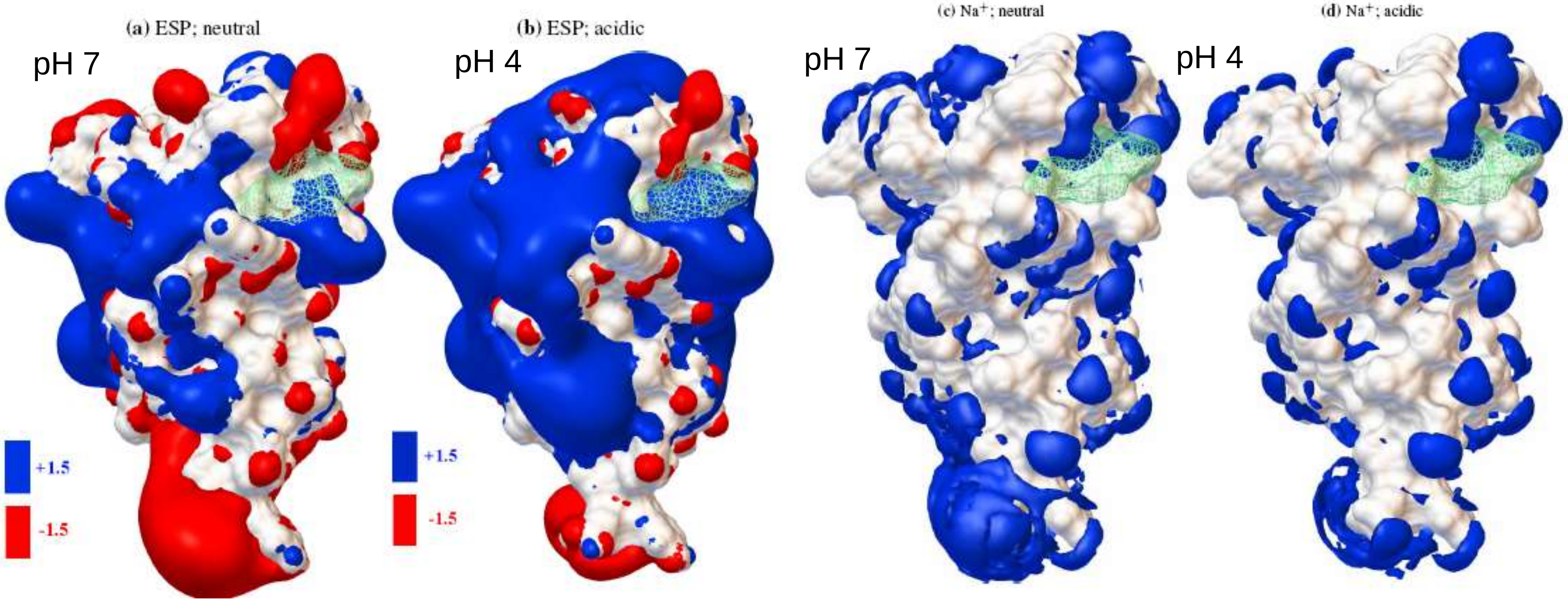
ASN159

N. Blinov *et al.*, *Mol. Sim*. **37**, 718 (**2011)**

3D-RISM-Dock: a new fragment-based drug design protocol

3D-RISM-KH scoring function: $\Delta G^{3D}_{RISM} = \Lambda(\boldsymbol{r}) + \Delta(\boldsymbol{r}) + N_{tor}W_{tor}$

$\Lambda(\boldsymbol{r}) \propto \sum_{\gamma} W_{\gamma}(\boldsymbol{r})$ (Sum over ligand sites) $\Delta(\boldsymbol{r}) \propto -\sum_{\sigma} W_{\sigma}(\boldsymbol{r})$ (Sum over solvent sites)



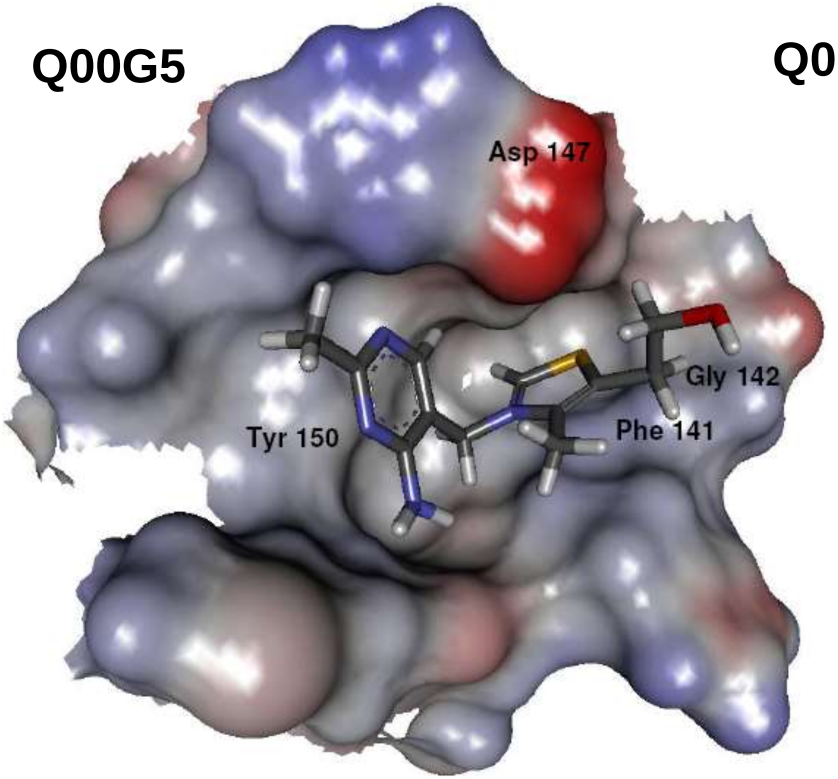
Electrostatic potential around prion protein **Na+ ion distribution around prion protein**

Nikolić, D.; Blinov, N.; Wishart, D. S.; Kovalenko, A. *J. Chem. Theory Comput.* **2012**, *8*, 3356.

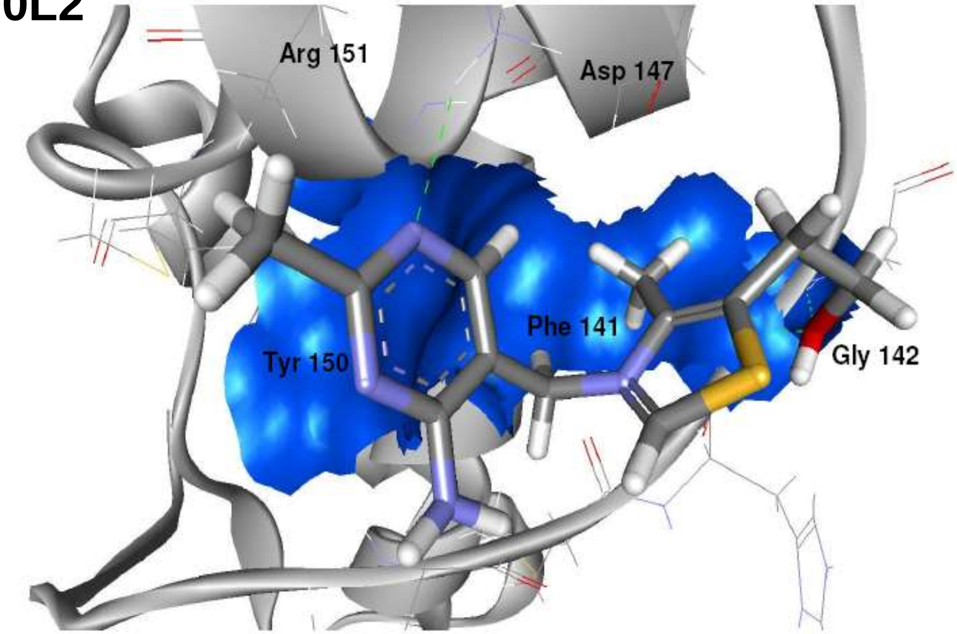
3D-RISM-Dock case study:
Thiamine binding to prion protein

Nikolić, D.; Blinov, N.; Wishart, D. S.; Kovalenko, A. *J. Chem. Theory Comput.* **2012**, *8*, 3356.

Q00G5



Q00L2



Electrostatic stabilization

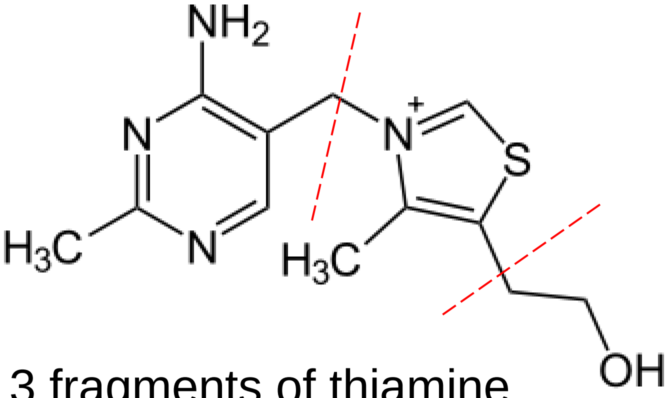
Hydrophobic stabilization

Binding modes are in agreement with the experimental NMR data

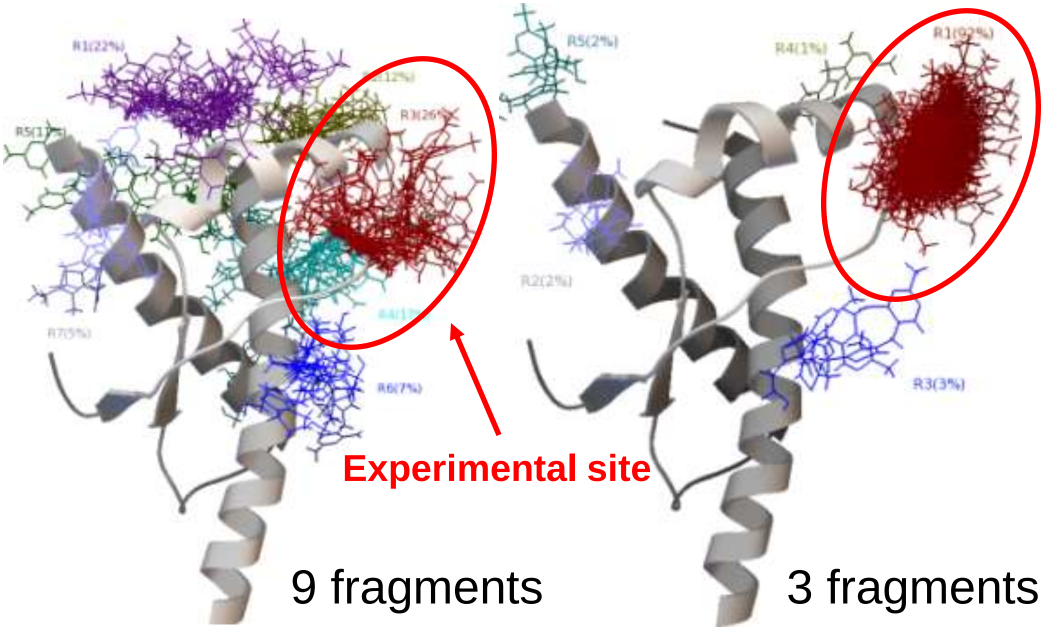
[R. Perez-Pineiro *et al., FEBS* journal ***2011***]

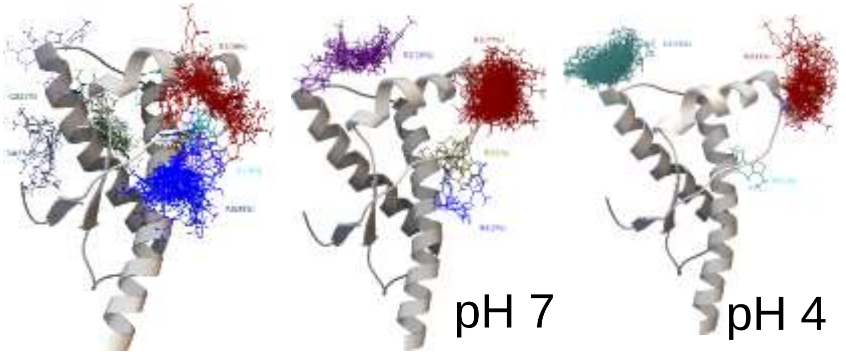
3D-RISM-Dock case study: Thiamine binding to prion protein Effects of ligand fragmentation on binding modes

[experiment: R. Perez-Pineiro *et al.*, *FEBS* journal ***2011***]

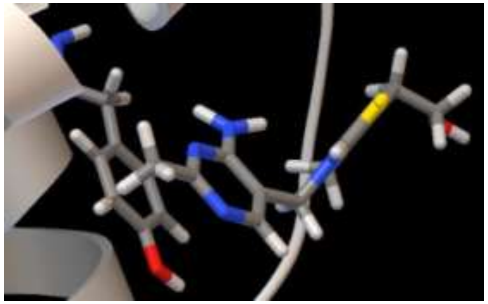


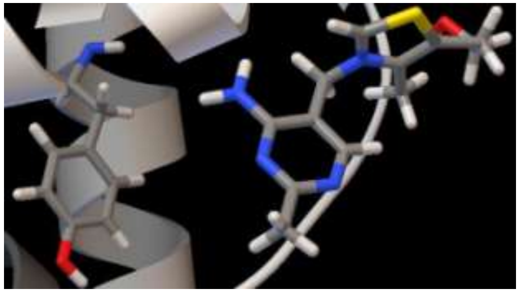
3 fragments of thiamine molecule used in docking experiments with prion protein





AutoDock 3D-RISM-DOCK

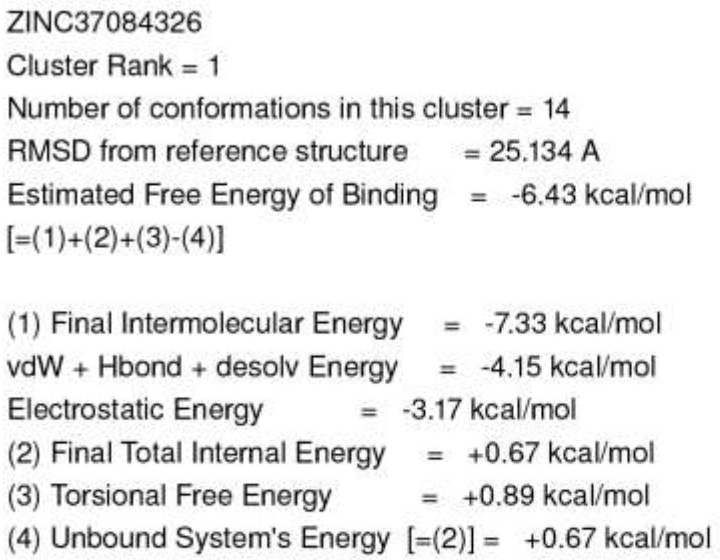




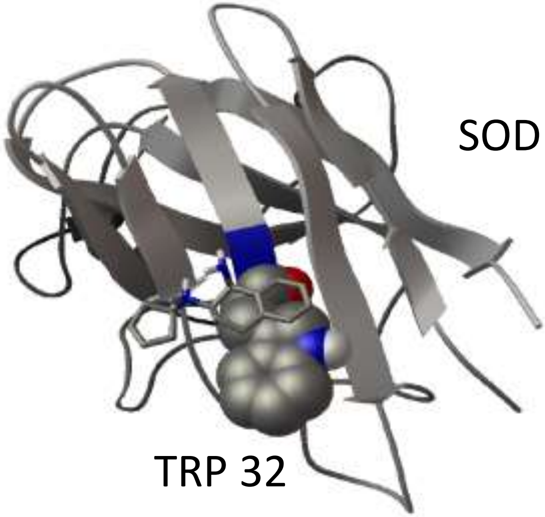
Superoxide Dismutase (SOD1): Inhibitors

High throughput virtual screening targeting the TRP-32 binding site of SOD1

Screening was performed with ZINC (UCSF) Fragment-Like library of compounds (in total, more than one million of compounds have been screened)

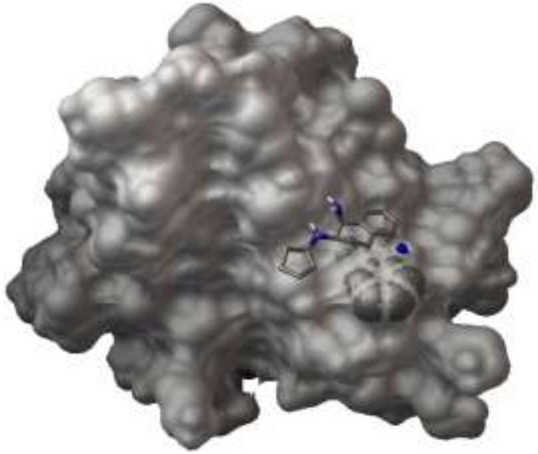


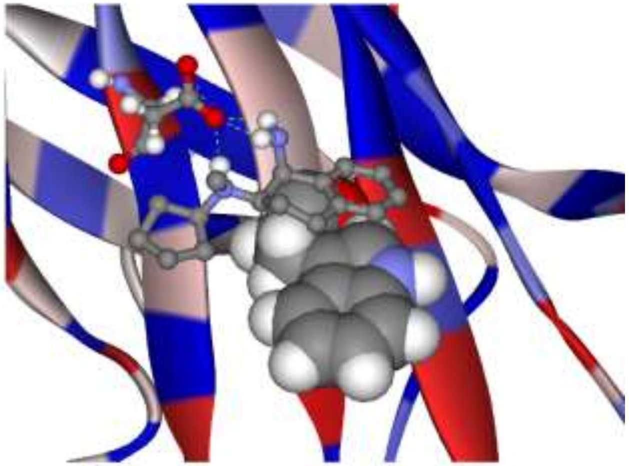
```
ZINC37084326
Cluster Rank = 1
Number of conformations in this cluster = 14
RMSD from reference structure       = 25.134 A
Estimated Free Energy of Binding    =  -6.43 kcal/mol
[=(1)+(2)+(3)-(4)]

(1) Final Intermolecular Energy     =  -7.33 kcal/mol
vdW + Hbond + desolv Energy         =  -4.15 kcal/mol
Electrostatic Energy           =  -3.17 kcal/mol
(2) Final Total Internal Energy     =  +0.67 kcal/mol
(3) Torsional Free Energy           =  +0.89 kcal/mol
(4) Unbound System's Energy  [=(2)] =  +0.67 kcal/mol
```



SOD1 receptor

TRP 32

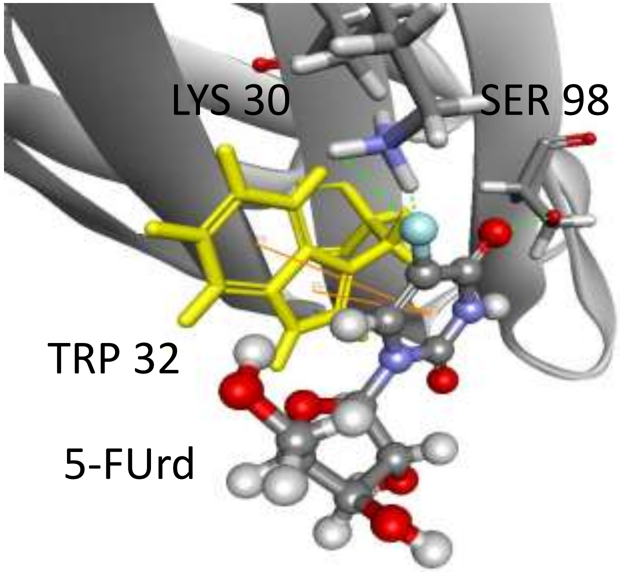




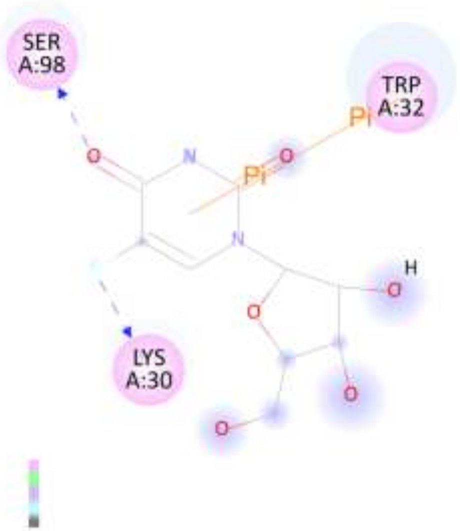
Optimization of the experimentally confirmed inhibitors targeting TRP32 binding site of SOD1 and rational design of new drug compounds

- Gareth S.A. Wright et al., Nature Comm. 4:1758 (2013): isoproterenol and 5-fluorouridine binds to TRP32 at SOD1
- Cashman’s lab: 5-fluorouridine and other compounds (such as uridine) block pathological misfolding of SOD1

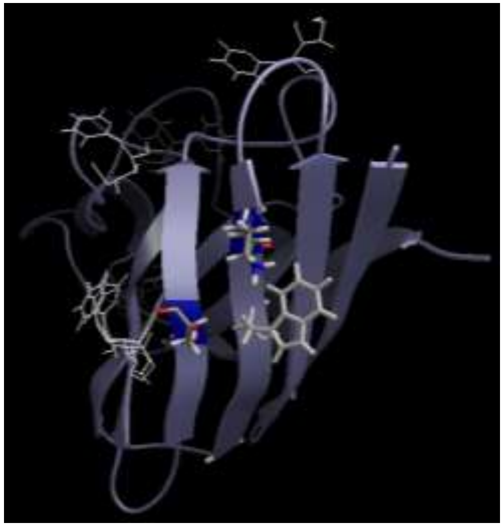
X-ray structure (PDB ID 4A7S) of SOD1 with bound 5-Fluorouridine



Molecular structure of SOD1 with bounds 5-FUrd

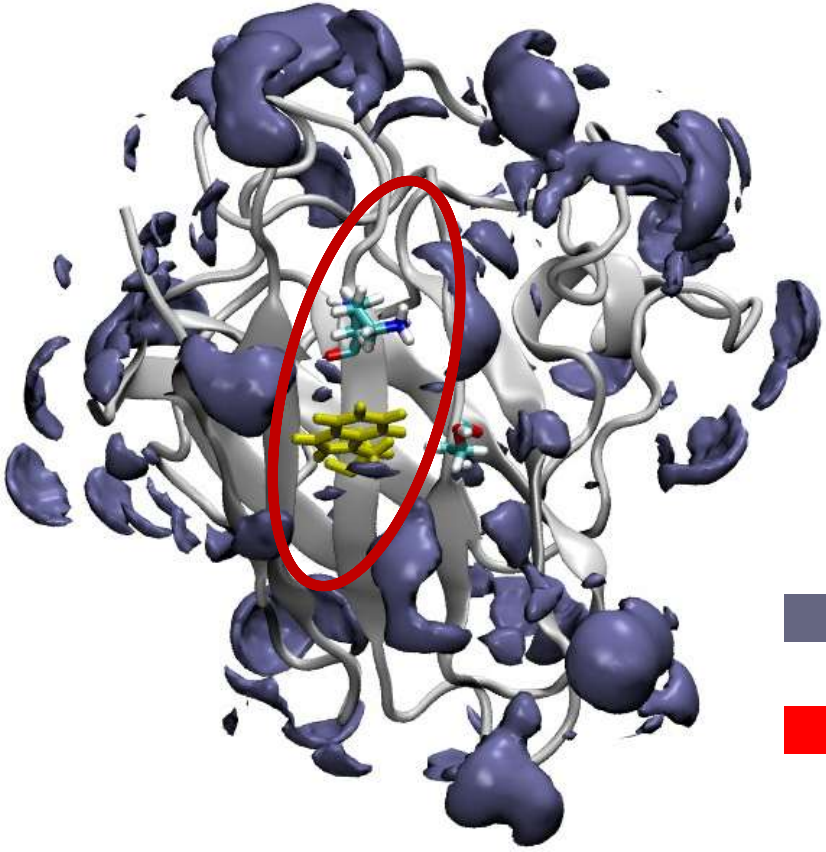


Protein-ligand interactions map



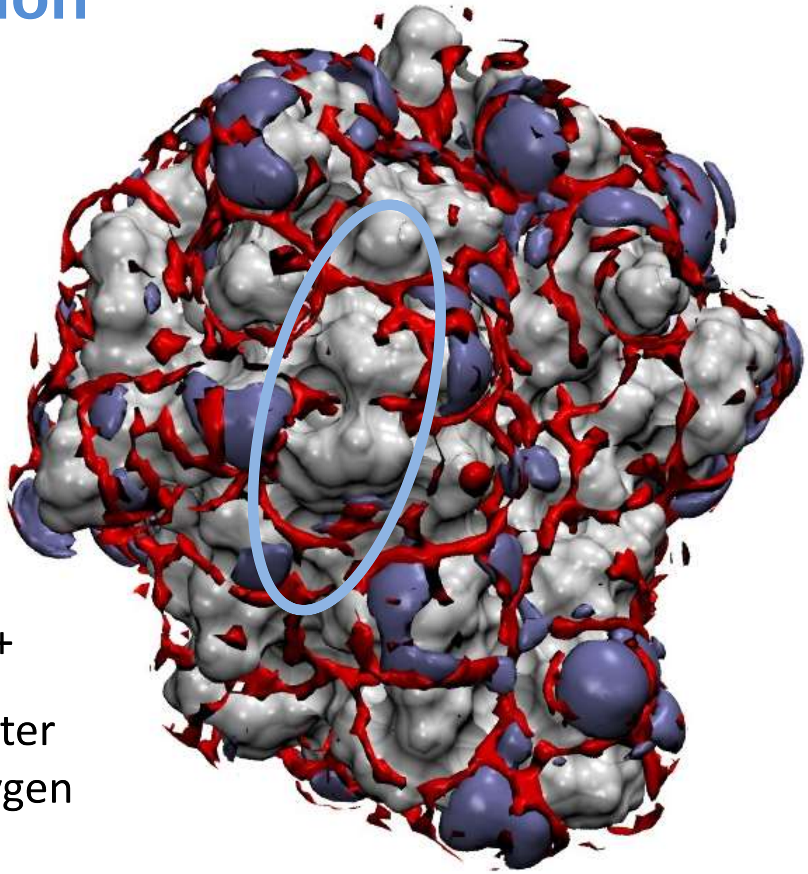
Blind docking with AutoDock Vina

Identification of structural solvent blocking unfavourable binding sites of SOD1 with 3D-RISM-KH molecular theory of solvation



▬ Na+

🟥 Water oxygen

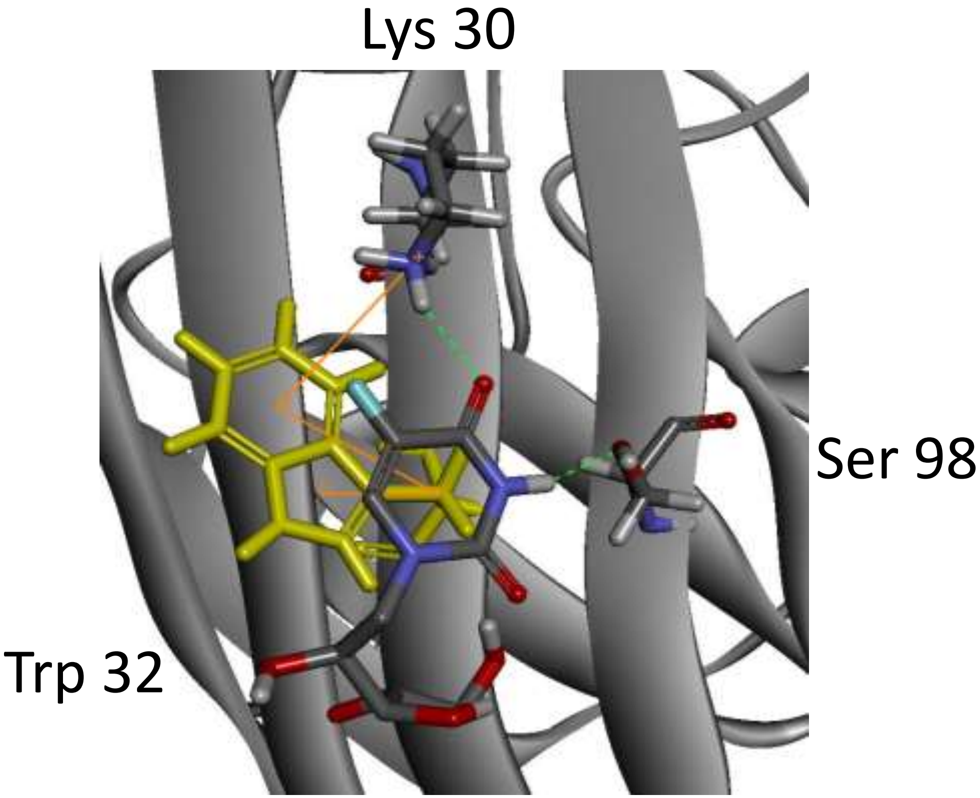


Sodium ions distribution
Density isolevel ρ=3 (times bulk density)

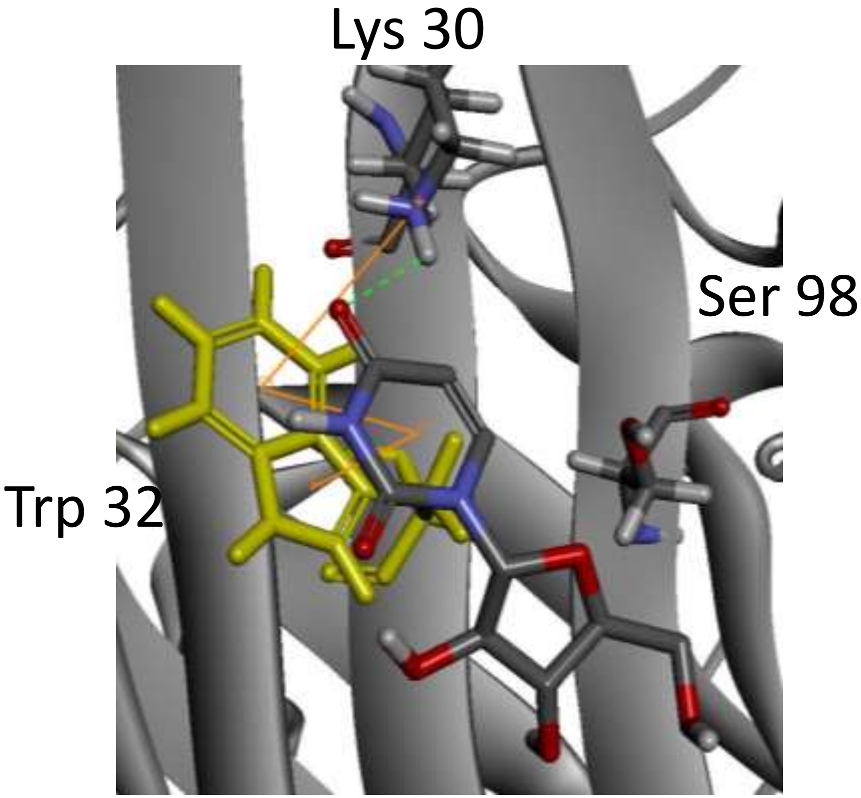
Water oxygen and Na+ ions distributions
Densities isolevel ρ=3

Experimental binding site is characterized by reduced solvent density: structural solvent need to be included into docking protocol

Top-ranked binding modes of 5-Fluorouridine and Uridine at SOD1 in the experimental cleft



5-FUrd, mode #1



Uridine, mode #2

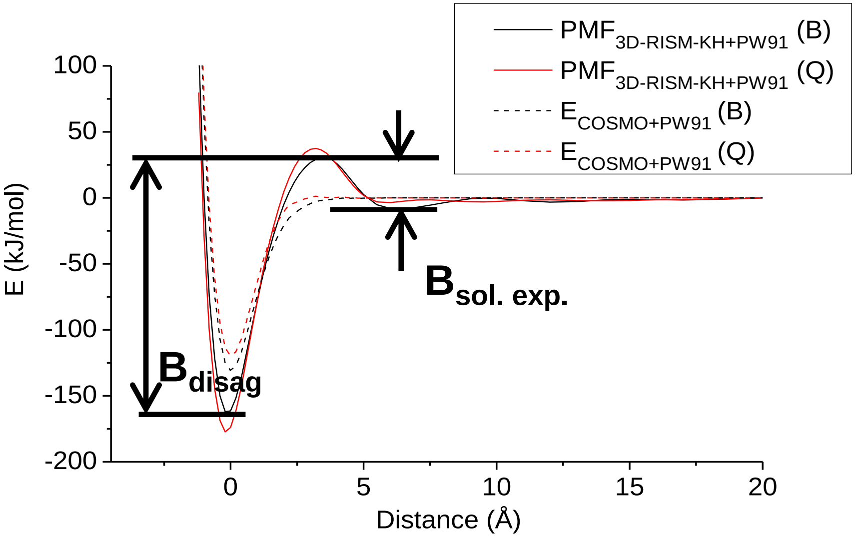
Adsorption from solution to solid surfaces by 3D-RISM-KH theory

Bitumen Fragment Adsorption on Chabazite: 3D-RISM-KH vs COSMO

$$\mathrm{PMF}(d)=\left(E_{\mathrm{PW91}}(d)+\mu_{\mathrm{solv}}(d)\right)-\left(E_{\mathrm{PW91}}(d_{\infty})+\mu_{\mathrm{solv}}(d_{\infty})\right)$$

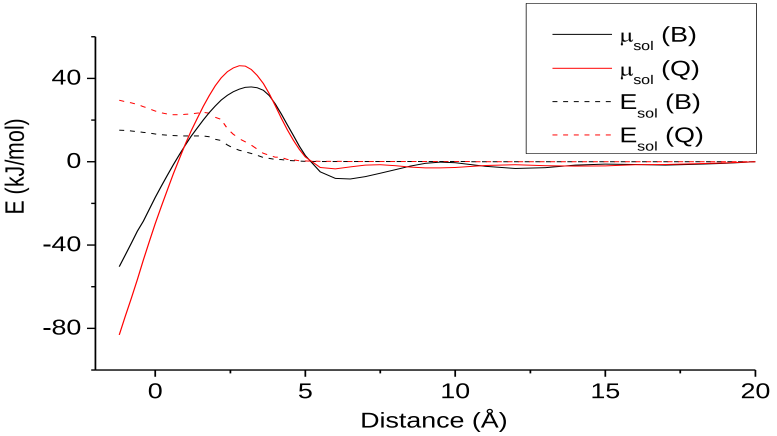
3D-RISM-KH molecular theory of solvation:

- **Accounts for solvation with accuracy comparable to explicit solvation model**
- **Predicts desolvation barriers that have to be overcome by bitumen molecules approaching solvated surface**
- **Predicts monolayer adsorption**



COSMO:

- **Continuum solvation method**
- **Accounts for solvent properties such as dielectric screening**

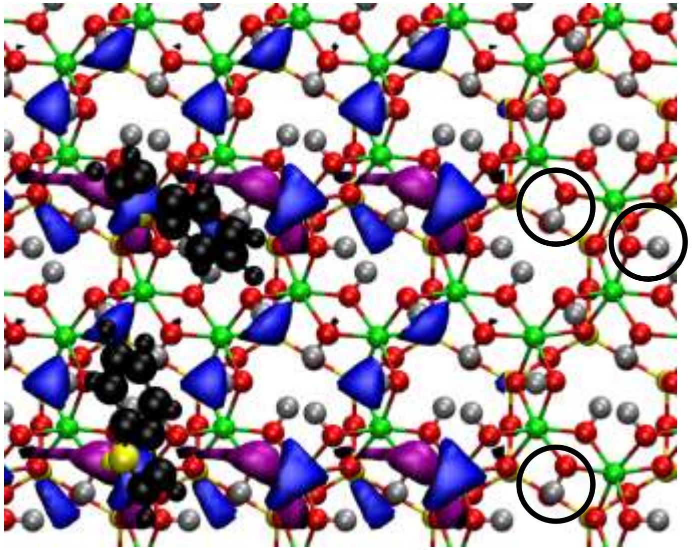


Thiophene disaggregation path

Optimized geometry

Stoyanov, S.R.; Gusarov, S.; Kovalenko, A. Multiscale Modeling of the Adsorption Interaction Between Model Bitumen Compounds and Zeolite Nanoparticles in Gas and Liquid Phase, in: *Industrial Applications of Molecular Simulations*, Meunier, M. (ed.); Taylor and Francis Books, Boca Raton, FL, USA, 2011, pp. 203-230.

Adsorption of indole on kaolinite surface from solution in toluene and in n-heptane

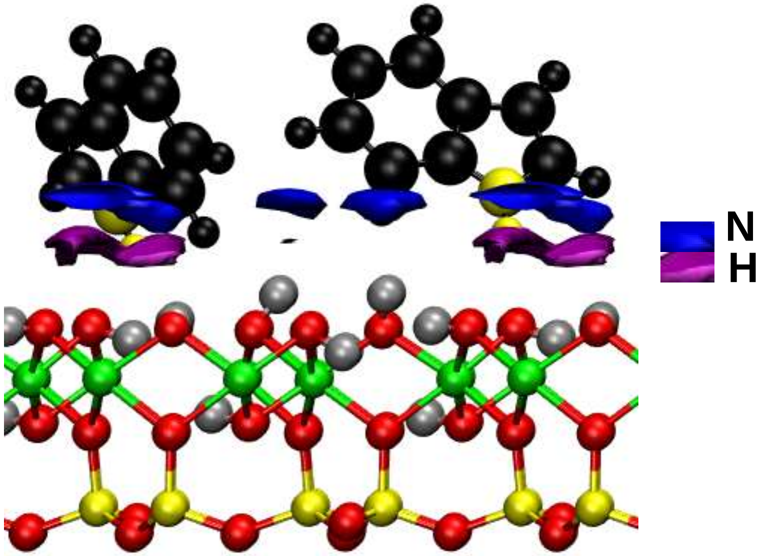


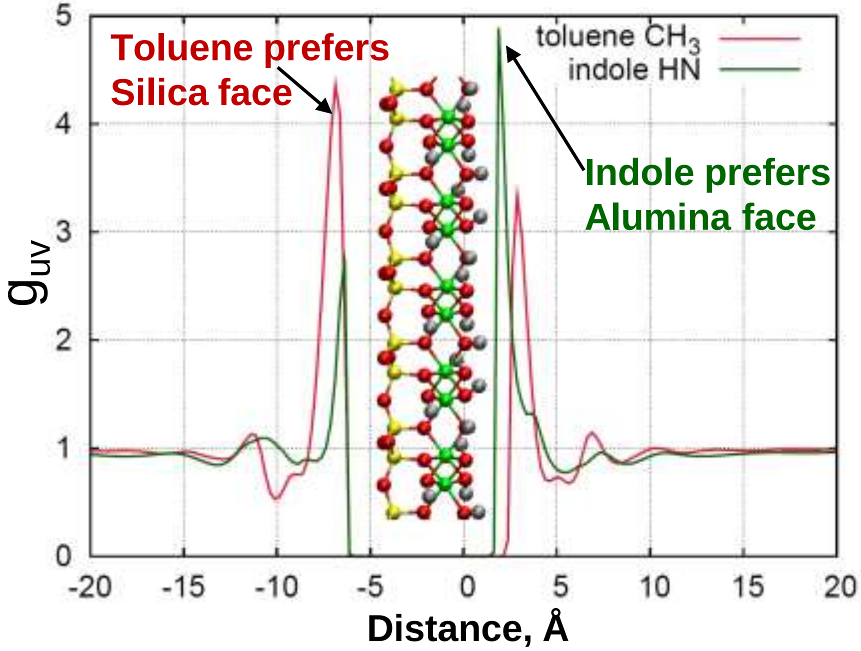
Aluminum hydroxide face

- Indole **HN** is oriented b/w two or one OH groups
- Orientation of toluene molecules is parallel

Silicon oxide face

- Indole molecules are almost parallel to the surface
- Toluene molecules are perpendicular to the surface, pointing with CH_3 site or ring H atoms towards SiO hexagon center



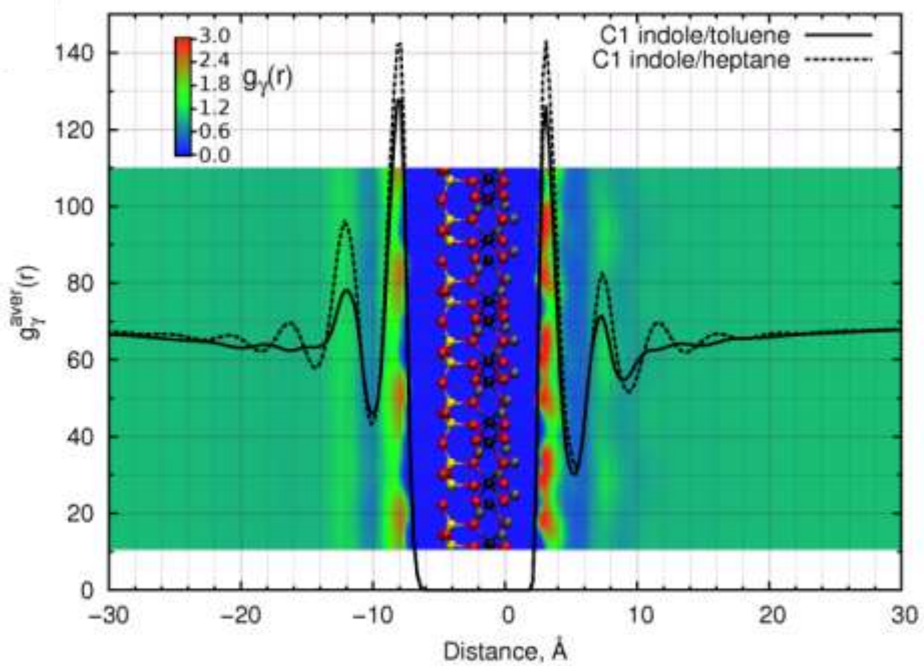


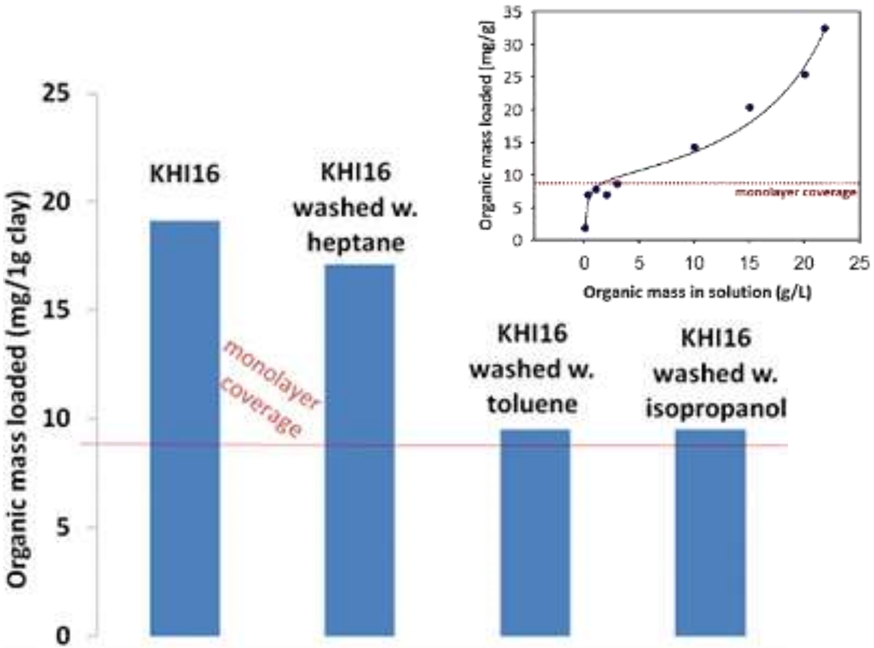
Toluene prefers Silica face

Indole prefers Alumina face

Fafard, J.; Lyubimova, O.; Stoyanov, S. R.; Kenne, G.; Gusarov, S.; Cuervo, J.; Kovalenko, A; Detellier, C. *J. Phys. Chem. C*, 2013, *117,* 18556

Multilayer adsorption of indole on kaolinite in toluene / n-heptane solution





Results

- $$g_\gamma^{\text{aver}}(z) = \int_{\text{surface}} \mathrm{dxdy}\; g_\gamma(x,y,z)$$

  predicts adsorption and depletion layers
- **Organic mass loading shows multilayer formation in organoclays**
- **Aromatic and hydrogen-bonding solvents leave only a monolayer of indole**

Prospectives:

○ **Rational design of additives for selective desorption of bitumen from kaolinite**

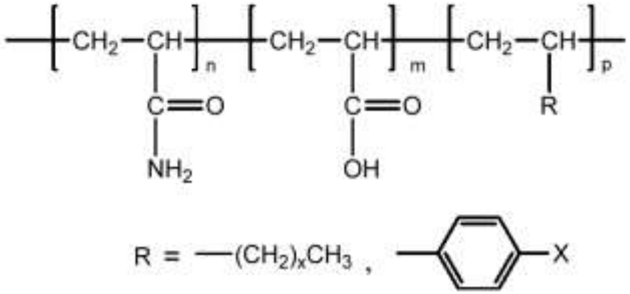
○ **Selection of green solvents for non-aqueous extraction of bitumen**

Fafard, J.; Lyubimova, O.; Stoyanov, S. R.; Kenne, G.; Gusarov, S.; Cuervo, J.; Kovalenko, A; Detellier, C. *J. Phys. Chem. C*, 2013, *117,* 18556

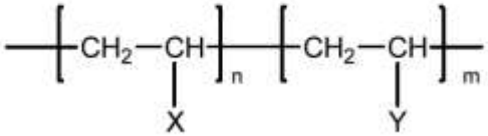
Effective interactions of clay nanocolloids in aqueous electrolyte solution with polymeric additives by 3D-RISM-KH theory

Background - Experiment

Experiment (S. Giasson, J. Zhu)

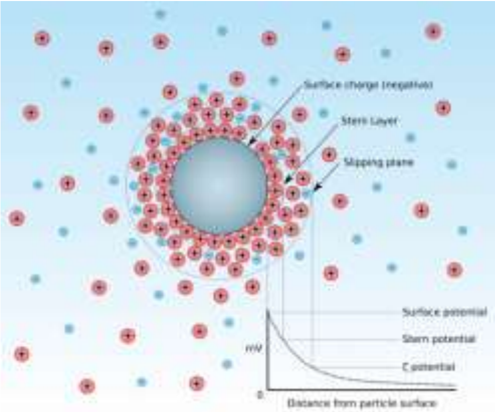


Copolymers of acrylamide, acrylic acid and hydrophobic comonomers. The R group on the 3rd monomer can be an alkyl or aromatic group, where X = H (styrene) or a polar group (such as hydroxystyrene).



$-[CH_2-CH(X)]_n-[CH_2-CH(Y)]_m-$

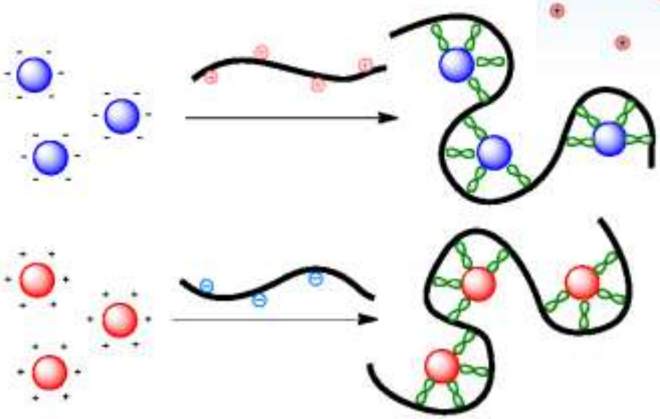
General structure of amphiphilic copolymer (either a random or a block copolymer) containing two vinilic comonomers with comparable reactivity but with different functional groups. X and Y are functional groups with different hydrophilicities.



Zeta potential will be determined experimentally from electrophoretic mobility measurements to provide estimates of the clay colloidal nanoparticle charges.

Light scattering will be used to estimate the structure and properties of clay nanoparticles so as to develop a coarse-grained clay force-field.

Possible ionic interactions between the functional groups of the polymers and the charged particles in aqueous media. Hydrogen bonds may also be formed. Upon association, the complexes become more hydrophobic than the ionic species and dehydration may take place to cause dewatering and sedimentation of the aggregates.



INSTITUTE FOR
OIL SANDS INNOVATION
AT THE UNIVERSITY OF ALBERTA

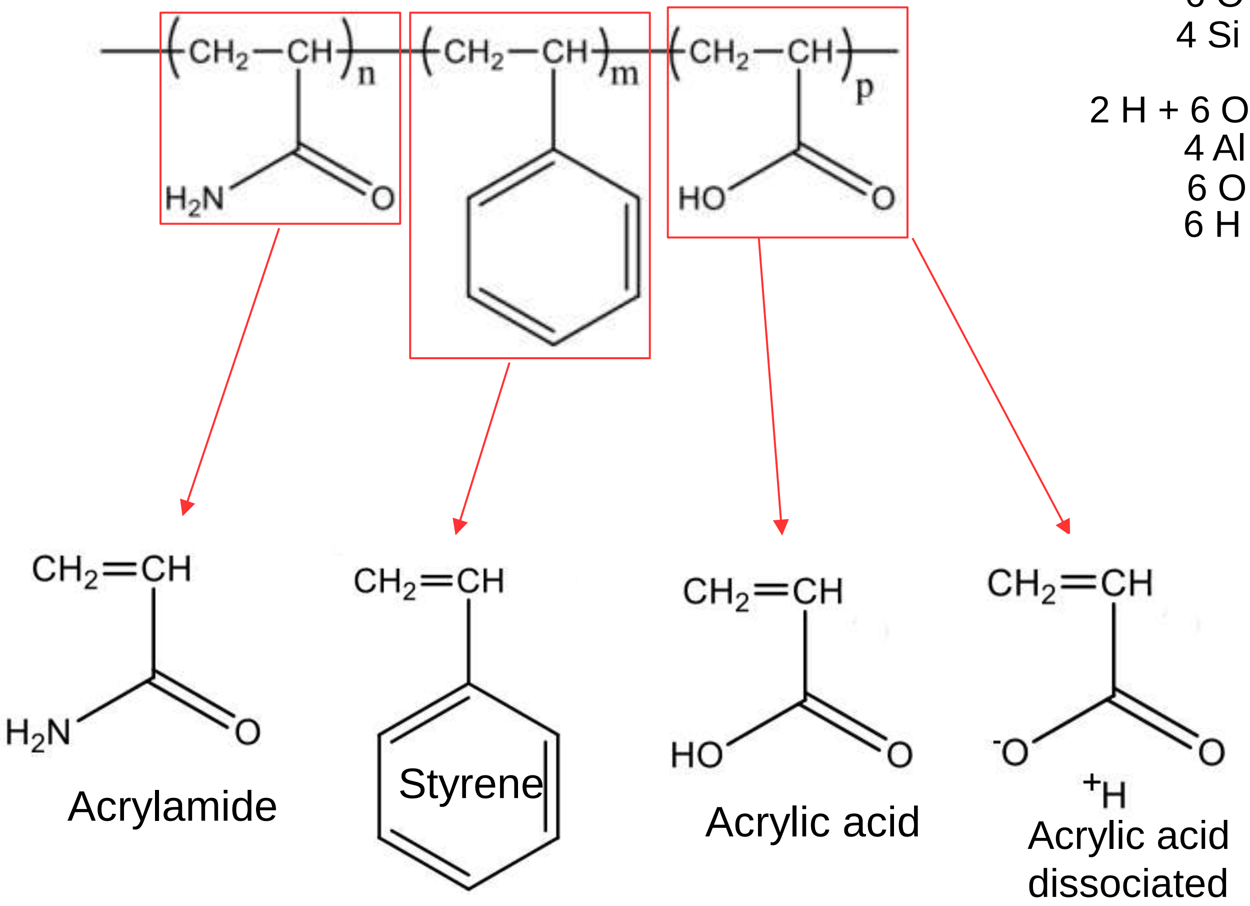
Imperial

Alberta Innovates
Energy and Environment Solutions

FACULTY OF ENGINEERING

Monomers and clay nanoparticle models

Water solution of salt and PAM monomers



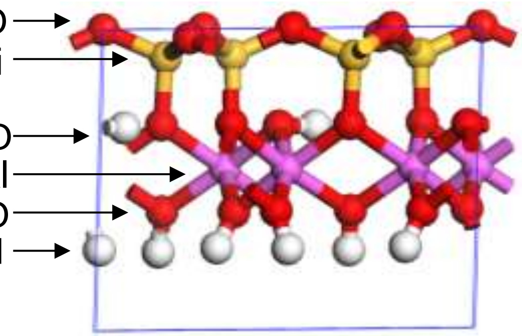
Acrylamide

Styrene

Acrylic acid

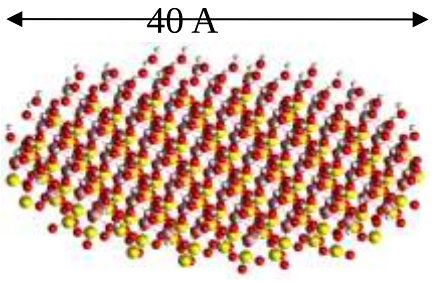
Acrylic acid dissociated

Kaolinite is the most abundant clay mineral in MFTs

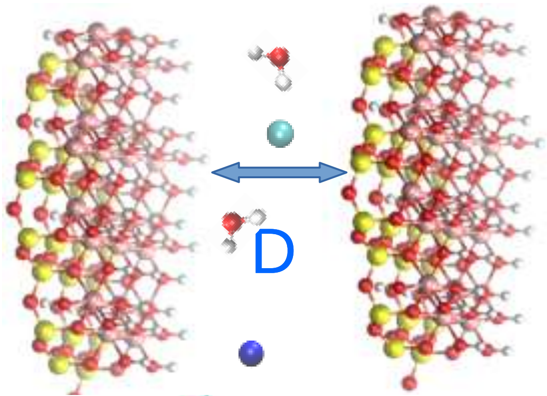


Kaolinite $Si_4Al_4O_{10}(OH)_8$

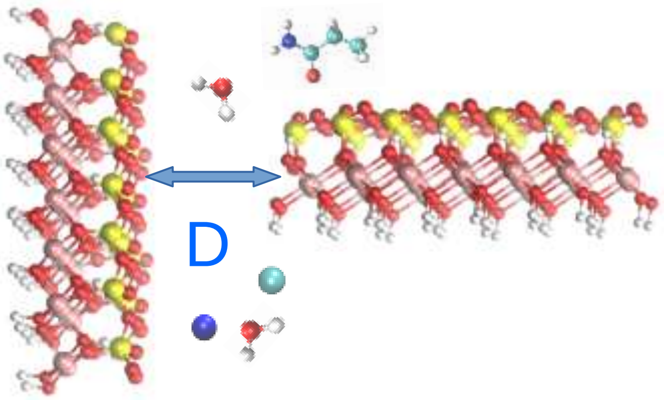
Colloidal suspension of single layer kaolinite platelets

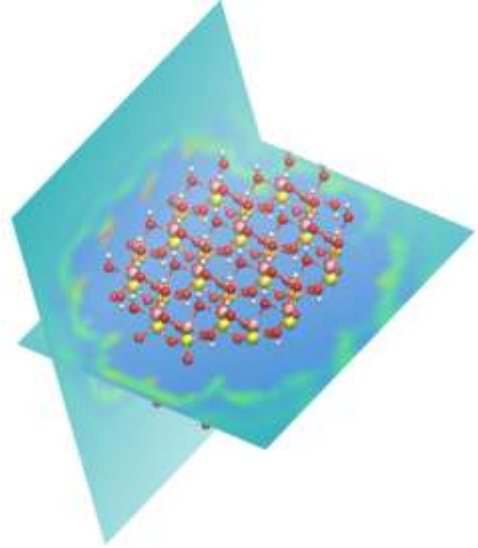


Atomistic model of clay nanocolloid for 3D-RISM-KH theory



Atomistic nanoparticles of kaolinite are used to compute the PMF between nanoparticles in different mutual arrangements and at various solvent compositions.





3D-RISM-KH molecular theory of solvation provides the solvation structure and thermodynamics, including PMFs that reflect molecular forces

Development of coarse-grained models of clay nanoparticles will based on the PMFs from 3D-RISM-KH theory is planned as a final results of the project.

Potential of Mean Force

$$U_{PMF}(D) = \mu_{solv}(D) - \mu_{solv}(\infty)$$

$$\mu_{solv} = \sum \int dr\,\Phi(r)$$

$$\Phi(r) = \rho_\gamma kT \left[\frac{1}{2} h_\gamma^2(r)\,\Theta\left(-h_\gamma(r)\right) - c_\gamma(r) - \frac{1}{2} h_\gamma(r)\,c_\gamma(r) \right]$$

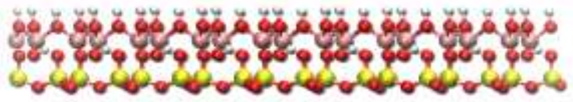
Results

- Preferential adsorption and binding of additives on specific faces and sites of clay nanoparticles
- Effect of salt and several additives on the PMFs as a descriptor for flocculation/aggregation phenomena.

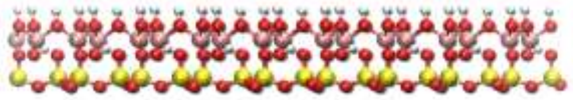
Arrangements of nanoparticles, monomers, water, and Na+ and Cl-

Platelets are assumed to be neutral, but posess a dipole moment oriented orthogonal to the surface.

Case 1



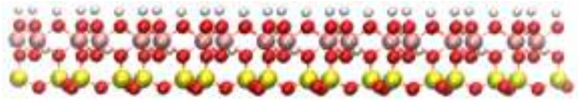
Hydrophilic-hydrophobic surfaces



Case 2

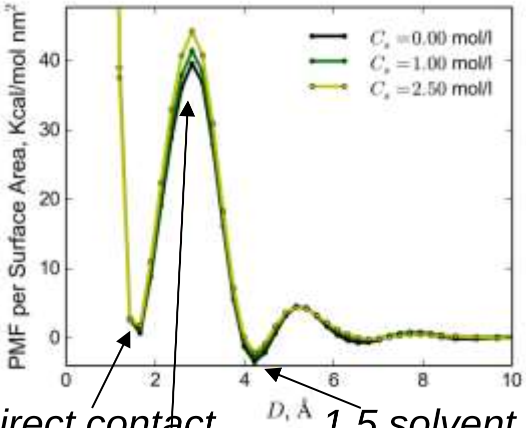
Hydrophilic-hydrophilic surfaces

Case 3



Hydrophobic-hydrophobic surfaces

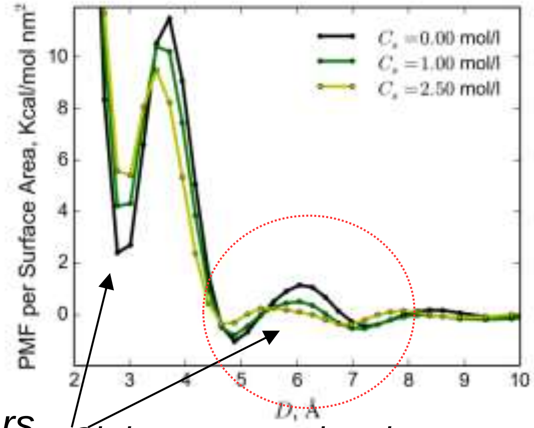
Results for aqueous NaCl solution without additives



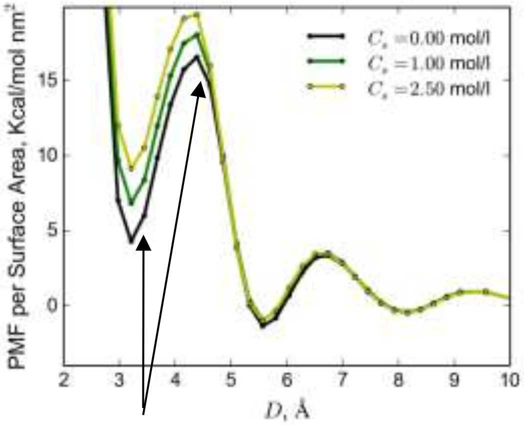
direct contact

one solvent layer

1.5 solvent layers



Cl- ions accumulate between the hydrophilic surfaces and disrupt the local order of water.



Na+ adsorbs strongly on the hydrophobic silica oxide surface
=>
the PMF in the contact area

Effect of monomer concentration on the PMFs

Acrylic acid prefers the hydrophobic silica oxide surface over the hydrophillic surface

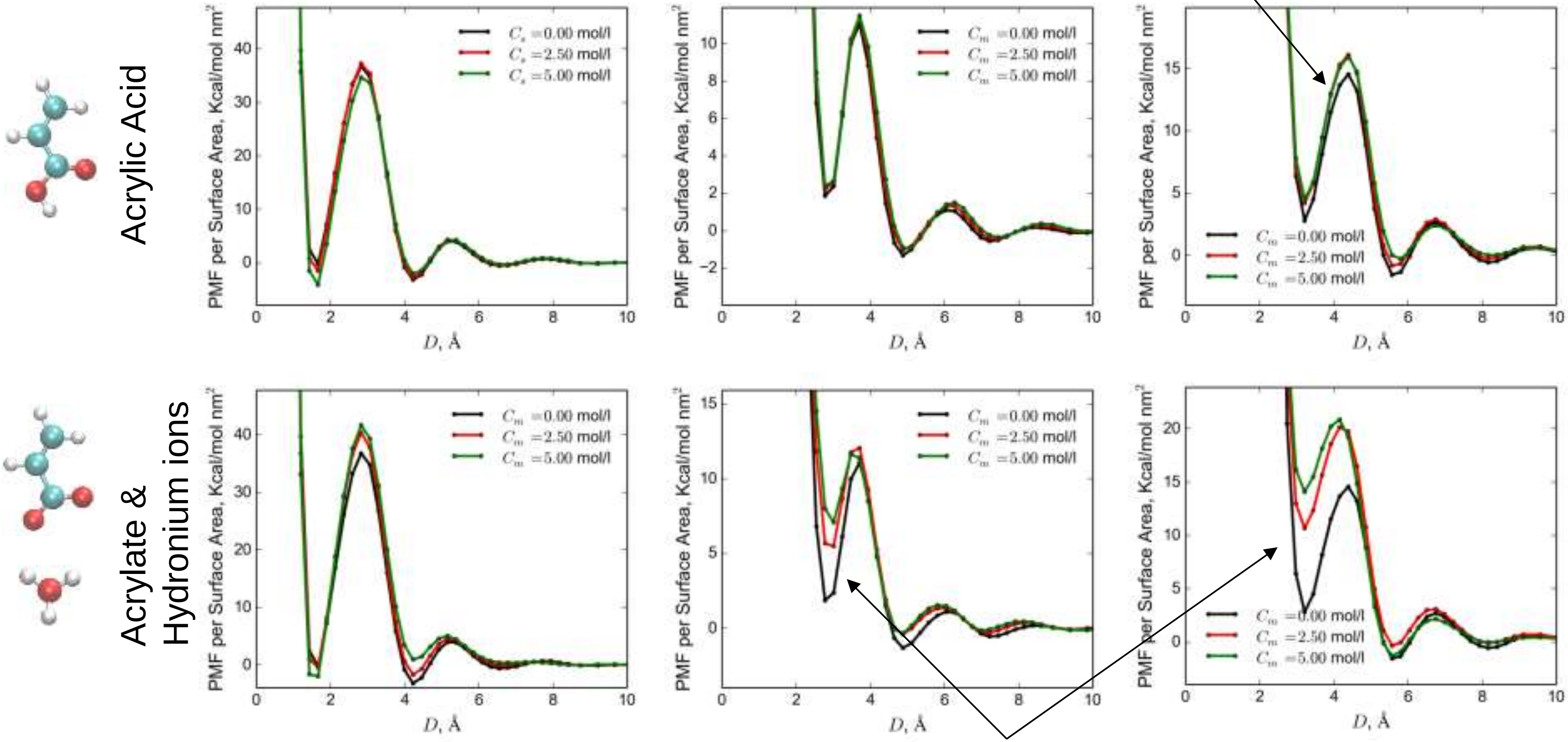
Case 1

Case 2

Case 3

Acrylic Acid

Acrylate & Hydronium ions



Acrylate and hydronium ions strongly adsorb => PMF increases

Multiscale coupling of DRISM-KH molecular theory of solvation with Dissipative Particle Dynamics (DPD) simulation

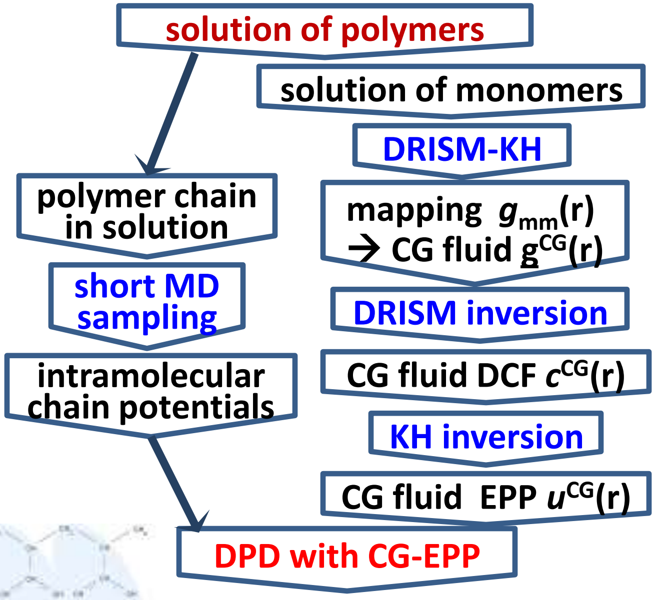
- **DRISM-KH is solved for all-atom species**
- **Coarse-grained effective pair potential (GC-EPP) is derived from all-atom RDFs by DRISM-KH invertion**
- **GC-EPP replaces Flory-Huggins heuristic conservative force in DPD**

Dissipative Particle Dynamics (DPD) simulation with a coarse-grained effective pair potential from DRISM-KH molecular theory of solvation

Kobryn, A. E.; Nikolić, D.; Lyubimova, O.; Gusarov, S.; Kovalenko, A.
J. Phys. Chem. B, **2014**, *118*, 12034

Coarse-Grained Effective Pair Potential from DRISM-KH theory:

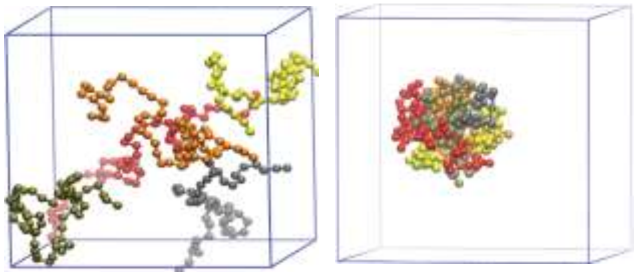
- Replaces the heuristically defined conservative force in DPD (Flory-Huggins)
- No restrictions on range, shape, equal volume of solute and solvent blobs
- Accounts for all chemical specificities from first principles of stat.mech.
- Flexible in specifying coarse-grained mapping of macromolecular species
- Enormously increases computational efficiency by eliminating solvent
- Applicable to polymer in melts, different solvent, and electrolyte solutions


solution of polymers
polymer chain in solution
short MD sampling
intramolecular chain potentials
solution of monomers
DRISM-KH
mapping $g_{mm}(r)$ → CG fluid $g^{CG}(r)$
DRISM inversion
CG fluid DCF $c^{CG}(r)$
KH inversion
CG fluid EPP $u^{CG}(r)$
DPD with CG-EPP

Validation: Polystyrene in Water and Toluene Solution

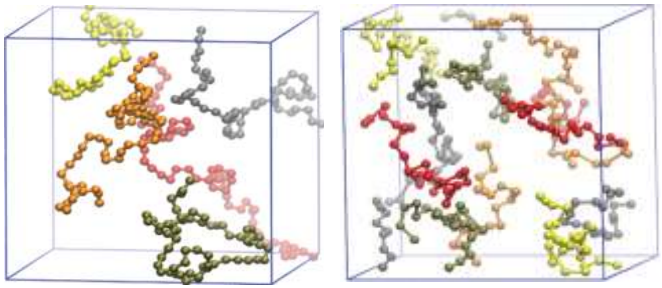
Flocculation in water

Initial snapshot **Resulting snapshot**

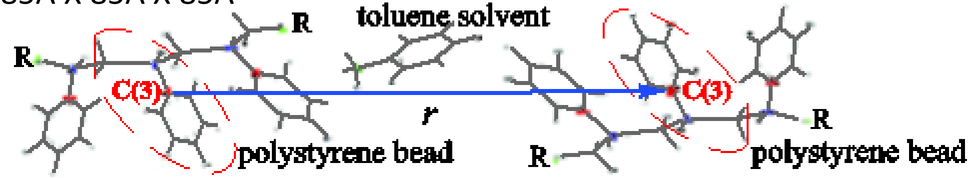


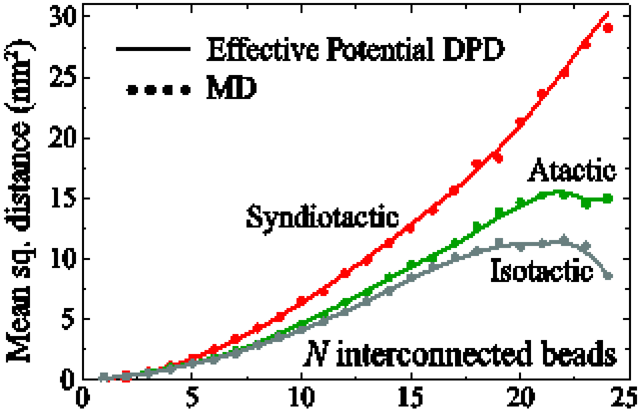
Dispersion in toluene

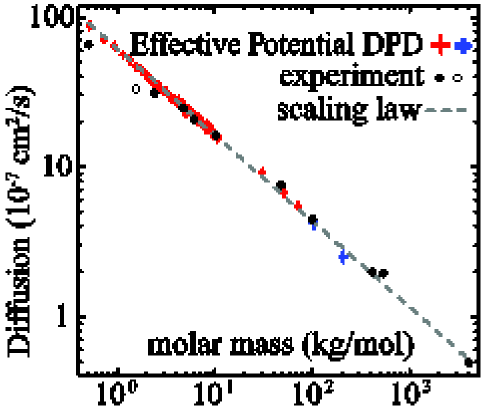
Initial snapshot **Resulting snapshot**



1 polymer chain = 60 beads
$V_{bead} = V_{monomer} \approx 5V_{H2O} \approx 150$ Å^3
$L_{Bead} \approx 8.0$ Å, $R_{Bead} = 3.3$ Å
System size: 85Å X 85Å X 85Å







Multiscale modeling of electronic and solvation structure: Self-consistent field, KS-DFT / 3D-RISM-KH theory

Multiscale modeling of electronic and solvation structure: Self-consistent field, KS-DFT / 3D-RISM-KH theory

A. Kovalenko and F. Hirata, JCP **110**, 10095 (1999).
A. Kovalenko, in: *Molecular Theory of Solvation*,
F. Hirata (ed.), Kluwer, 2003, pp. 169-213.

Free energy functional

$$A[n_e(\mathbf{r}),\{\rho_\gamma(\mathbf{r})\}] = E_{\text{solute}}[n_e(\mathbf{r})] + \Delta\mu_{\text{solvation}}[n_e(\mathbf{r}),\{\rho_\gamma(\mathbf{r})\}]$$

$$\Delta\mu^{\text{KH}}_{\text{solvation}} = \rho k_{\text{B}} T \sum_\gamma \int d\mathbf{r}\left[\tfrac{1}{2} h_\gamma^2(\mathbf{r})\Theta(-h_\gamma(\mathbf{r})) - \tfrac{1}{2} h_\gamma(\mathbf{r})c_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})\right]$$

Variational principle
Electron linear response
Solvent mean field

SCF equations

KS-DFT equation: $\left(-\tfrac{1}{2}\Delta + v_i(\mathbf{r}) + v_{\text{H}}(\mathbf{r}) + v_{\text{xc}}(\mathbf{r}) + v_{\text{solv}}(\mathbf{r})\right)\psi_j(\mathbf{r}) = \varepsilon_j \psi_j(\mathbf{r})$

3D-RISM equation: $h_\gamma(\mathbf{r}) = \sum_\alpha \int d\mathbf{r}'\, c_\alpha(\mathbf{r}-\mathbf{r}')\left(\omega_{\alpha\gamma}(r') + \rho_\alpha h_{\alpha\gamma}(r')\right)$

Kovalenko-Hirata closure relation: $g_\gamma(\mathbf{r}) = \exp\left(-\beta u_\gamma(\mathbf{r}) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})\right)$ for $g_\gamma(\mathbf{r}) \le 1$

$= 1 - \beta u_\gamma(\mathbf{r}) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})$ for $g_\gamma(\mathbf{r}) > 1$

Coupling potentials: $u^e_\gamma(\mathbf{r}) = \int d\mathbf{r}'\, n_e(\mathbf{r}-\mathbf{r}') v^{(\text{ps})}_\gamma(r')$ where $n_e(\mathbf{r}) = \sum \left|\psi_j(\mathbf{r})\right|^2$

$$v_{\text{solv}}(\mathbf{r}) = \sum_\gamma \int d\mathbf{r}'\, \rho_\gamma h_\gamma(\mathbf{r}-\mathbf{r}') v^{(\text{ps})}_\gamma(r')$$

KS-DFT / 3D-RISM-KH: Analytical Gradients

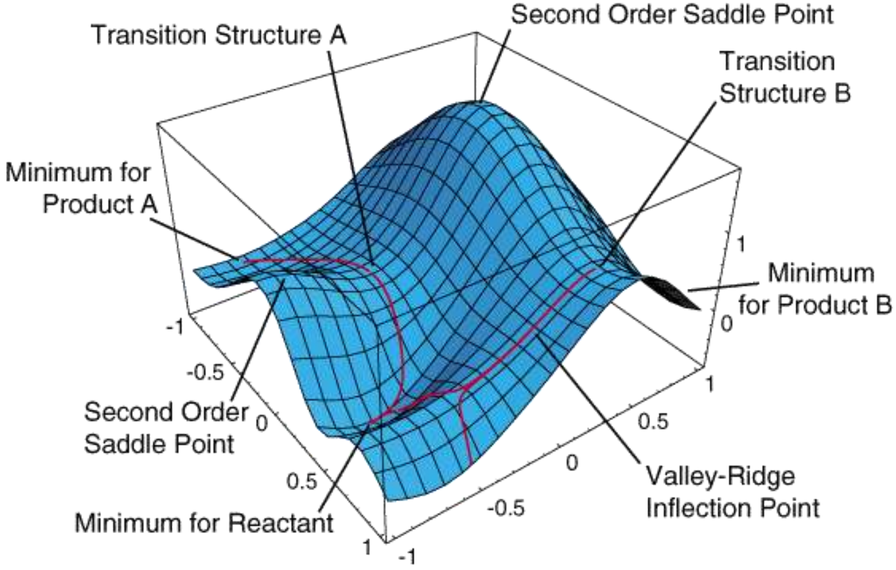
S. Gusarov, T. Ziegler, and A. Kovalenko, J. Phys. Chem. A **110**, 6083 (2006)

Derivatives of the free energy with respect to nuclei coordinates:

$$\frac{\partial(A)}{\partial \mathbf{R}} = \frac{\partial E_{\text{el}}}{\partial \mathbf{R}} + \frac{\partial \mu_{\text{solv}}}{\partial \mathbf{R}}$$

Gas-phase term

Solvation term



Using $$v_{\text{solv}}(r) = \frac{\delta \mu_{\text{solv}}}{\delta n_{\text{e}}(r)}$$

we have

$$\left(\frac{\partial A}{\partial \mathbf{R}_\alpha}\right)_{\text{solv}} = \frac{\partial(\mu_{\text{solv}})}{\partial \mathbf{R}_\alpha} = \int d\mathbf{r}\left[\rho^V g_\gamma(\mathbf{r})\frac{\partial u_\gamma^{(\text{sr})}(\mathbf{r})}{\partial \mathbf{R}_\alpha} + v_{\text{solv}}(\mathbf{r})\frac{\partial n_{\text{e}}(\mathbf{r})}{\partial \mathbf{R}_\alpha}\right] + q_\alpha \left.\frac{\partial v_{\text{solv}}(\mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{R}_\alpha}$$

Electronic & solvation structure of carbon monoxide in water: CASSCF / 3D-RISM-KH theory

Sato, Kovalenko, and Hirata, JCP **112**, 9463 (2000)

Fock operator with the solvation term

$$F_i^{(\mathrm{solv})} = F_i^{(0)} - f_i \langle V^{(\mathrm{solv})} \,|\, P \rangle$$

Solvation potential on solute electrons

$$V^{(\mathrm{solv})}(\mathbf{r}) = \rho_\gamma h_\gamma^{(\mathrm{solv})}(\mathbf{r}) * v_\gamma^{(\mathrm{ps})}(r)$$

Classical potential on solvent site γ

$$u_\gamma(\mathbf{r}) = u_\gamma^{(\mathrm{sr})}(\mathbf{r}) + q_\gamma \left(\phi_\gamma^{(\mathrm{n})}(\mathbf{r}) - \phi_\gamma^{(\mathrm{e})}(\mathbf{r}) \right)$$

Electron electrostatic potential

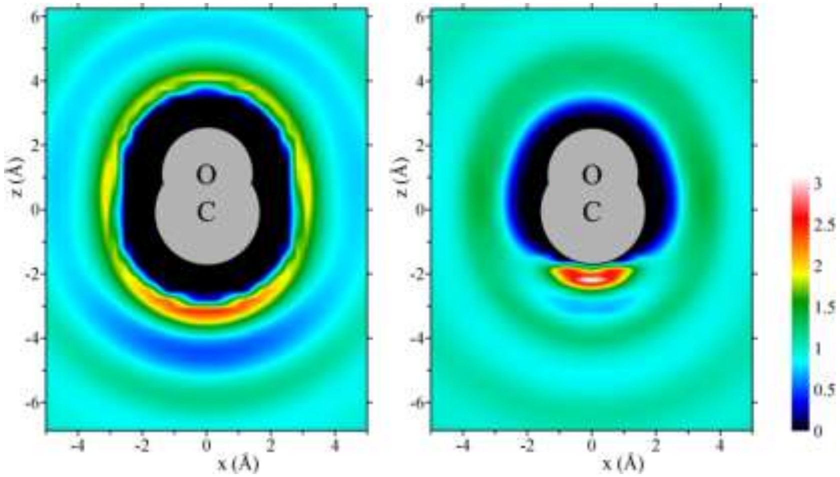
$$-k^2 \phi^{(\mathrm{e})}(\mathbf{k}) = 4\pi \rho^{(\mathrm{e})}(\mathbf{k})$$

Electron density on the reciprocal grid in terms of Gaussian orbitals

$$\begin{aligned}\rho^{(\mathrm{e})}(\mathbf{k}) &= \int \mathrm{d}\mathbf{r}\, \rho^{(\mathrm{e})}(\mathbf{r}) \exp(i\mathbf{k}\cdot\mathbf{r}) \\ &= \int \mathrm{d}\mathbf{r} \sum_{\mu\nu} D_{\mu\nu} \chi_\mu(\mathbf{r}) \chi_\nu^*(\mathbf{r}) \exp(i\mathbf{k}\cdot\mathbf{r}) \\ &= \sum_{\mu\nu} D_{\mu\nu} I_{\mu\nu}(\mathbf{k})\end{aligned}$$

Hydration structure of the CO molecule

water Oxygen   water Hydrogen



CASSCF with
2 core, 8 active orbitals, 10 electrons
in DZP basis set (9*s*5*p*1*d*/4*s*2*p*1*d*)

Multiscale modeling of nanochemistry in solution: (OFE)KS-DFT / 3D-RISM-KH, including analytical gradients

S. Gusarov, T. Ziegler, A. Kovalenko, J.Phys.Chem.A, **110**, 6083 (2006)
D. Casanova, S. Gusarov, A. Kovalenko, T. Ziegler, J.Chem.Theory Comput. **3**, 458 (2007)

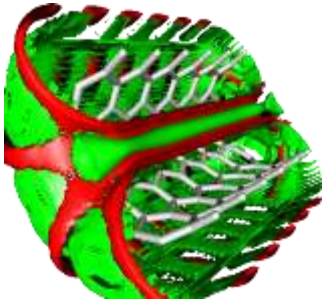
IMPLEMENTATION

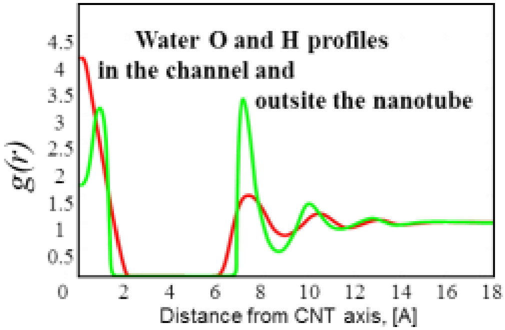
- **Amsterdam Density Functional (ADF) quantum chemistry package**
- **SCF *ab initio* KS-DFT / 3D-RISM-KH theory of molecular solvation at the Car-Parrinello level**
- **3D resolution of molecular solvation shells at molecular simulations level, but with NO noice**
- **Prediction of solvation physics and chemistry, including inner nanospaces**
- **Analytical gradients including solvation: geometry optimization, transition states, reactions and (nano)catalysis in solution**

Implemented in
ADF Computational Chemistry package:

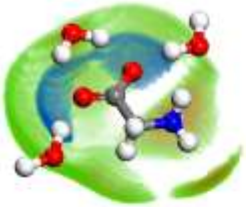
- Optimal structure
- Properties (NMR,CD,IR,…)
- Reactions / Nanocatalysis
- Spectroscopic / Photochemical

Solvation Structure and Thermodynamics



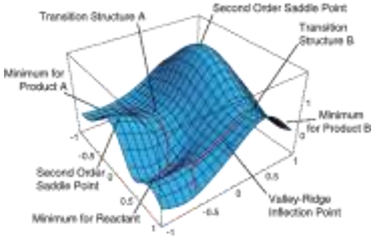


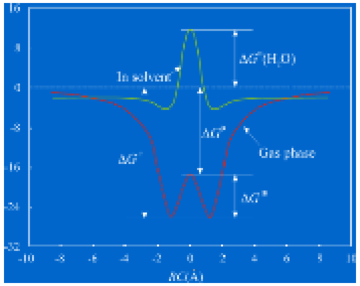
Analytical gradients: Gas phase + Solvation



$$\frac{\partial(A)}{\partial \mathbf{R}} = \frac{\partial E_{\text{el}}}{\partial \mathbf{R}} + \frac{\partial \mu_{\text{solv}}}{\partial \mathbf{R}}$$

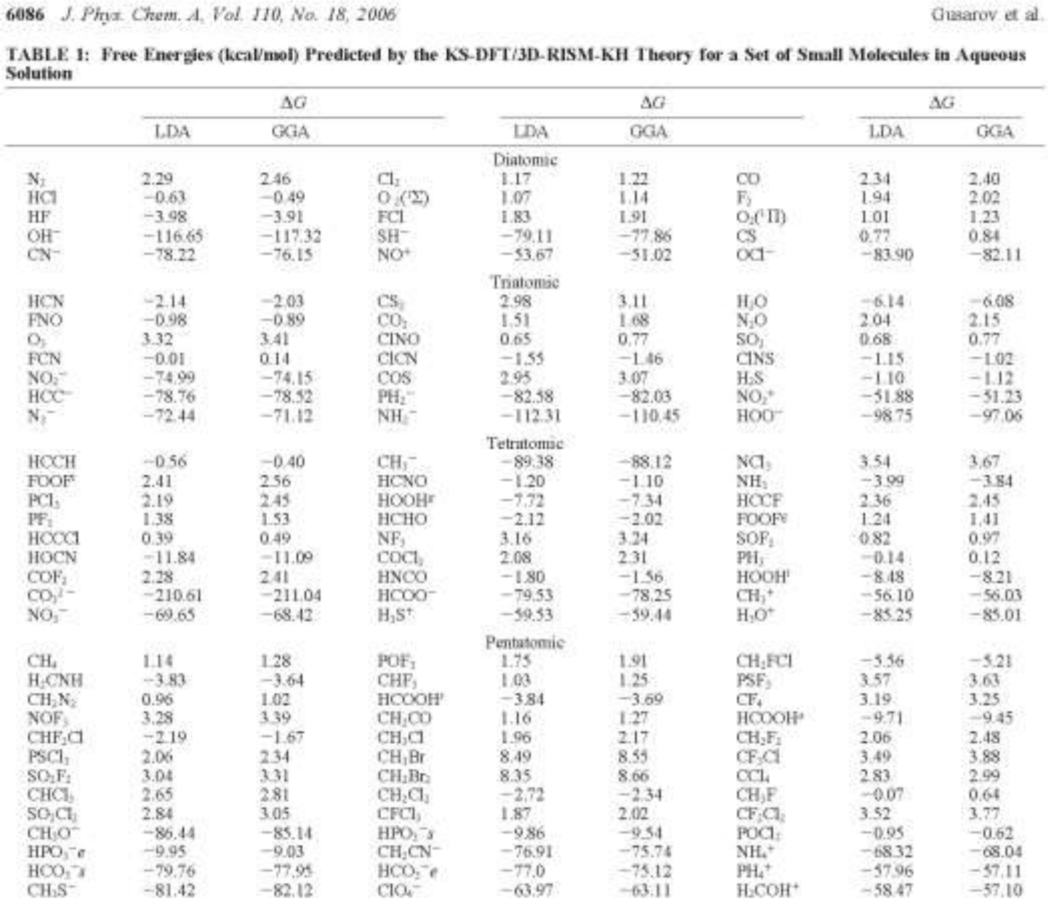
Potential Energy States (PES) analysis:





Hydration free energy: KS-DFT/3D-RISM-KH vs Experiment

S. Gusarov, T. Ziegler, A. Kovalenko,
J. Phys. Chem. A, **110**, 6083 (2006)



6086 *J. Phys. Chem. A, Vol. 110, No. 18, 2006* Gusarov et al.

TABLE 1: Free Energies (kcal/mol) Predicted by the KS-DFT/3D-RISM-KH Theory for a Set of Small Molecules in Aqueous Solution

	ΔG			ΔG			ΔG	
	LDA	GGA		LDA	GGA		LDA	GGA
				Diatomic				
N_2	2.29	2.46	Cl_2	1.17	1.22	CO	2.34	2.40
HCl	−0.63	−0.49	$O_2(^1\Sigma)$	1.07	1.14	F_2	1.94	2.02
HF	−3.98	−3.91	FCl	1.83	1.91	$O_2(^1\Pi)$	1.01	1.23
OH^-	−116.65	−117.32	SH^-	−79.11	−77.86	CS	0.77	0.84
CN^-	−78.22	−76.15	NO^+	−53.67	−51.02	OCl^-	−83.90	−82.11
				Triatomic				
HCN	−2.14	−2.03	CS_2	2.98	3.11	H_2O	−6.14	−6.08
FNO	−0.98	−0.89	CO_2	1.51	1.68	N_2O	2.04	2.15
O_3	3.32	3.41	ClNO	0.65	0.77	SO_2	0.68	0.77
FCN	−0.01	0.14	ClCN	−1.55	−1.46	ClNS	−1.15	−1.02
NO_2^-	−74.99	−74.15	COS	2.95	3.07	H_2S	−1.10	−1.12
HCC^-	−78.76	−78.52	PH_2^-	−82.58	−82.03	NO_2^+	−51.88	−51.23
N_3^-	−72.44	−71.12	NH_2^-	−112.31	−110.45	HOO^-	−98.75	−97.06
				Tetratomic				
HCCH	−0.56	−0.40	CH_3^-	−89.38	−88.12	NCl_3	3.54	3.67
$FOOF^f$	2.41	2.56	HCNO	−1.20	−1.10	NH_3	−3.99	−3.84
PCl_3	2.19	2.45	$HOOH^g$	−7.72	−7.34	HCCF	2.36	2.45
PF_3	1.38	1.53	HCHO	−2.12	−2.02	$FOOF^g$	1.24	1.41
HCCCl	0.39	0.49	NF_3	3.16	3.24	SOF_2	0.82	0.97
HOCN	−11.84	−11.09	$COCl_2$	2.08	2.31	PH_3	−0.14	0.12
COF_2	2.28	2.41	HNCO	−1.80	−1.56	$HOOH^i$	−8.48	−8.21
CO_3^{2-}	−210.61	−211.04	$HCOO^-$	−79.53	−78.25	CH_3^+	−56.10	−56.03
NO_3^-	−69.65	−68.42	H_3S^+	−59.53	−59.44	H_3O^+	−85.25	−85.01
				Pentatomic				
CH_4	1.14	1.28	POF_3	1.75	1.91	CH_2FCl	−5.56	−5.21
H_2CNH	−3.83	−3.64	CHF_3	1.03	1.25	PSF_3	3.57	3.63
CH_2N_2	0.96	1.02	$HCOOH^i$	−3.84	−3.69	CF_4	3.19	3.25
NOF_3	3.28	3.39	CH_2CO	1.16	1.27	$HCOOH^a$	−9.71	−9.45
CHF_2Cl	−2.19	−1.67	CH_3Cl	1.96	2.17	CH_2F_2	2.06	2.48
$PSCl_3$	2.06	2.34	CH_3Br	8.49	8.55	CF_3Cl	3.49	3.88
SO_2F_2	3.04	3.31	CH_2Br_2	8.35	8.66	CCl_4	2.83	2.99
$CHCl_3$	2.65	2.81	CH_2Cl_2	−2.72	−2.34	CH_3F	−0.07	0.64
SO_2Cl_2	2.84	3.05	$CFCl_3$	1.87	2.02	CF_2Cl_2	3.52	3.77
CH_3O^-	−86.44	−85.14	HPO_3^-s	−9.86	−9.54	$POCl_3$	−0.95	−0.62
HPO_3^-e	−9.95	−9.03	CH_2CN^-	−76.91	−75.74	NH_4^+	−68.32	−68.04
HCO_3^-a	−79.76	−77.95	HCO_2^-e	−77.0	−75.12	PH_4^+	−57.96	−57.11
CH_3S^-	−81.42	−82.12	ClO_4^-	−63.97	−63.11	H_2COH^+	−58.47	−57.10

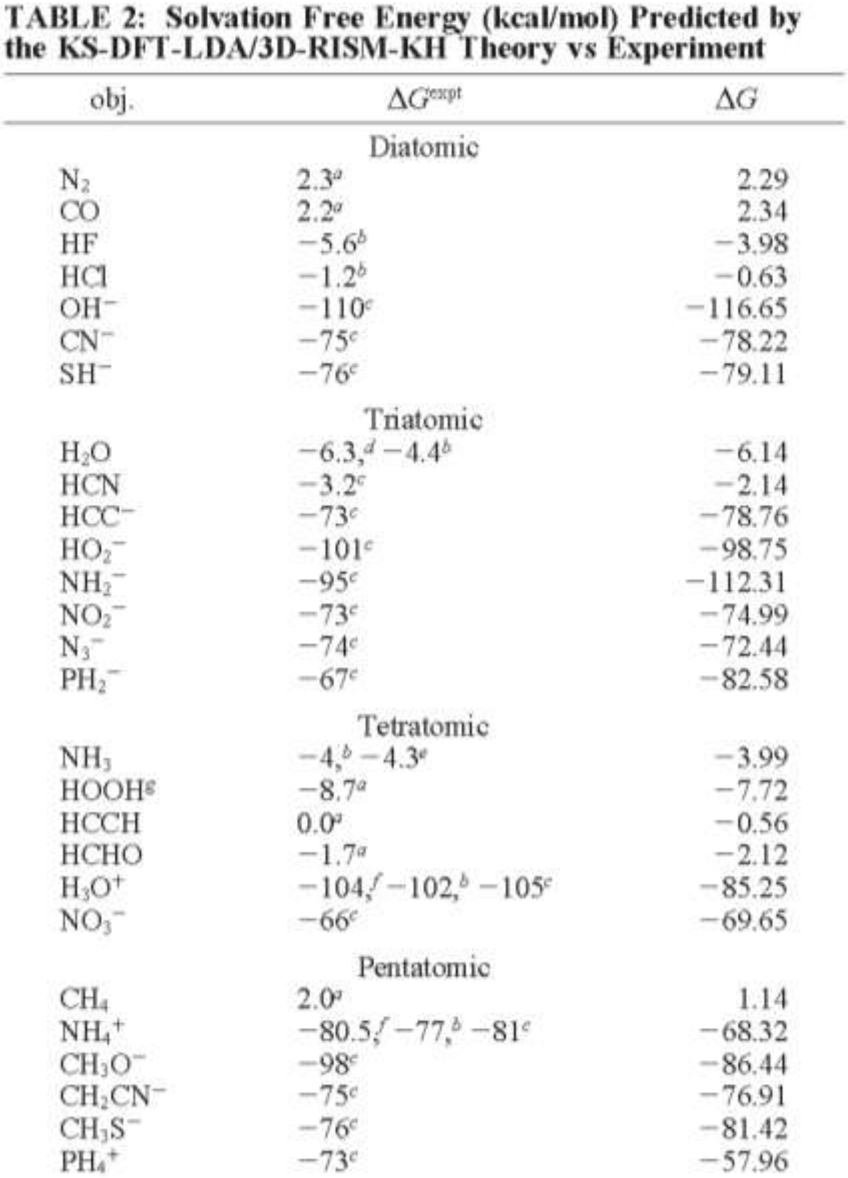
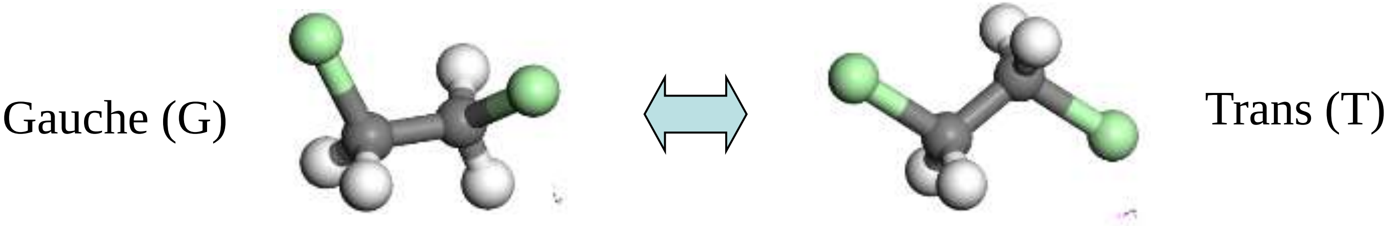


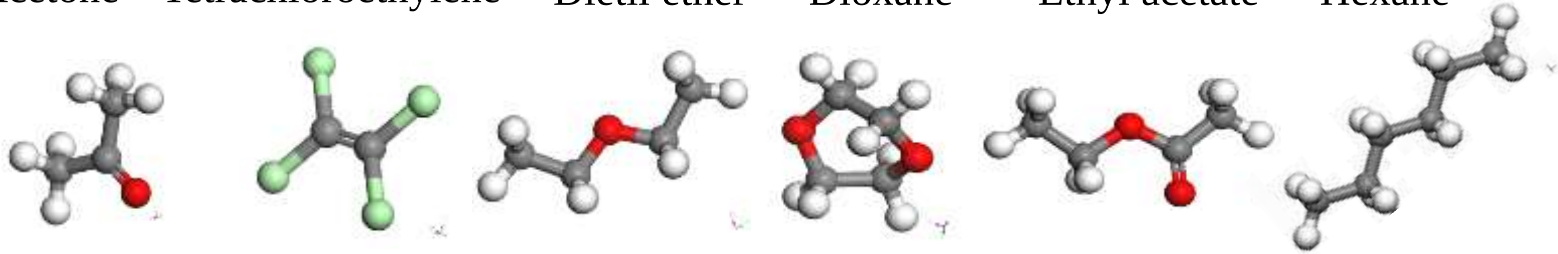
TABLE 2: Solvation Free Energy (kcal/mol) Predicted by the KS-DFT-LDA/3D-RISM-KH Theory vs Experiment

obj.	ΔG^{expt}	ΔG
	Diatomic	
N_2	2.3^a	2.29
CO	2.2^a	2.34
HF	-5.6^b	−3.98
HCl	-1.2^b	−0.63
OH^-	-110^c	−116.65
CN^-	-75^c	−78.22
SH^-	-76^c	−79.11
	Triatomic	
H_2O	$-6.3,^d -4.4^b$	−6.14
HCN	-3.2^c	−2.14
HCC^-	-73^c	−78.76
HO_2^-	-101^c	−98.75
NH_2^-	-95^c	−112.31
NO_2^-	-73^c	−74.99
N_3^-	-74^c	−72.44
PH_2^-	-67^c	−82.58
	Tetratomic	
NH_3	$-4,^b -4.3^e$	−3.99
$HOOH^g$	-8.7^a	−7.72
HCCH	0.0^a	−0.56
HCHO	-1.7^a	−2.12
H_3O^+	$-104,^f -102,^b -105^c$	−85.25
NO_3^-	-66^c	−69.65
	Pentatomic	
CH_4	2.0^a	1.14
NH_4^+	$-80.5,^f -77,^b -81^c$	−68.32
CH_3O^-	-98^c	−86.44
CH_2CN^-	-75^c	−76.91
CH_3S^-	-76^c	−81.42
PH_4^+	-73^c	−57.96

Solvent effect on conformational equilibria of 1,2-dichloroethane: KS-DFT / 3D-RISM-KH

D. Casanova, S. Gusarov, A. Kovalenko, T. Ziegler, J. Chem. Theory Comput., **3**, 458 (2007)

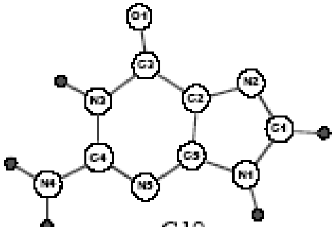




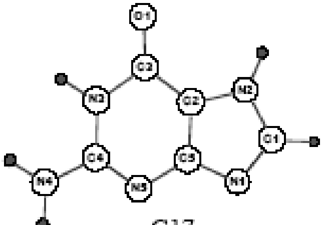
Solvent:	Acetone	Tetrachloroethylene	Dietil ether	Dioxane	Ethyl acetate	Hexane
COSMO	**0.40**	**1.18**	**0.84**	**1.28**	**0.70**	**1.34**
3D-RISM-KH	**0.16**	**1.20**	**0.54**	**0.04**	**0.63**	**1.57**
PCM	**0.13**	**0.79**	**0.61**	**0.82**	**0.48**	**0.95**
Experiment	**0.18**	**0.95**	**0.77**	**0.54**	**0.50**	**1.15**

KS-DFT / 3D-RISM-KH for solvent effect on tautomerism: Guanine

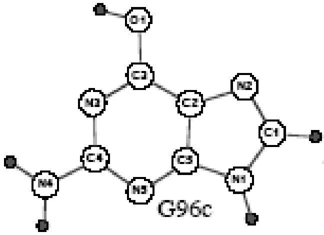
D. Casanova, S. Gusarov, A. Kovalenko, T. Ziegler,
J. Chem. Theory Comput., **3**, 458 (2007)



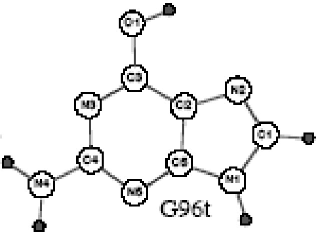
G19



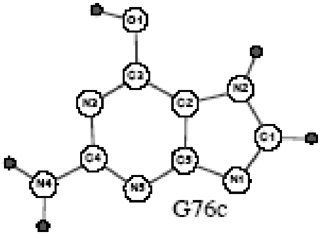
G17



G96c



G96t



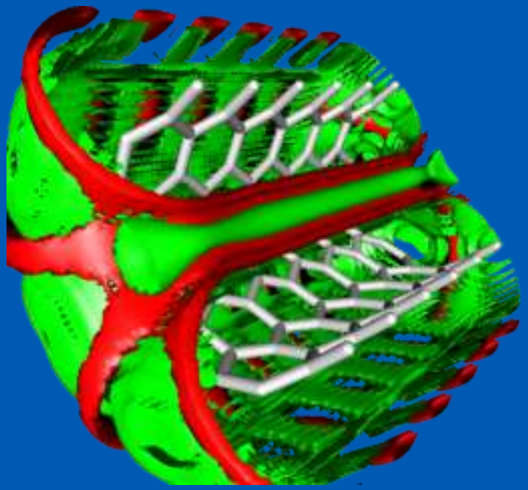
G76c

Relative free energies between guanine tautomers in aqueous solution (kcal/mol)

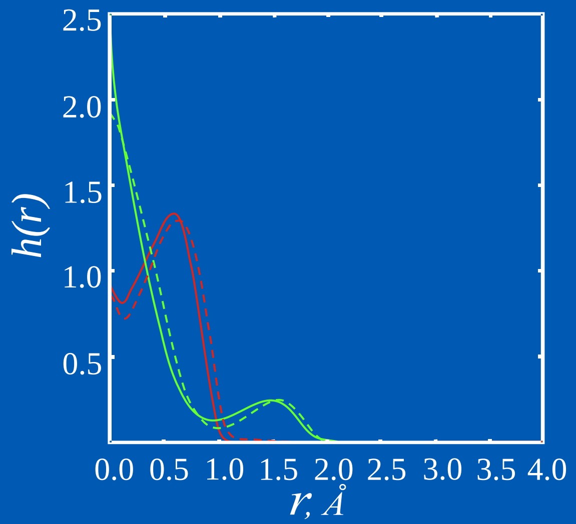
	COSMO	3D-RISM	Colomians[a]
G19	0.0	0.0	0.0
G17	0.40	1.10	1.00
G96c	10.11	7.33	7.20
G96t	10.57	7.18	8.00
G97c	10.94	9.32	8.80

[a] *J. Am. Chem. Soc.* **1996**, *118*, 6811

Water density profiles inside (6,6) carbon nanotube.
KS-DFT/3D-RISM-KH (dashed lines & 3D) vs MD (solid lines)

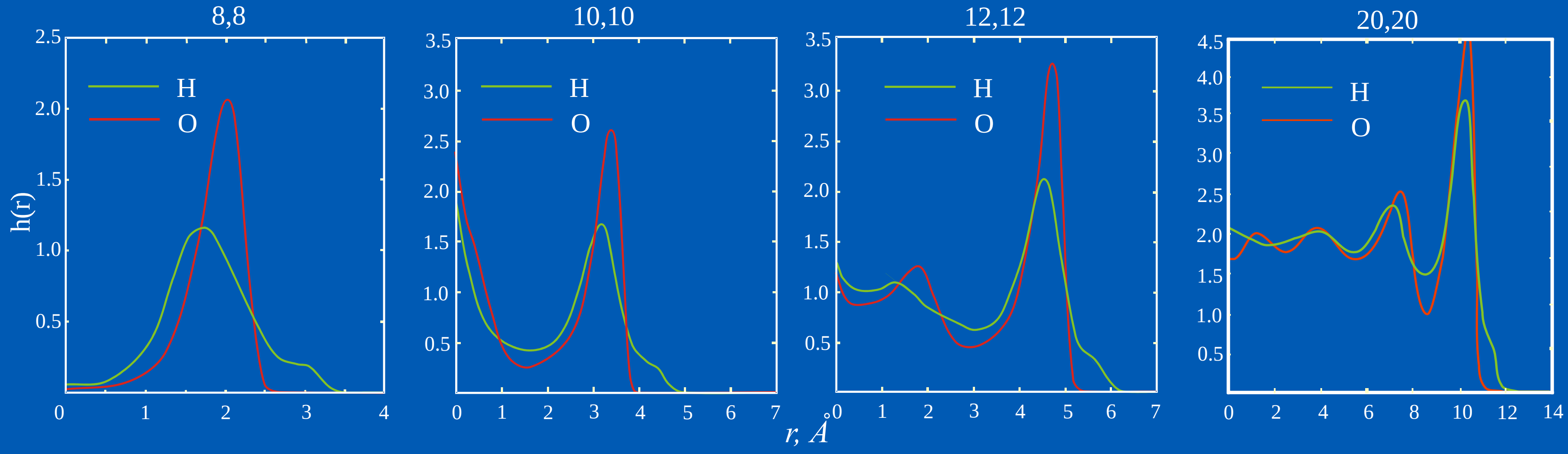


Ab initio DFT and 3D molecular theory of solvation:
D. Casanova, S. Gusarov, A. Kovalenko, T. Ziegler,
J. Chem. Theory Comput., **3**, 458 (2007)



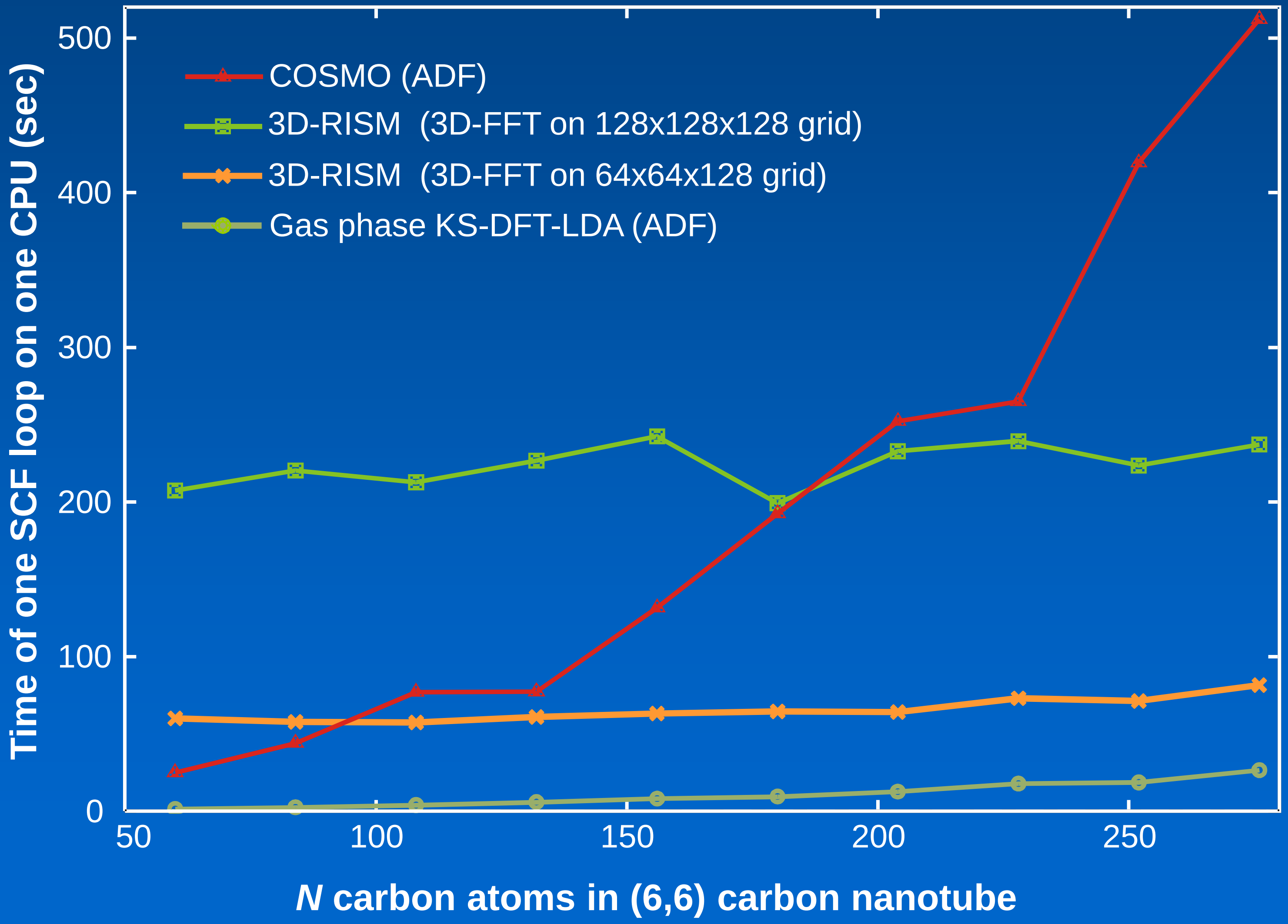
MD (dashed lines): Chem. Phys. Lett. **329**, 341 (2000);
SPC water and (6,6) CNT with soft potential of Lennard-Jones type

Density profiles of water oxygen and hydrogen inside the carbon nanotube



Timings of KS-DFT-LDA with different solvation schemes: 3D-RISM-KH molecular theory vs COSMO continuum model

S. Gusarov, T. Ziegler, and A. Kovalenko, J. Phys. Chem. A **110**, 6083 (2006)



Scales as N^2

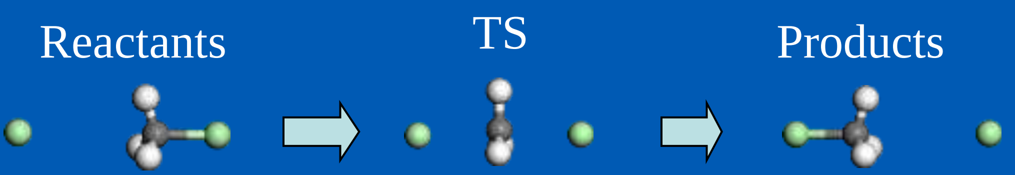
- number of COSMO cavity surface charges $M=N^{2/3}$
- charges matrix inversion ~ M^3

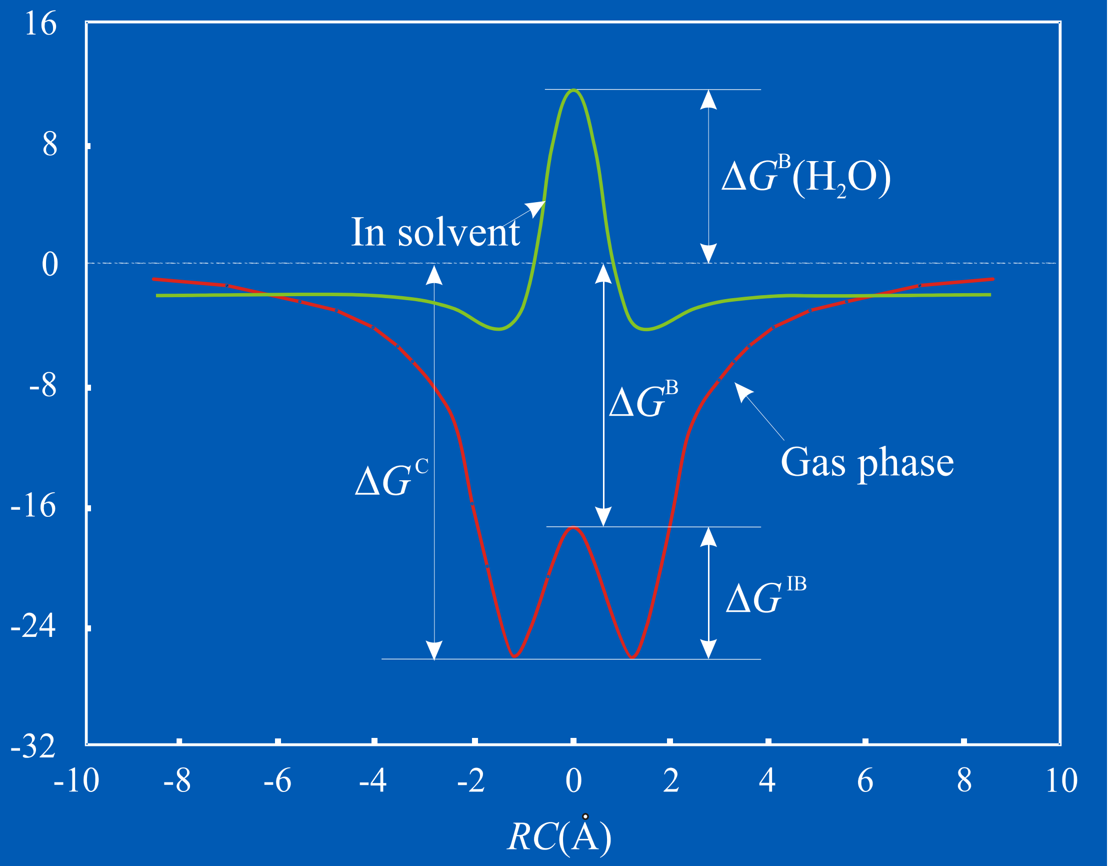
Scales as $N \log N$

(with 3D-FFT)

Free energy profile for the identity S_N2 reaction in gas phase and in solution by KS-DFT/3D-RISM-KH vs continuum solvation and experiment

D. Casanova, S. Gusarov, A. Kovalenko, T. Ziegler,
J. Chem. Theory Comput., **3**, 458 (2007)





$F^- + CH_3F \Rightarrow FCH_3 + F^-$

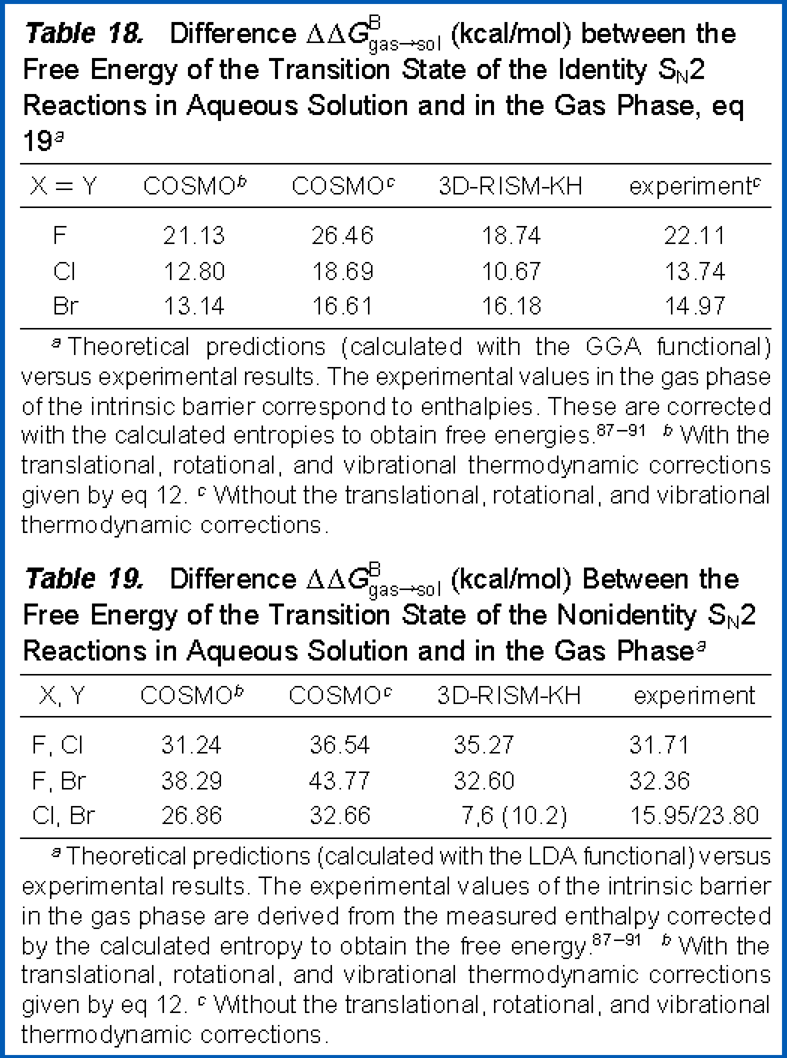


Table 18. Difference $\Delta\Delta G^{B}_{gas\rightarrow sol}$ (kcal/mol) between the Free Energy of the Transition State of the Identity S_N2 Reactions in Aqueous Solution and in the Gas Phase, eq 19[a]

X = Y	COSMO[b]	COSMO[c]	3D-RISM-KH	experiment[c]
F	21.13	26.46	18.74	22.11
Cl	12.80	18.69	10.67	13.74
Br	13.14	16.61	16.18	14.97

[a] Theoretical predictions (calculated with the GGA functional) versus experimental results. The experimental values in the gas phase of the intrinsic barrier correspond to enthalpies. These are corrected with the calculated entropies to obtain free energies.[87−91] [b] With the translational, rotational, and vibrational thermodynamic corrections given by eq 12. [c] Without the translational, rotational, and vibrational thermodynamic corrections.

Table 19. Difference $\Delta\Delta G^{B}_{gas\rightarrow sol}$ (kcal/mol) Between the Free Energy of the Transition State of the Nonidentity S_N2 Reactions in Aqueous Solution and in the Gas Phase[a]

X, Y	COSMO[b]	COSMO[c]	3D-RISM-KH	experiment
F, Cl	31.24	36.54	35.27	31.71
F, Br	38.29	43.77	32.60	32.36
Cl, Br	26.86	32.66	7,6 (10.2)	15.95/23.80

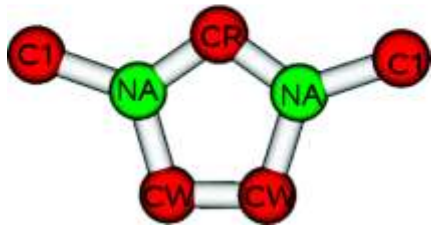
[a] Theoretical predictions (calculated with the LDA functional) versus experimental results. The experimental values of the intrinsic barrier in the gas phase are derived from the measured enthalpy corrected by the calculated entropy to obtain the free energy.[87−91] [b] With the translational, rotational, and vibrational thermodynamic corrections given by eq 12. [c] Without the translational, rotational, and vibrational thermodynamic corrections.

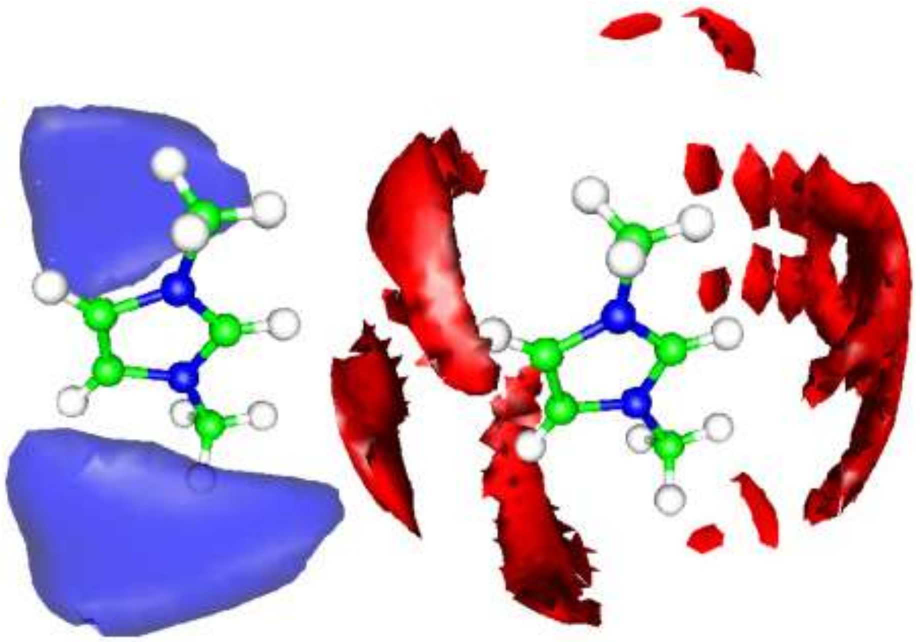
Ionic liquids: Multiscale KS-DFT / 3D-RISM-KH theory

M. Malvaldi, S. Bruzzone, C. Chiappe, S. Gusarov, and A. Kovalenko, J. Phys. Chem. B, **113**, 3536 (2009).

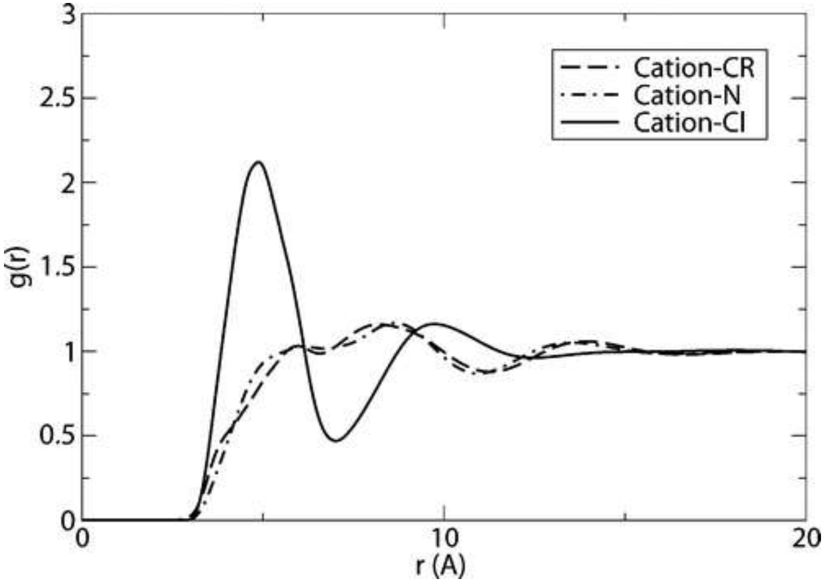
- **Electronic structure of IL in solution – at Car-Parrinello level (e.g. dipole moment enhancement)**
- **Solvation structure (e.g. stacking of neighboring cations)**
- **Thermochemistry**
- **Reactivity**

$mmim^+$ cation





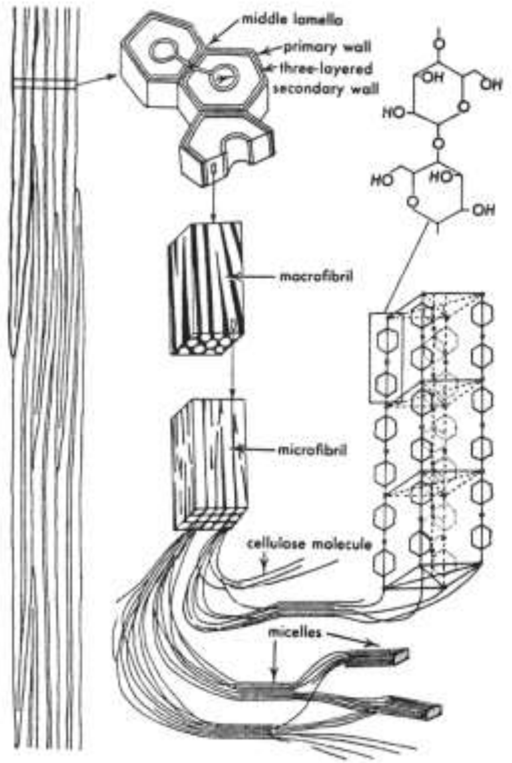
Radial distribution functions of $mmim^+Cl^-$



Solvation structure of methyl-methyl imidazolium ion in bulk liquid of [mmim][Cl] at T=400 K, obtained by self-consistent field KS-DFT / 3D-RISM-KH theory of electronic structure in solution. Isosurfaces of the nitrogen of $[mmim]^+$ cations at $g_N(r)>2$ (blue) and of Cl^- anions at $g_{Cl}(r)>5$ (red).

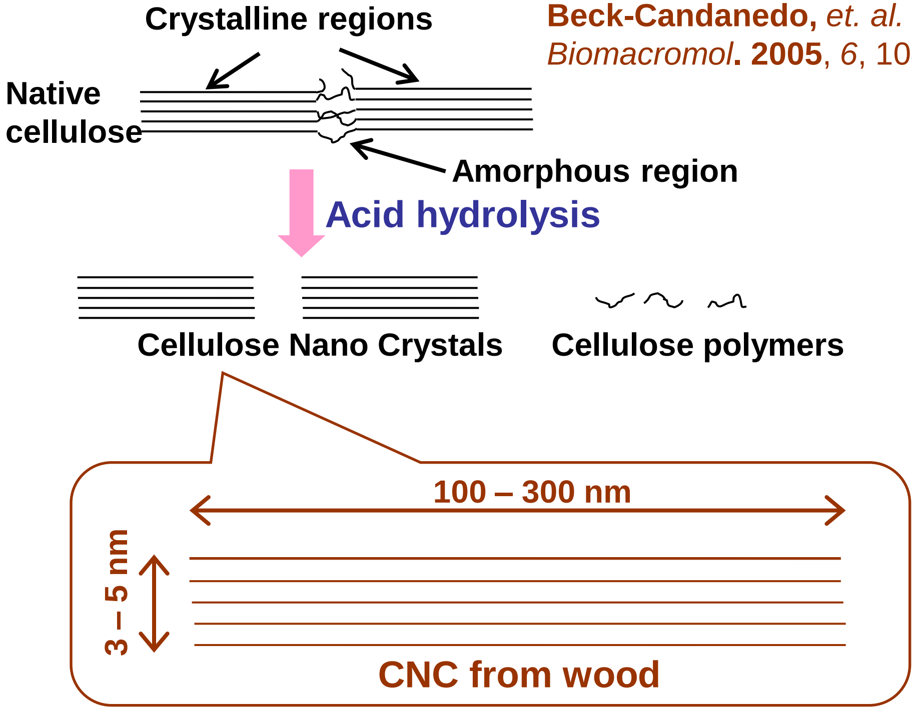
Cellulose Nano Crystals (CNC) in aqueous electrolyte solutions, organic solvents, and ionic liquids: Prediction of solubility and dispersion by 3D-RISM-KH theory

Cellulose Nano Crystals



Wadood, H. *Can. J. Chem. Eng.* **2006,** *84*, 513

Beck-Candanedo, *et. al.* *Biomacromol*. **2005**, *6*, 1048.



Crystalline regions

Native cellulose

Amorphous region

Acid hydrolysis

Cellulose Nano Crystals

Cellulose polymers

100 – 300 nm

3 – 5 nm

CNC from wood

Cellulose Nano Crystals (CNC):

- Reinforcing additive in high-value products
- Allows the deepest level of structure control
- Has reduced defect occurrence
- Easy to functionalize
- Need tuning of hydrophilicity/hydrophobicity

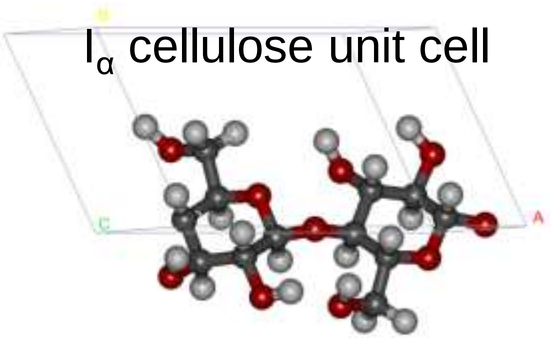
CNC modifications:

- Tune surface properties
- Retain hydrogen bonding network
- Preserve mechanical properties
- Control colloidal properties (dispersion)
- Control polymer melt / solution flow

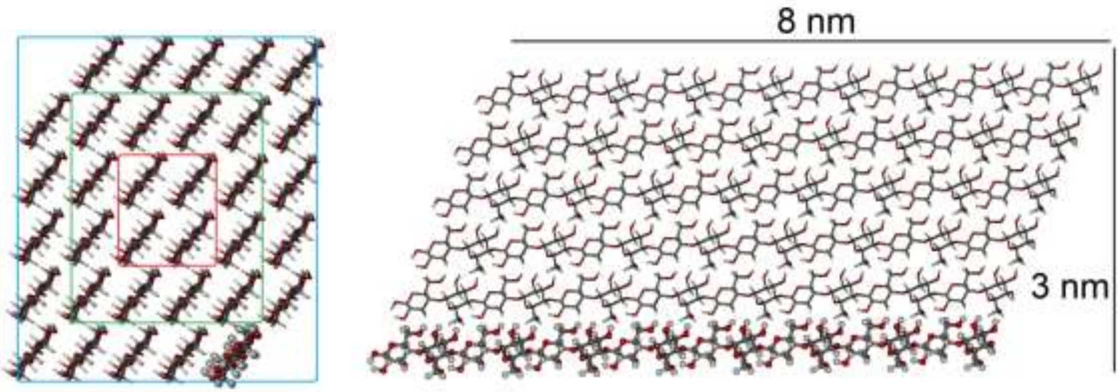
Multiscale theory, modeling, and simulations (DFT, MD, and sD-RISM-KH)
to develop quantitative structure-property relationships (QSPR)
between surface chemical modifications and properties
for rational design of CNC-based materials

CNC particles: Models and Methods

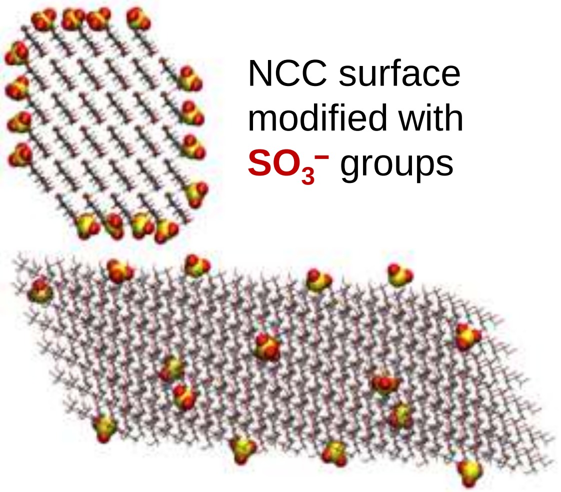
- Quantum Chemistry (DFT)
  - Geometrical and electronic structure
  - Missing force field parameters
- Molecular Dynamics (Amber11 software)
  - GLYCAM06 (carbohydrate force field)
  - SPC/E water model
- 3D-RISM-KH solvation structure and thermodynamics of nanoparticles in solution



I_α cellulose unit cell



I_α cellulose crystal structure has been used to build the initial fibril: 34 chains (every chain consists of 16 glucose units)



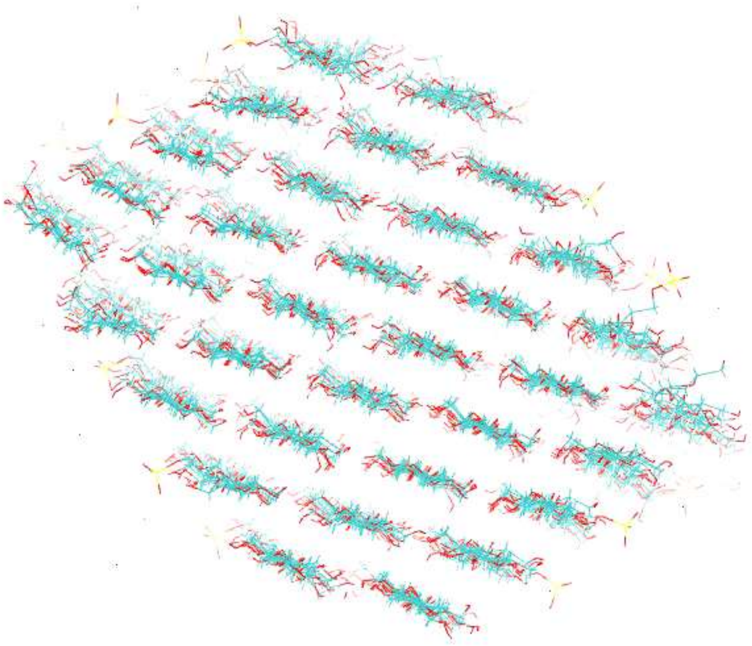
NCC surface modified with **SO_3^-** groups

Sulfate concentration:
Boluk, Y. *Elemental analysis*

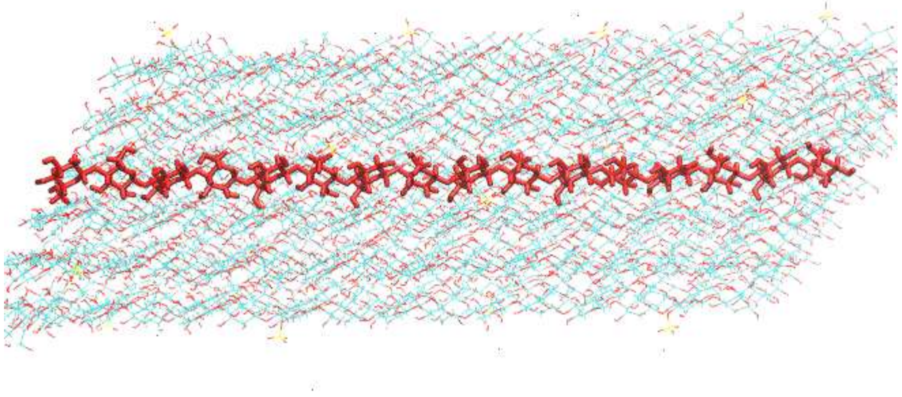
S. R. Stoyanov, S. Gusarov, and A. Kovalenko, in: *Production and Application of Cellulose Nanoparticles*, M. T. Postek, R. J. Moon, A. Rudie, M. Bilodeau (Eds.), (TAPPI Press, 2013), pp. 147-150.

S. R. Stoyanov, O. Lyubimova, S. Gusarov, and A. Kovalenko, *Nordic Pulp & Paper Res. J.*, **29**. 144 (2014).

Relaxation and twisting of CNC particle



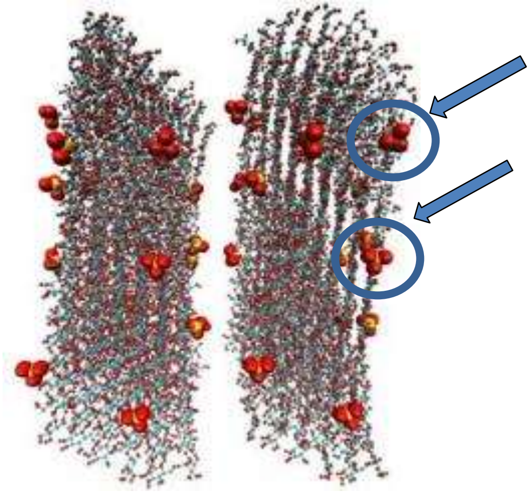
top view



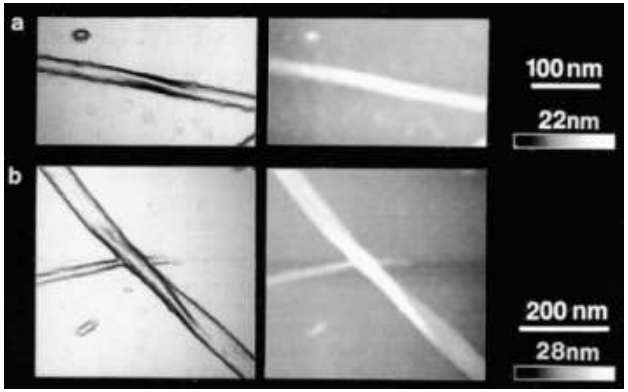
side view

S. R. Stoyanov, O. Lyubimova, S. Gusarov, and A. Kovalenko, *Nordic Pulp & Paper Res. J.*, **29**. 144 (2014)

CNC particles twisting: Analysis

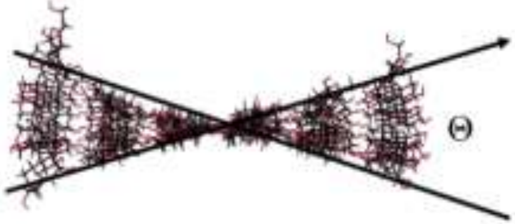


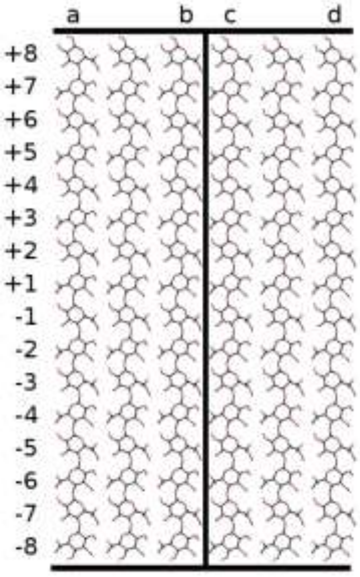
CNC surface modified with SO_3^- groups

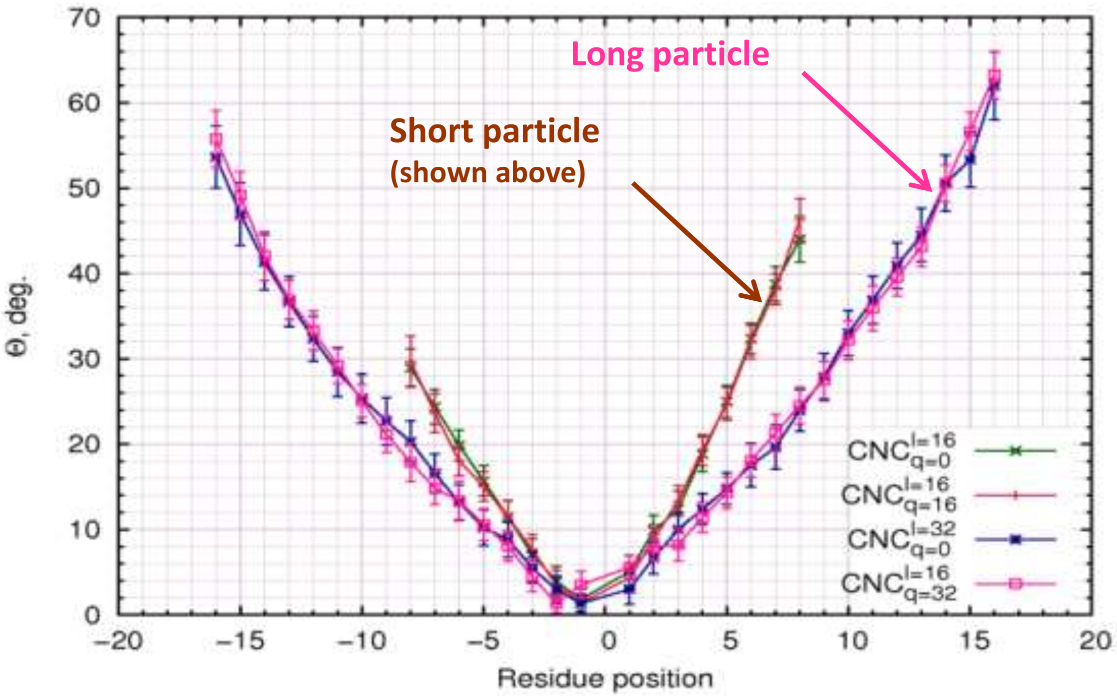


TM-AFM images of microfibrils of Micrasterias denticulata. In both (a) and (b) the microfibrils can be seen to undergo right-handed twist.

S.J. Hanley, J-F. Revol, L. Godbout, D. G. Gray. **Cellulose** (1997), 4, 209-220







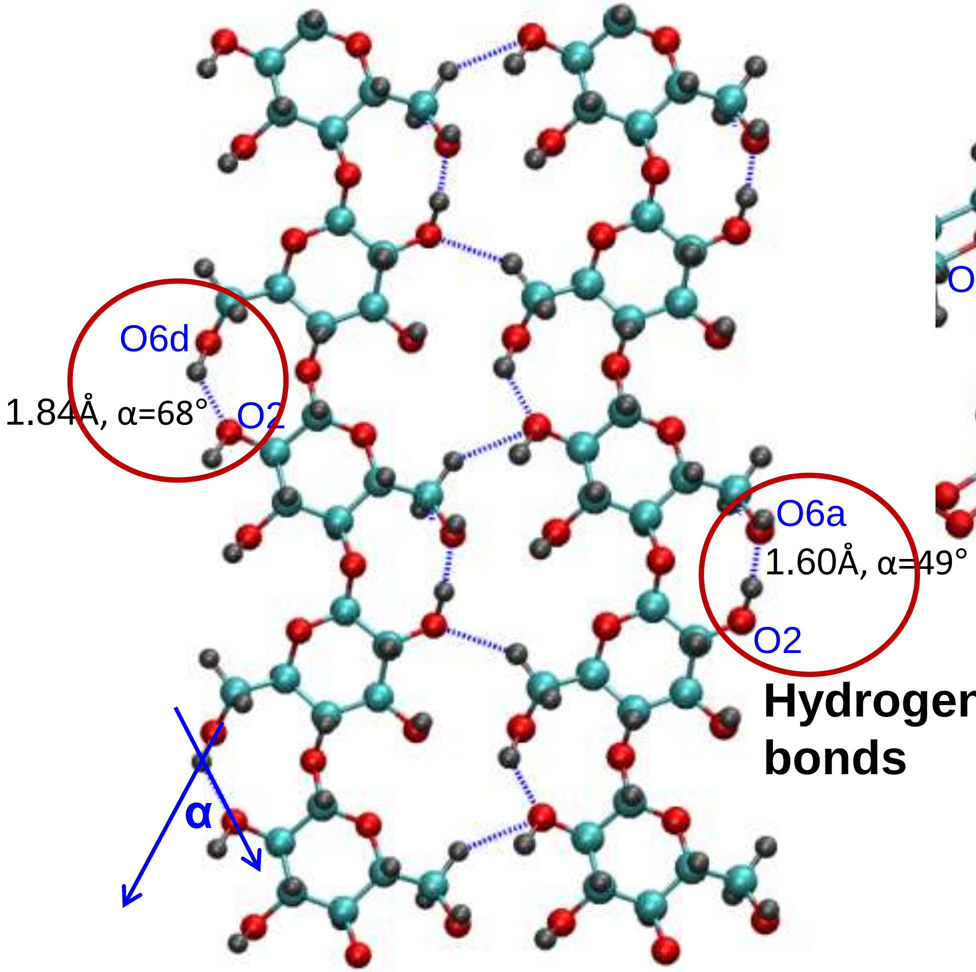
Long particle

Short particle
(shown above)

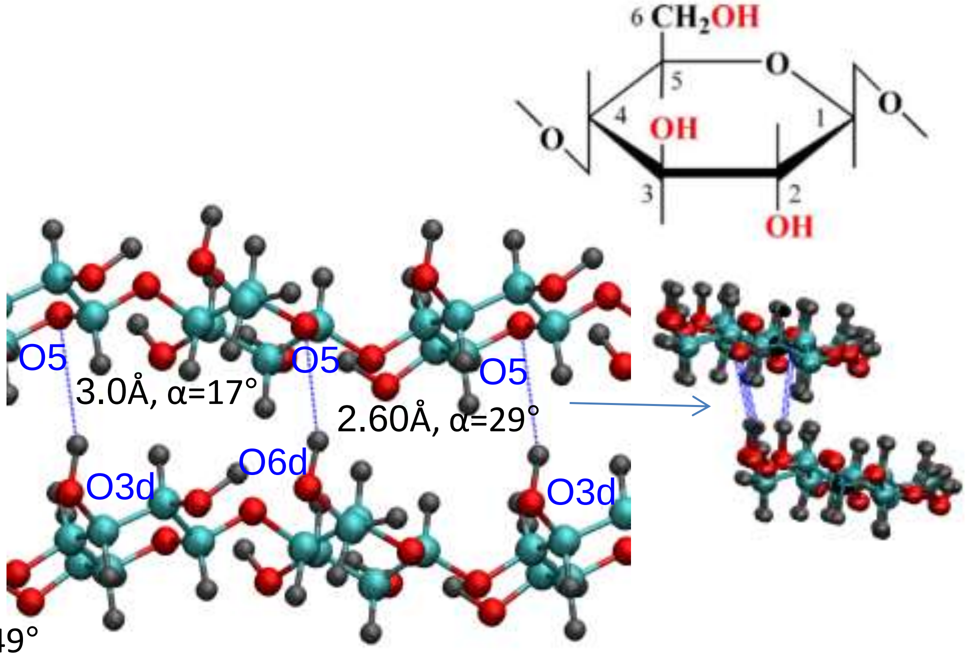
S. R. Stoyanov, S. Gusarov, and A. Kovalenko, in: *Production and Application of Cellulose Nanoparticles*, M. T. Postek, R. J. Moon, A. Rudie, M. Bilodeau (Eds.), (TAPPI Press, 2013), pp. 147-150.

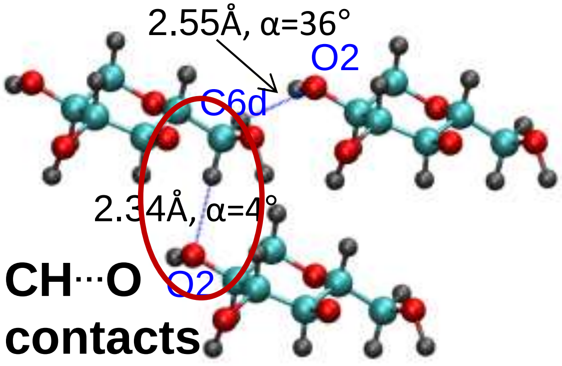
S. R. Stoyanov, O. Lyubimova, S. Gusarov, and A. Kovalenko, *Nordic Pulp & Paper Res. J.*, **29**. 144 (2014).

Experimental data: hydrogen bonding in Iα cellulose



Hydrogen bonds



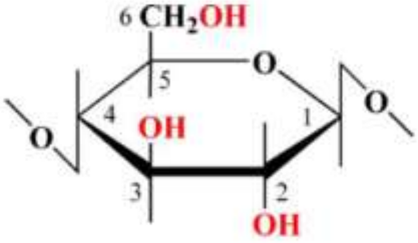


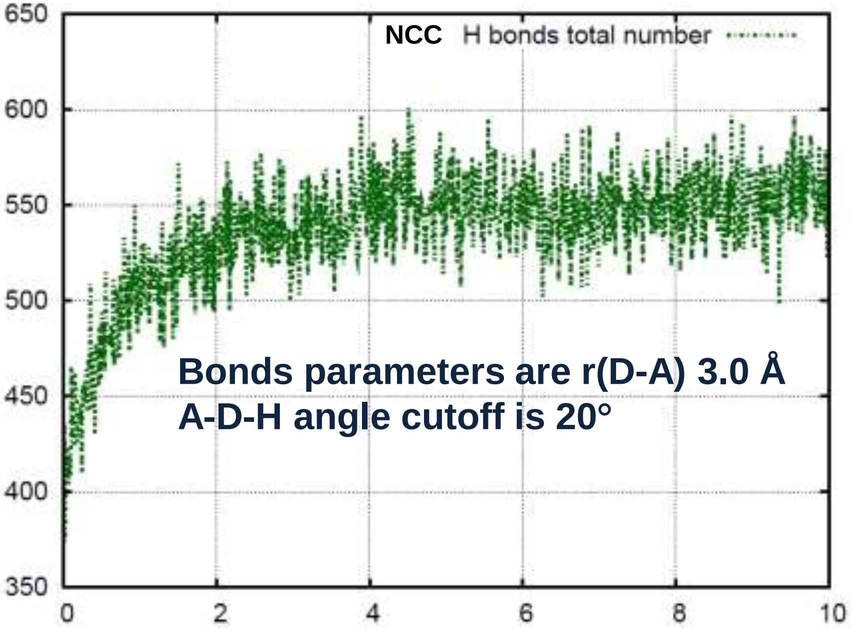
CH···O contacts

Hydrogen bonding parameters:

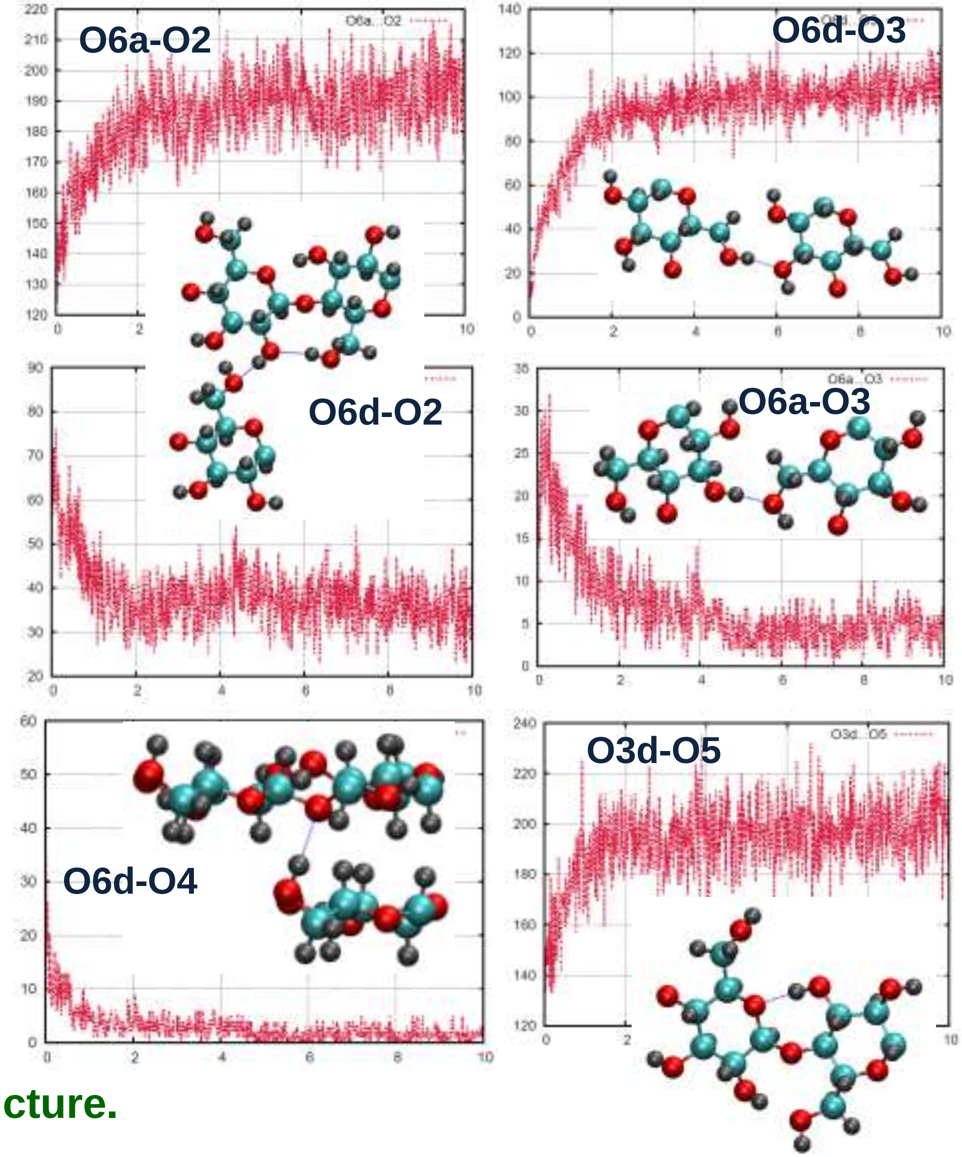
- r(D-A) 3.0 Å
- A-D-H angle cutoff of 20°

Reorganization of CNC intramolecular hydrogen bonds





Bonds parameters are r(D-A) 3.0 Å
A-D-H angle cutoff is 20°



- ➢ **Hydrogen bonds stabilize twisted CNC structure.**

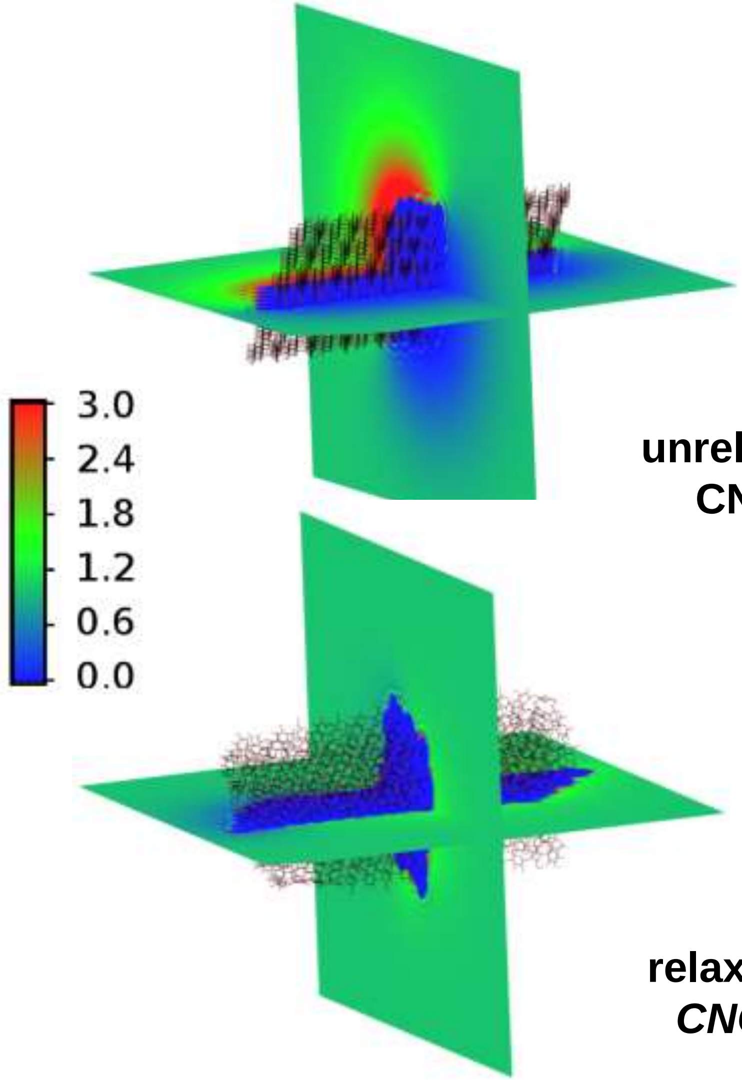
S. R. Stoyanov, O. Lyubimova, S. Gusarov, and A. Kovalenko, *Nordic Pulp & Paper Res. J.*, **29**. 144 (2014)

Charge redistribution upon relaxation of neutral CNC

Sections of 3D maps of Na^+ distributions $g_{Na+}(r)$ in solution around CNC particle

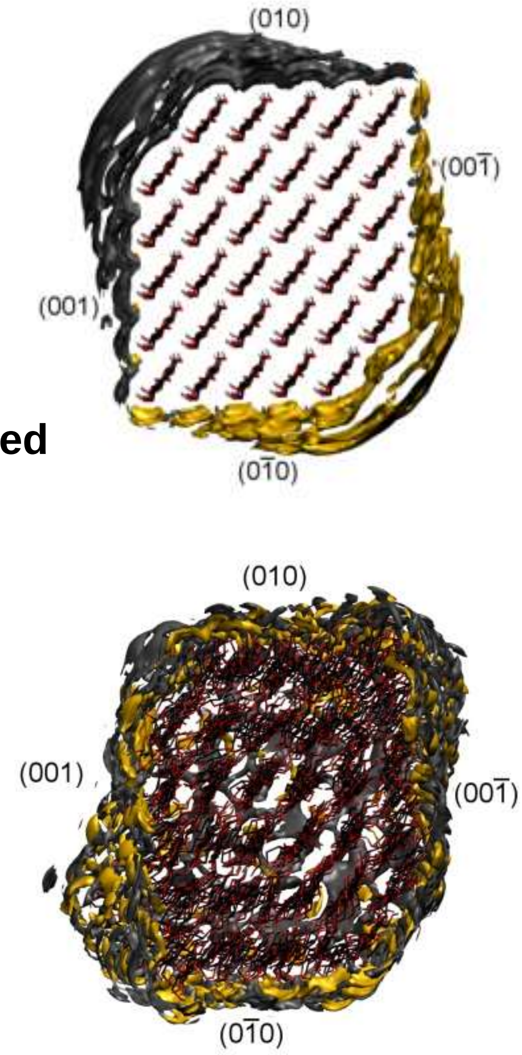
$g_{Na+}(r)$=3.0 (in yellow)
$g_{Cl-}(r)$=3.0 (in gray)

0.030 mol/kg NaCl



unrelaxed CNC

relaxed *CNC*



Before relaxation:

- **Thick interfacial layer of Na^+ counterions**

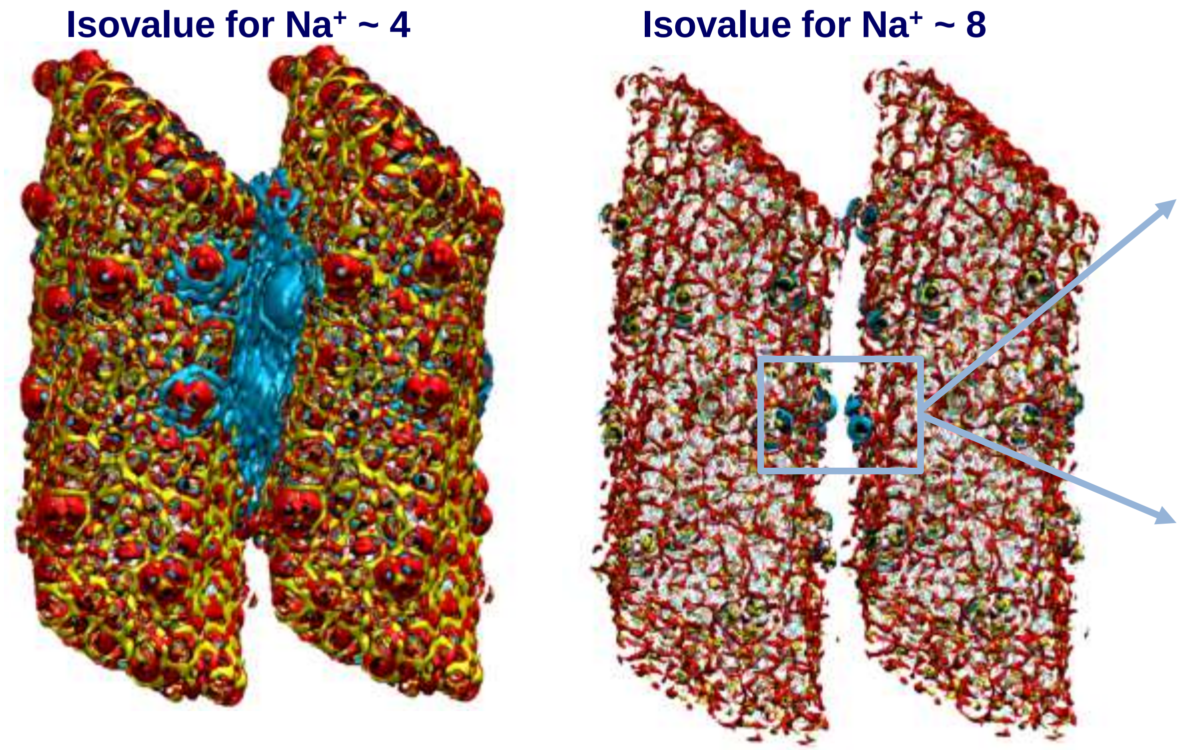
After relaxation:

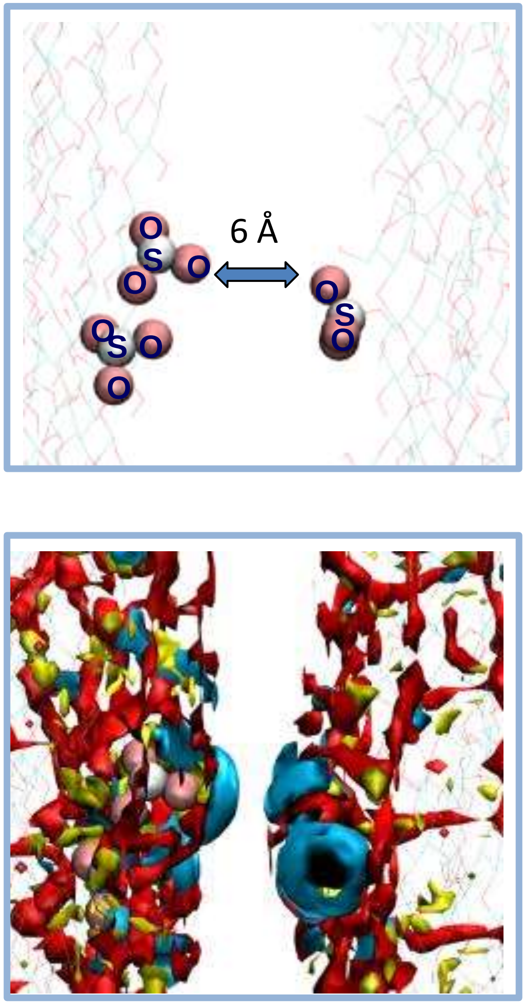
- **Thin interfacial layer of Na^+ counterions**
- **Even distribution of Na^+ and $Cl^–$ ions**

S. R. Stoyanov, O. Lyubimova, S. Gusarov, and A. Kovalenko, *Nordic Pulp & Paper Res. J.*, **29**. 144 (2014)

Effective Interaction of CNC rods bridged by Na$^+$ counterions

- 3D distribution functions **at isoelectric point** show formation of Na$^+$ bridges stabilizing NCC aggregates





🟦 Na$^+$
🟨 H$_w$
🟥 O$_w$

Towards understanding and control of:

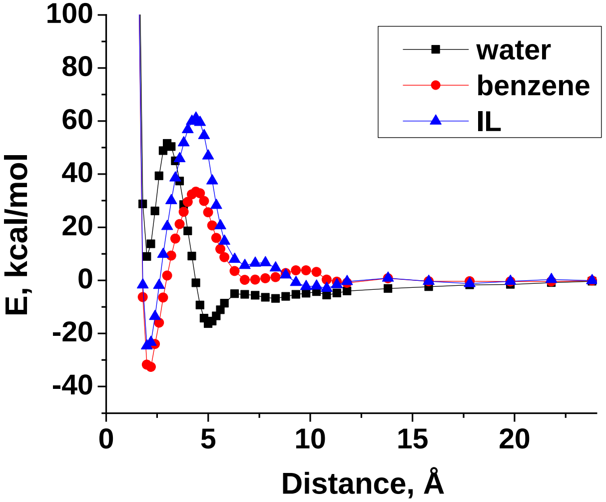
- ✓ **cholesteric ordering**
- ✓ **NCC dispersion in various solvents**

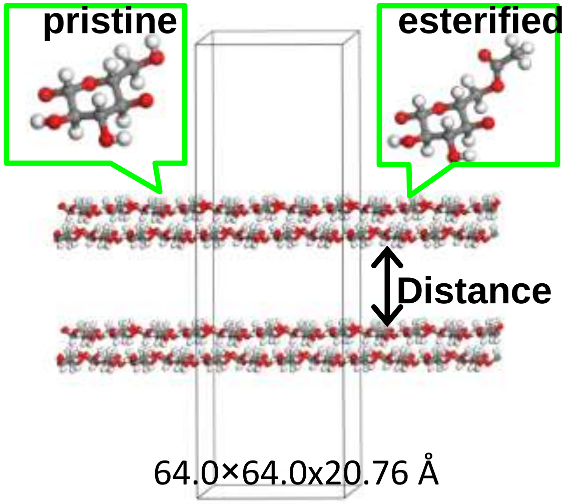
S. R. Stoyanov, S. Gusarov, and A. Kovalenko, in: *Production and Application of Cellulose Nanoparticles*, M. T. Postek, R. J. Moon, A. Rudie, M. Bilodeau (Eds.), (TAPPI Press, 2013), pp. 147-150.

S. R. Stoyanov, O. Lyubimova, S. Gusarov, and A. Kovalenko, *Nordic Pulp & Paper Res. J.*, **29**. 144 (2014).

Dispersion of modified CNC in water, benzene, and ionic liquid

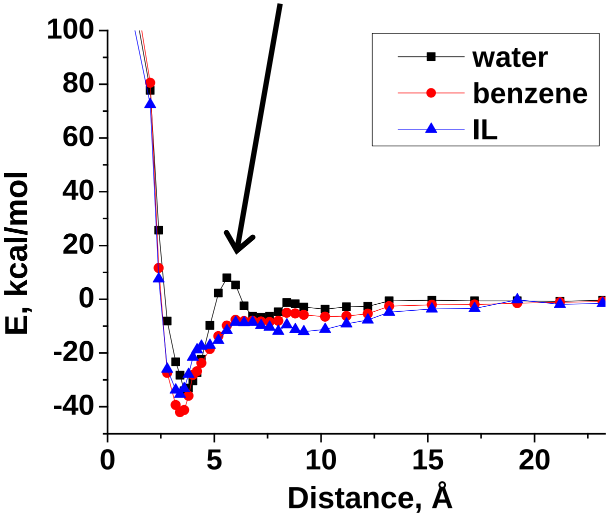
Disaggregation barriers in benzene and IL are higher than in water for pristine CNC

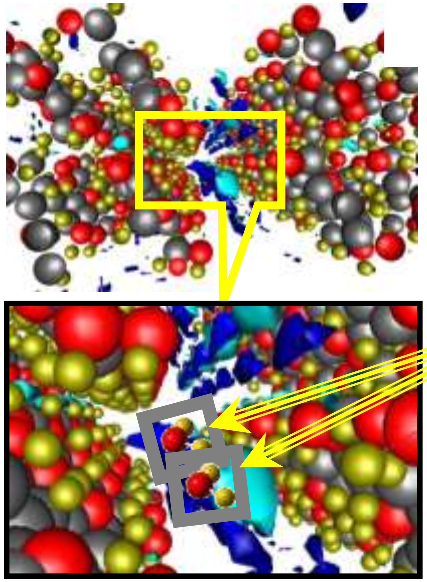




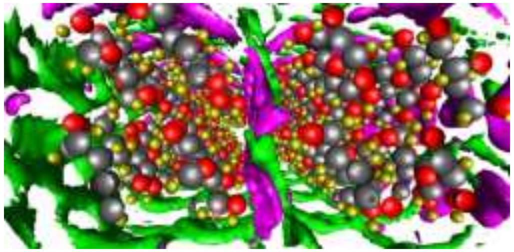
Structures are relaxed using PCFF, with CNC backbone atoms fixed, in Accelrys Materials Studio®

No solvent expulsion barrier and a lower disaggregation barrier in benzene or in ionic liquid for esterified CNC





$g_{Na} = 2.0, g_{Cl} = 2.4$

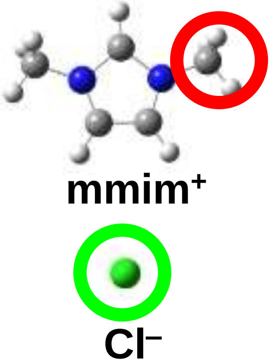


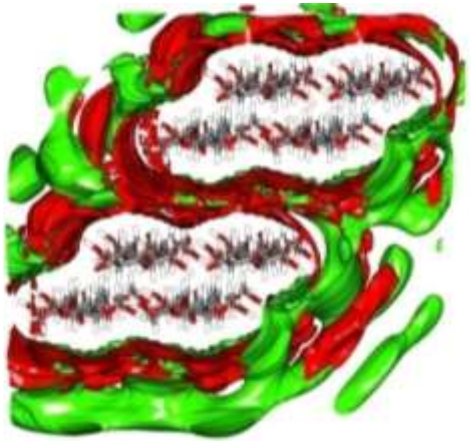
CNC-water hydrogen bonds replace CNC-CNC hydrogen bonding network

● H ● C ● O

🟦 H 🟦 O 🟩 Cl^- 🟪 Na^+

$g_O = 5.0$
$g_H = 2.8$
$g_{Na} = 2.0$
$g_{Cl} = 2.4$





Stoyanov, S. R.; Gusarov, S., Kovalenko, A.
in: *Production and Application of Cellulose Nanoparticles*, Postek, M. T.; Moon, R. J.; Rudie, A.; Bilodeau, M., Eds.
TAPPI Press, 2013, pp. 147-150.

Engineering modified cellulose nanocrystals (CNCs) as a compound carrier with programmable release

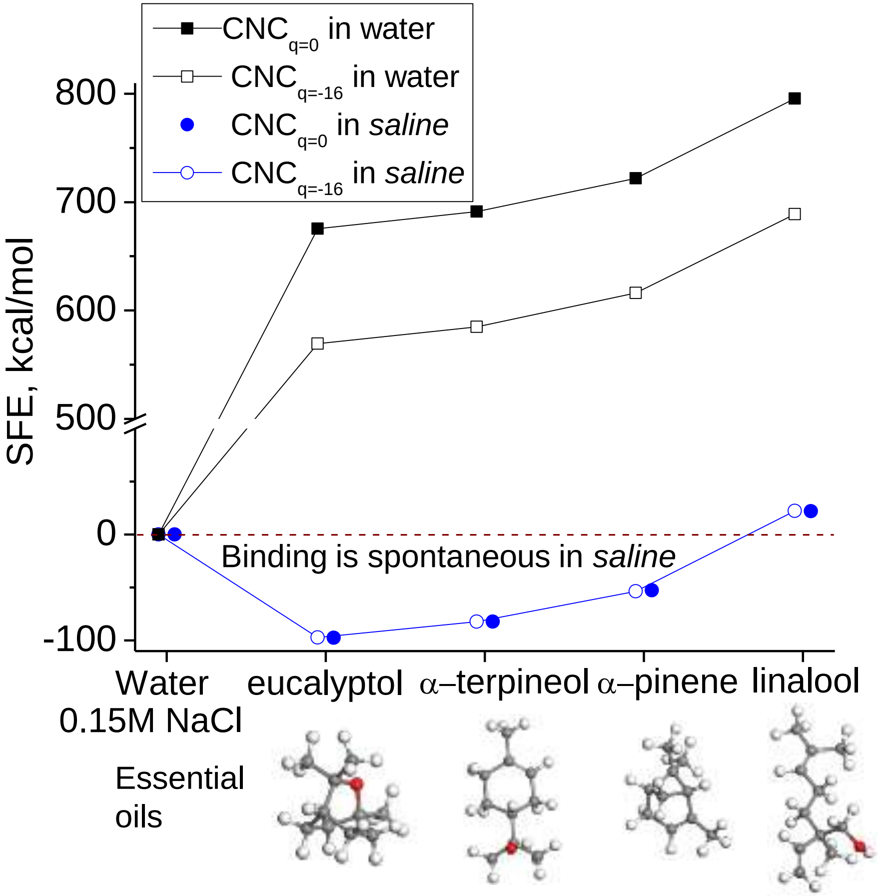
Cellulose nanocrystals (CNC) – biocompatible and biodegradable natural nanoparticles – are attractive for advanced applications in composites, foams and gels as well as in drug delivery as compound carriers.

Essential oils (EO) are natural products with antiradical, antioxidant, anti-inflammatory and antimicrobial properties, and are effective for inhibiting different human diseases. A recent structure-activity relationship study shows the following:

- The antimicrobial and antioxidant activities of terpene EOs are highly affected by their chemical nature and interactions;
- The biological activity of terpene EO components is related to synergistic and additive effects on the common prominent pathogenic and spoilage food-related bacteria;
- Characterization of the mechanisms of changes in antibacterial activities in systems of mixtures of terpenes is needed to design blends for the best medical and food safety applications.

We intend to explore the mechanism of adsorption and desorption of EOs on CNC in solution so as help engineer CNC surface modifications for effective binding and programmable release of EOs and other compounds.

Here, we employ the 3D-RISM-KH molecular theory of solvation to predict the thermodynamics of adsorption of essential oils on CNC in water and saline solution.



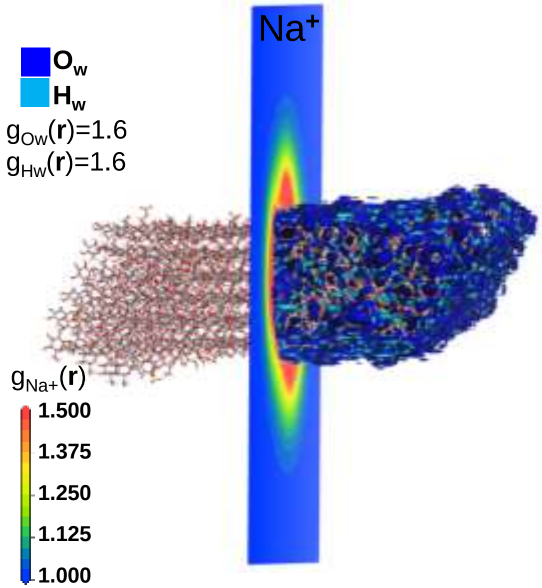
Zengin, H.; Baysal, A. H. *Molecules* **2014**, *19*, 17773-17798.

Lyubimova, O.; Stoyanov, S.R.; Gusarov, S.; Kovalenko, A. *Langmuir* **2015**, *31*, 7106-7116.

Stoyanov, S. R.; Lyubimova, O.; Gusarov, S.; Kovalenko, A. *Nordic Pulp Paper Res. J* **2014**, *29*, 144–155.

Stoyanov, S. R.; Gusarov, S.; Kovalenko, A. Multiscale Modeling of Solvation and Effective Interactions of Functionalized Cellulose Nanocrystals. In *Production and Applications of Cellulose Nanomaterials.* Postek, M. T.; Moon, R. J.; Rudie, A.; Bilodeau, M., Eds.; TAPPI Press, Atlanta, GA, USA, **2013**, pp. 147–150.

Preferred Binding Configurations of CNCs in *saline*



In *saline*, Na^+ and Cl^- form a spatially-extended electric double layer that stabilizes the negatively-charged CNC grafted with surface sulfate ester groups.

The models of pristine and sulfate-ester grafted CNC particles are developed and studied in aqueous NaCl solutions [Lyubimova *et al.*]. The EOs are modeled using UAM [Kobryn *et al.*]

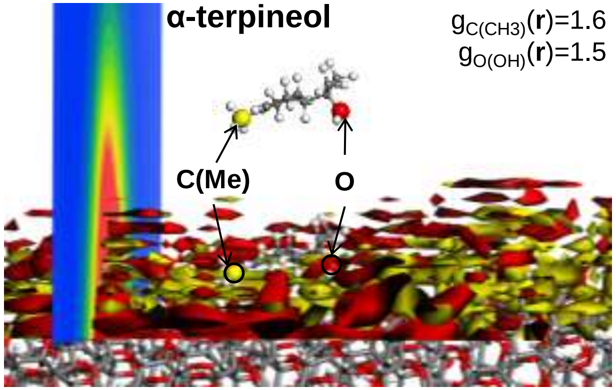
Kovalenko, A. Three-dimensional RISM theory for molecular liquids and solid-liquid interfaces, In *Understanding Chemical Reactivity: Molecular Theory of Solvation*, F. Hirata, ed., Kluwer Academic Publishers, 24 (2003), pp.169-275.

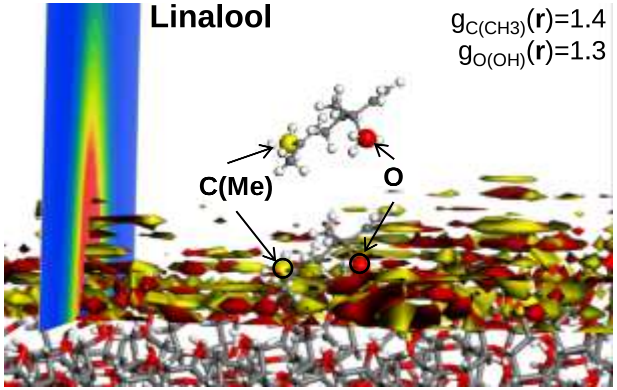
Kobryn, A. E.; Kovalenko, A. *J. Chem. Phys.* **2008**, *129*, 134701.

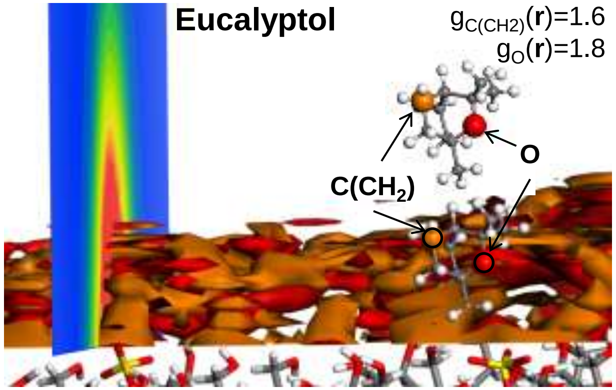
In saline, the binding of EOs is spontaneous on both neutral and negatively-charged CNC at room temperature. The binding preferences show the following:

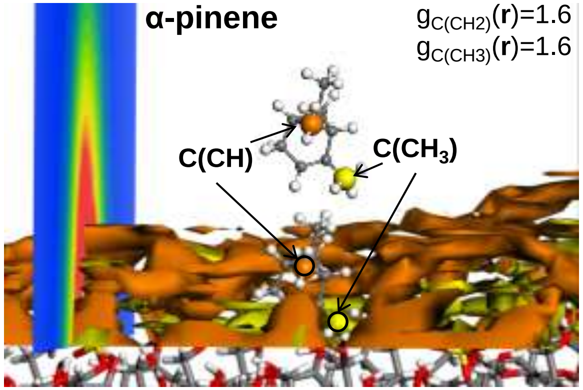
- Functional groups containing heteroatoms bind preferentially to hydrogen-bonding sites and tend to be closer to the CNC surface, as evinced by localized isosurfaces (red);
- Methyl groups interact mainly by van der Walls forces that are less specific, as evidenced by the more diffuse isosurfaces (yellow);
- CH_2 group interactions are even less specific, as shown by the diffuse isosurfaces (orange);
- The preferred binding configurations are with the heteroatom moiety towards the CNC hydrophilic CNC surface and the hydrocarbon groups away from the surface.

Such "dressed" CNC would be more hydrophobic and could become lipid-soluble, suggesting enhanced intake. Release studies in lipid media could be the next steps in this investigation.


α-terpineol


Linalool

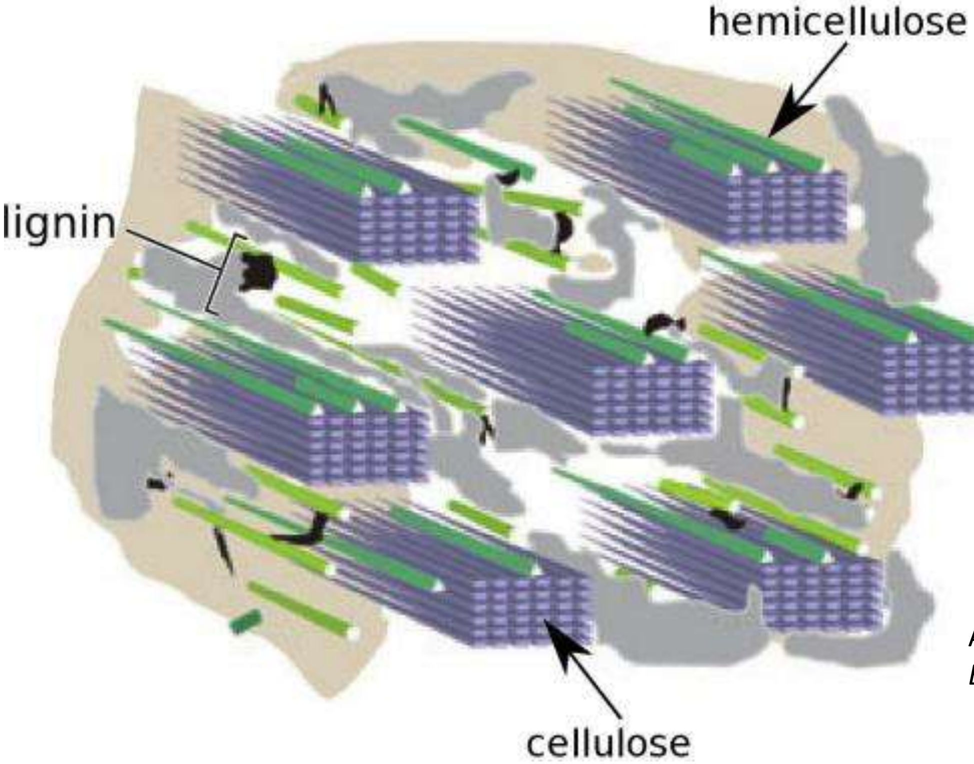

Eucalyptol


α-pinene

Lyubimova, O.; Stoyanov, S.R.; Gusarov, S.; Kovalenko, A. *Langmuir* **2015**, *31*, 7106-7116.

Overcoming plant biomass recalcitrance

The Plant Cell Wall



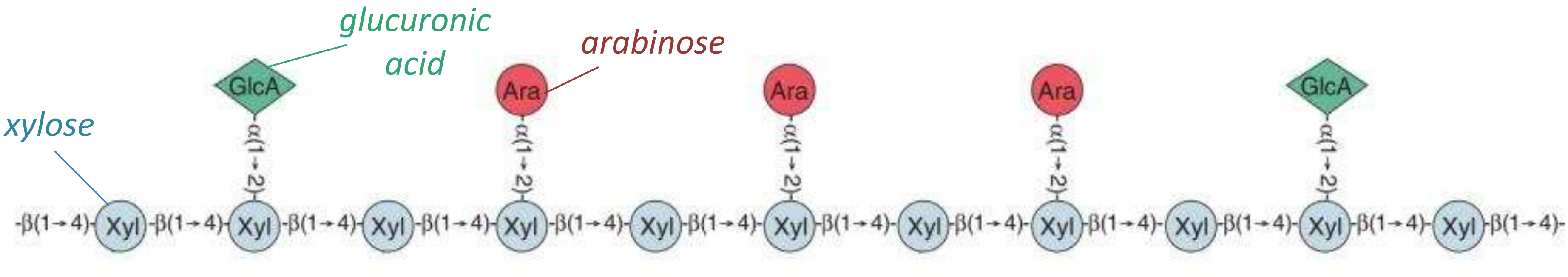
Adapted from: *Annu. Rev. Chem. Biomol. Eng.* (2011) **2**, 121-45

- **Cellulose is embedded in a matrix of hemicellulose and lignin**
- **The non-cellulosic compounds are responsible for providing the cohesive forces that stabilize the cell wall structure**

A molecular model of hemicellulose

Hemicellulose is a highly branched heteropolysaccharide

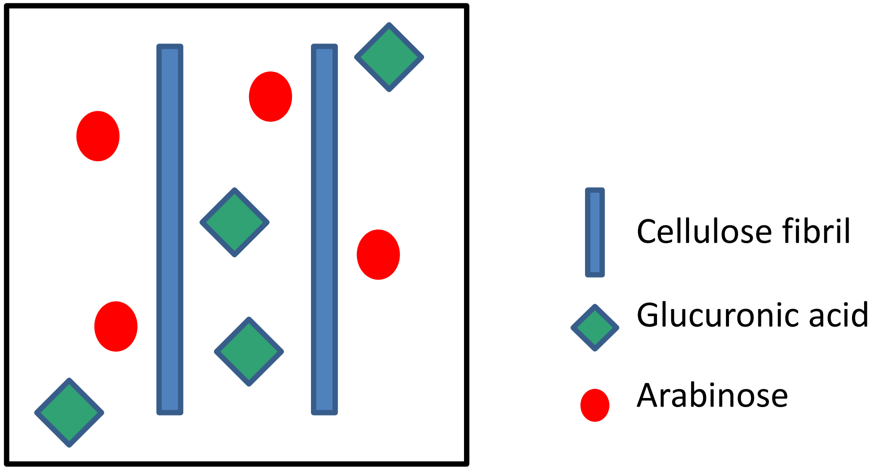
For example, Xylan occurs in sugar cane, corn and other crops



All of the chemical changes in the xylan structure regard the branches, mainly arabinose and glucuronic acid.

Our cell wall model

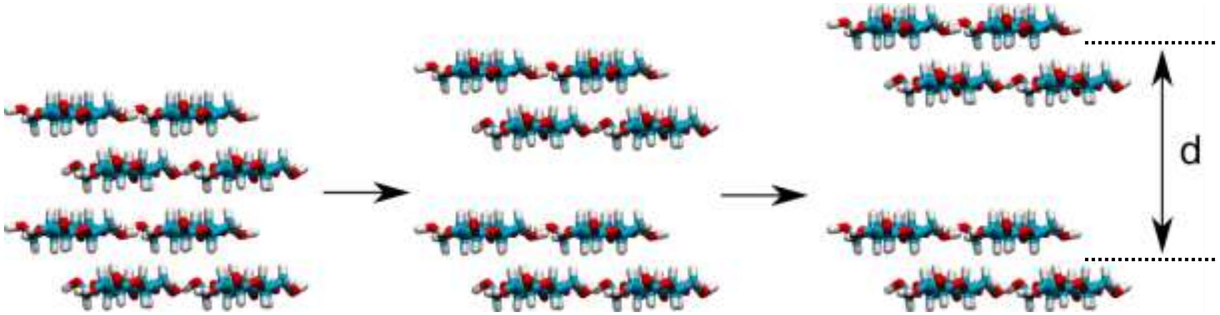
Two 4-chain cellulose fibrils immersed in a hydrogel of hemicellulose branches, e.g., arabinose and glucuronic acid monomers, which are considered as the solvent for the 3D-RISM-KH calculations



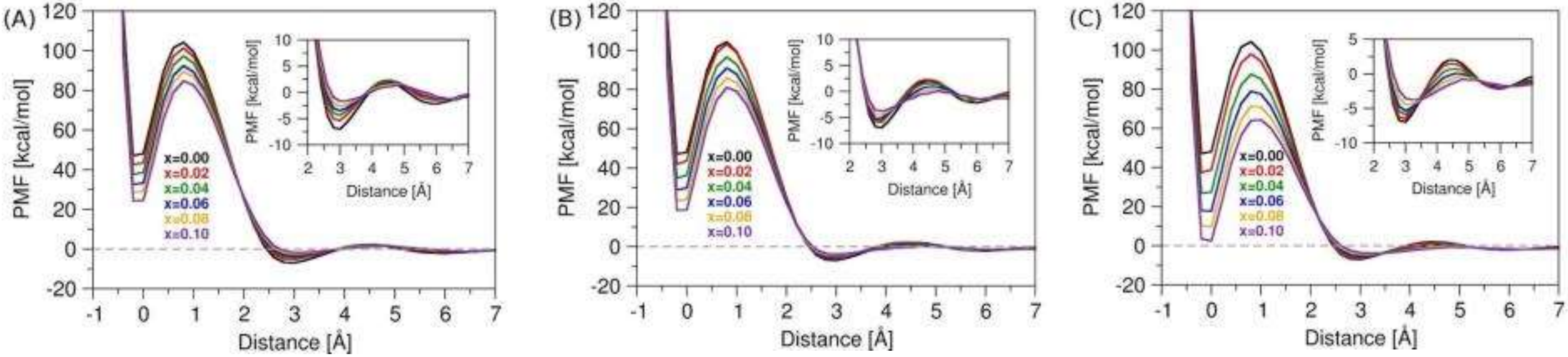
Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Am. Chem. Soc.* 203, *135*, 19048

Effective interactions (PMF): hydrophobic faces

Separation of the hydrophobic faces of two cellulose fibrils



Different compositions and concentrations of the hemicellulose branches


(A) arabinose

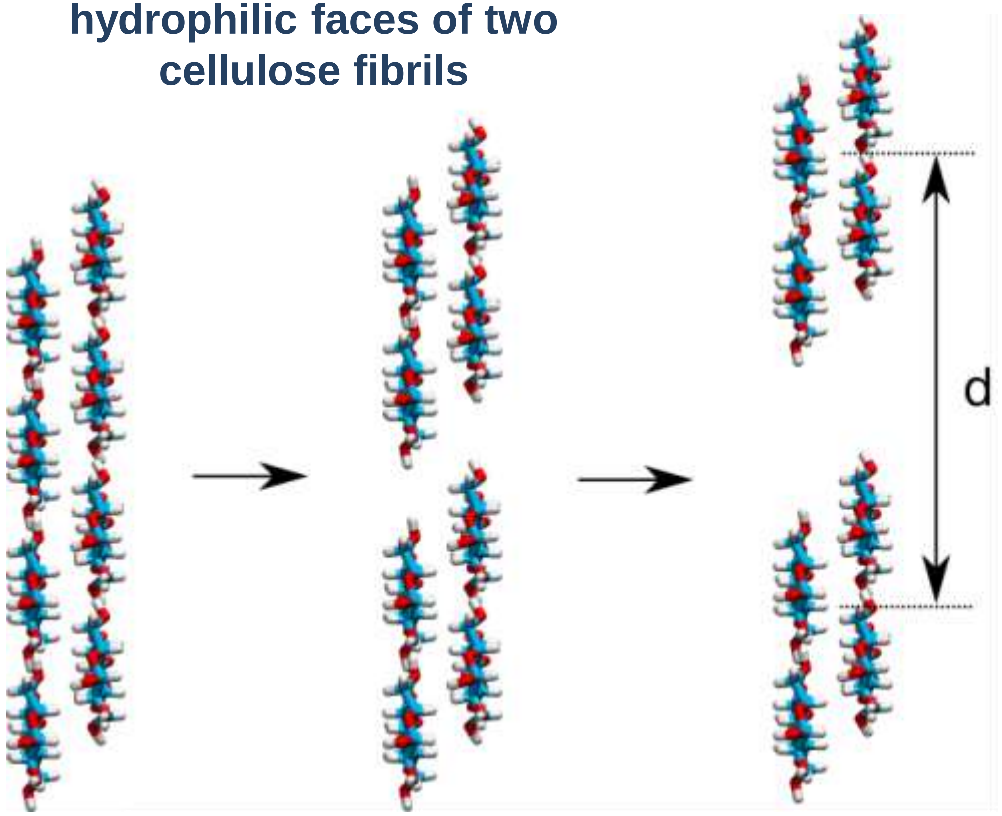
(B) glucuronic acid

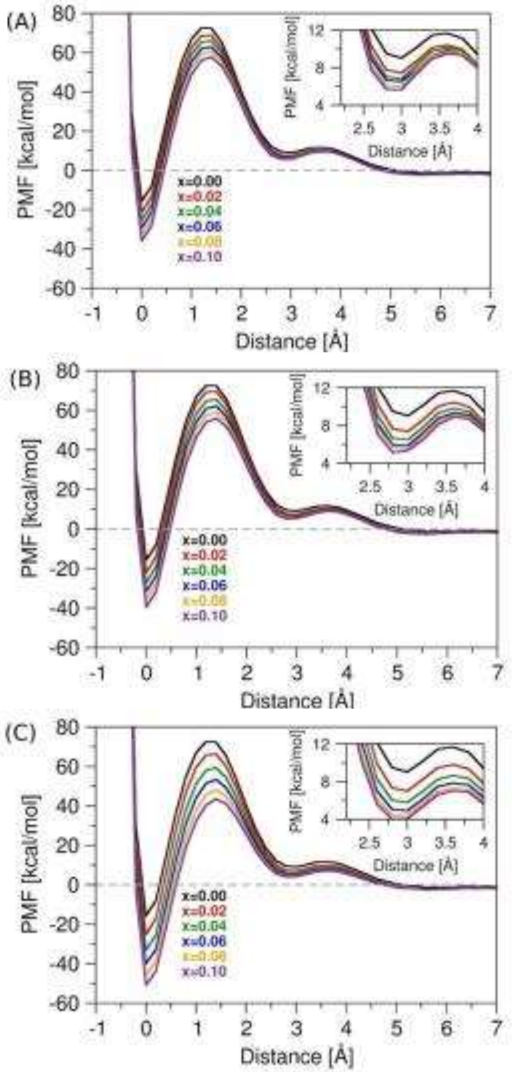
(C) glucuronate

Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Am. Chem. Soc.* 203, *135*, 19048

Effective interactions (PMF): hydrophilic faces

Separation of the hydrophilic faces of two cellulose fibrils





(A) arabinose

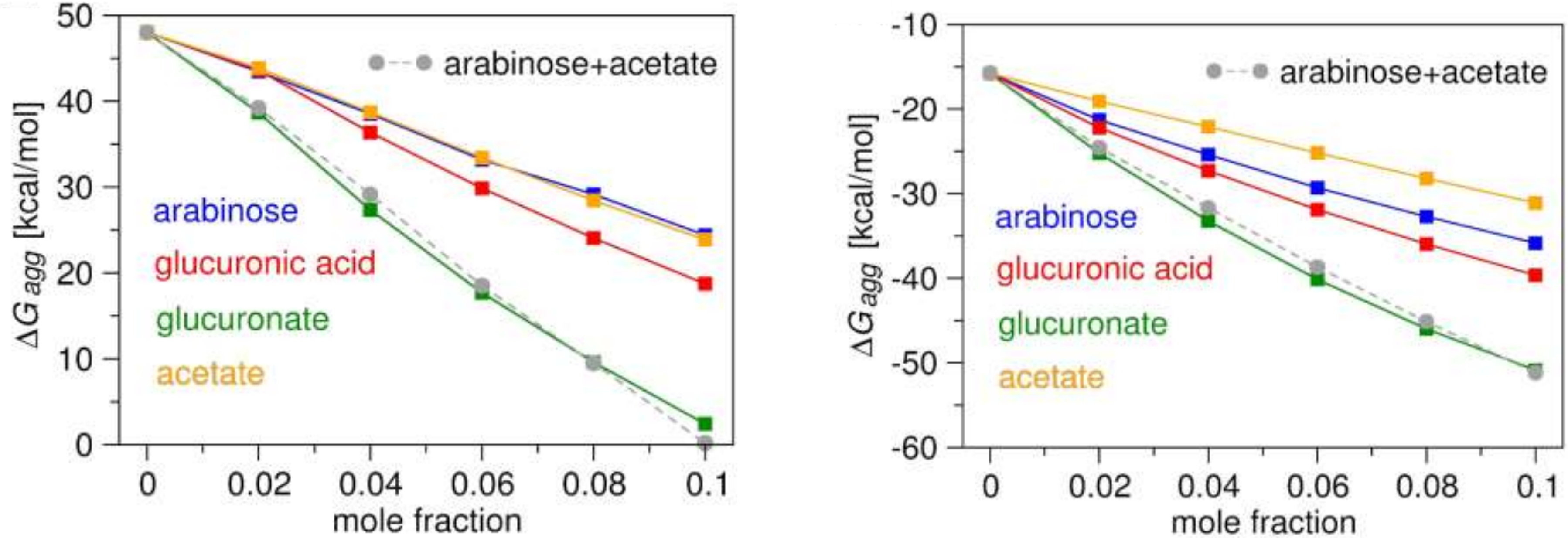
(B) glucuronic acid

(C) glucuronate

Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Am. Chem. Soc.* 203, *135*, 19048

Changes in hemicellulose composition affect cell wall stability

Aggregation Free Energy



Hydrophobic faces

Hydrophilic faces

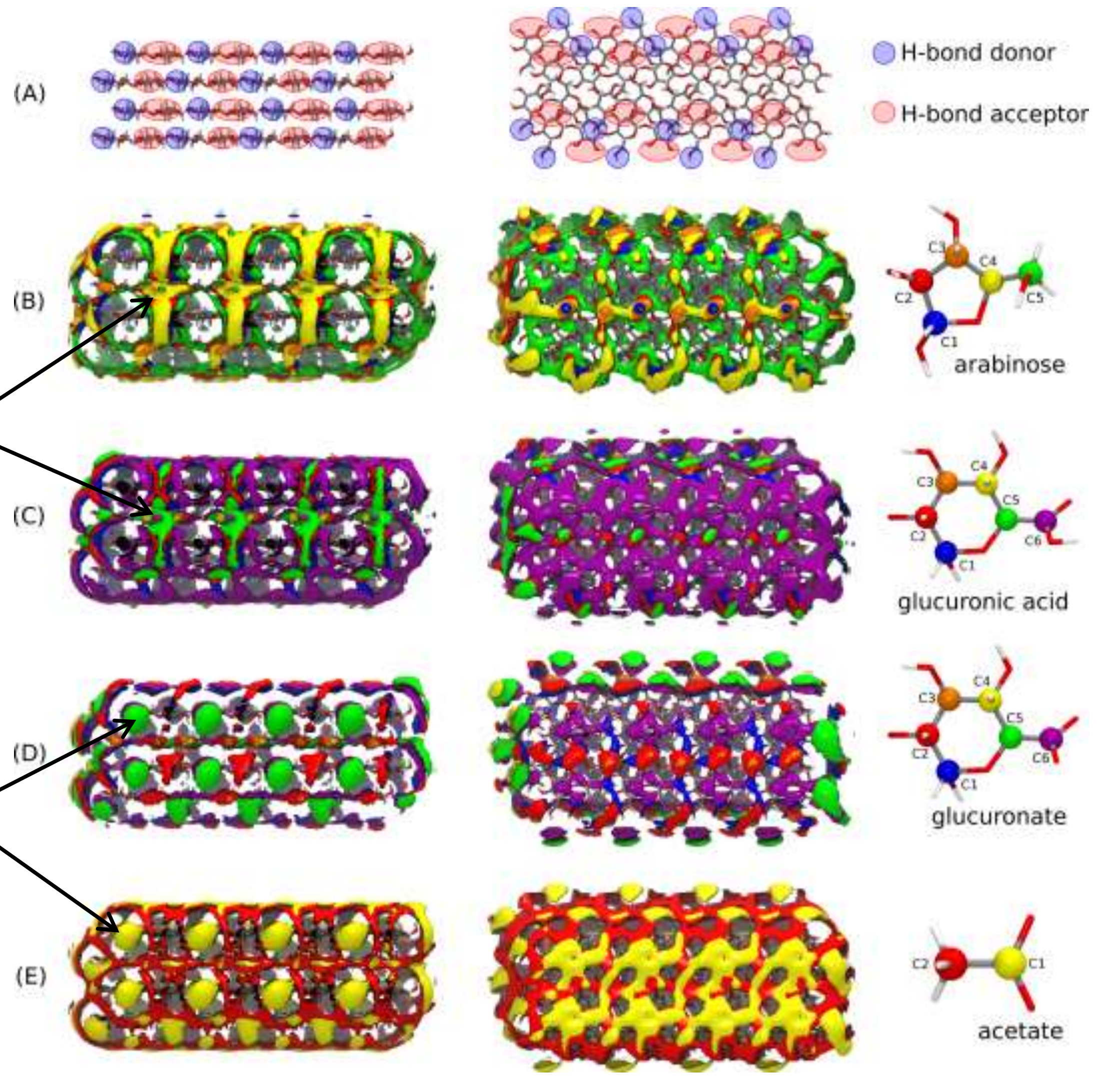
The hemicellulose composition strongly modulates the cell wall recalcitrance

Glucuronic acid & glucuronate are mainly responsible for the recalcitrance

Hydrophilic and basic interactions have comparable contributions to recalcitrance

Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Am. Chem. Soc*. 203, *135*, 19048

Site-specific cellulose-hemicellulose interactions



Secondary alcohol

Primary alcohol

Arabinose and glucuronic acid bind the H-bond acceptors on the cellulose surface (secondary alcohols)

Glucuronate and acetate bind the H-bond donors on the cellulose surface (primary alcohols)

Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Am. Chem. Soc*. 203, *135*, 19048

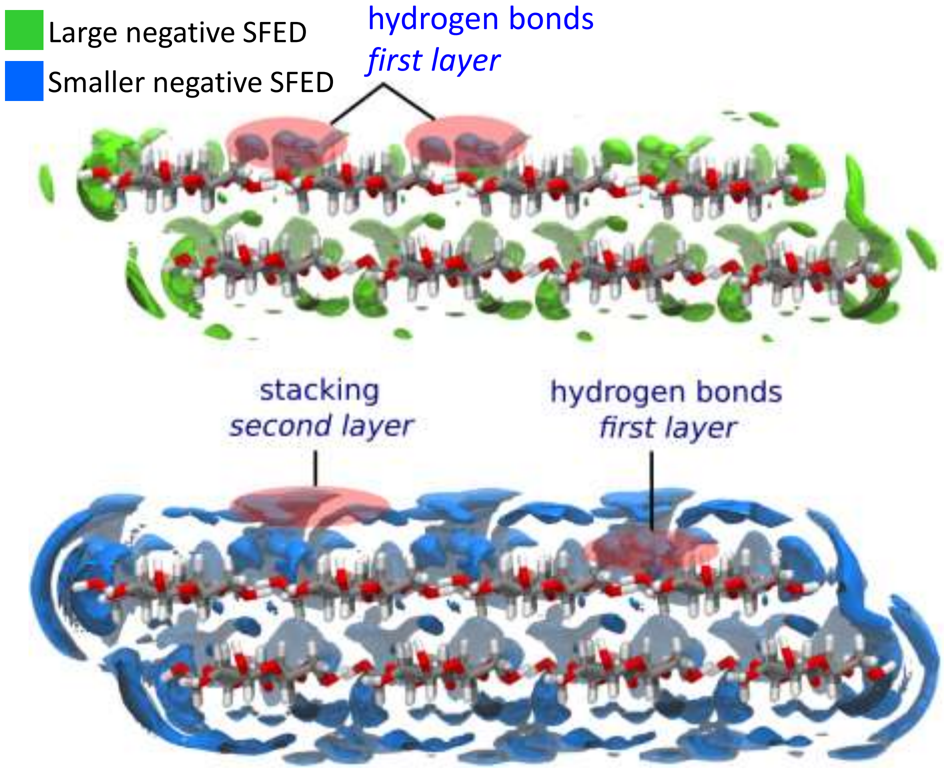
Cellulose-hemicellulose interactions: 3D solvation free energy density

$$\Delta\mu = \int \mathrm{d}\mathbf{r} \sum_{\gamma} \Phi_{\gamma}(\mathbf{r})$$

$$\Phi_{\gamma}(\mathbf{r}) = \rho_{\gamma} k_B T \left[\frac{1}{2} h_{\gamma}^{2}(\mathbf{r}) - \frac{1}{2} h_{\gamma}(\mathbf{r}) c_{\gamma}(\mathbf{r}) - c_{\gamma}(\mathbf{r}) \right]$$

Intensity of solute-solvent effective adhesion over 3D spatial regions of the solvation shells

$\Phi_{\gamma}(r) < 0$ - adhesion
$\Phi_{\gamma}(r) > 0$ - expulsion

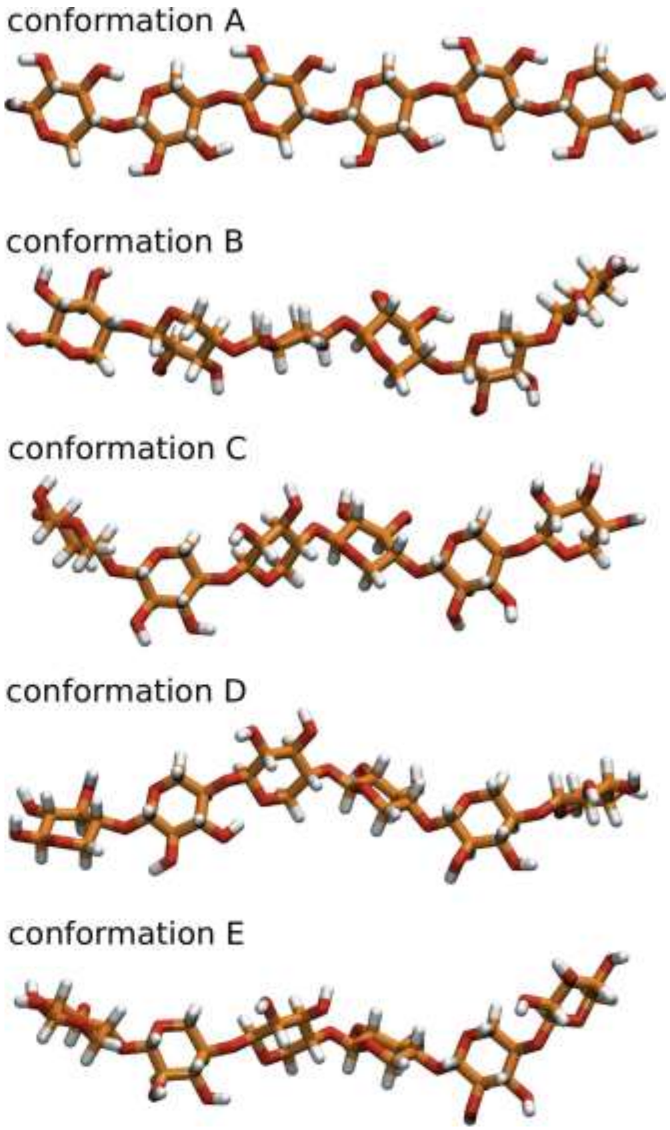


3D-SFED isosurfaces:

- **are highly localized around polar sites on the cellulose microfiber surface**
- **indicate hemicellulose-cellulose H-bonding**
- **show diffuse second layer of hemicellulose monomers stacking over cellulose surface**
- **show weaker stacking interactions in the second layer, less specific than H-bonding**
- **show stacking due to hydrophobic and enhanced intermolecular C—H·····O interactions**

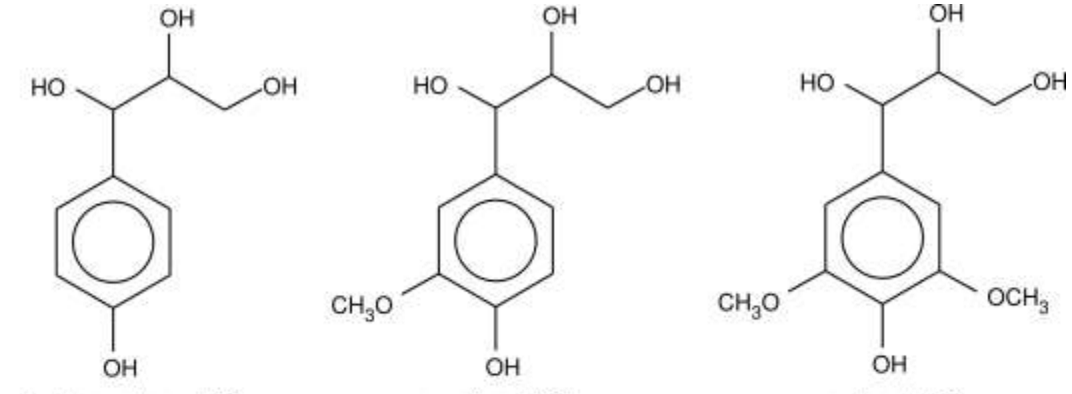
Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Am. Chem. Soc.* 203, *135*, 19048

Hemicellulose oligomers in lignin hydrogel



Hemicellulose model (solute): xylohexaose – a simplified model of xylan - in different conformations obtained from MD simulations

Lignin model (solvent): three different lignin monomers in a aqueous solution

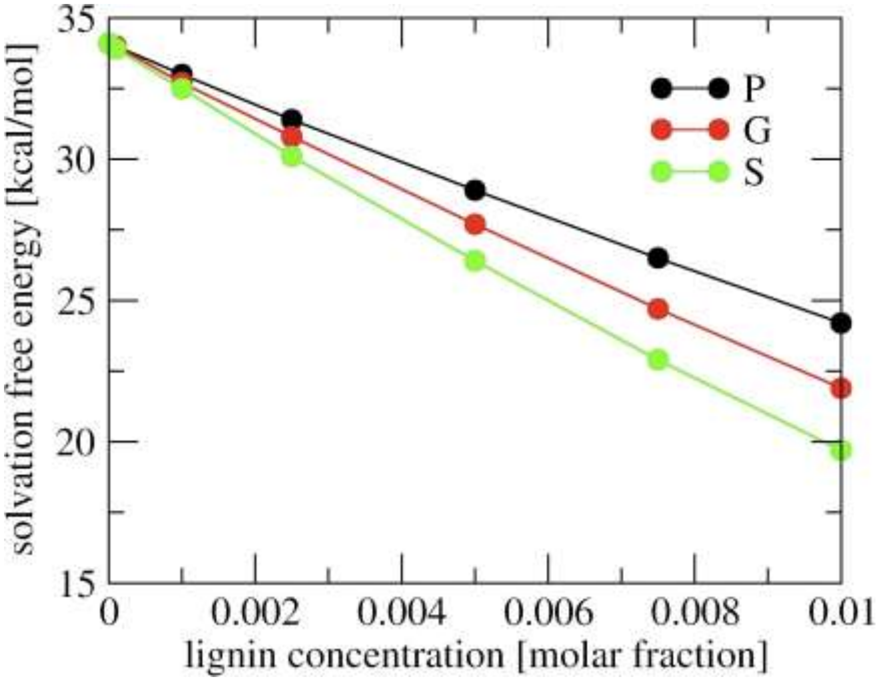


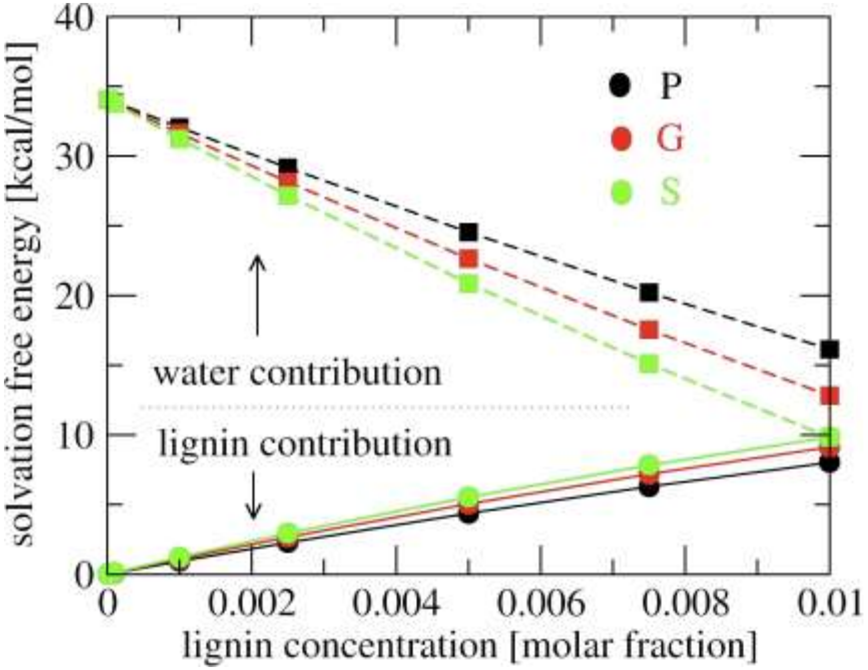
p-hydroxyphenyl (P) guaiacyl (G) syringyl (S)

Lignin monomers. The -OH in the middle of the aliphatic chain removed to mimic beta-O-4 lignin-lignin bond.

Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Phys. Chem. Lett.* 2015, 6, 206.

Hemicellulose in lignin hydrogel: Solvation Free Energy





- ☐ **SFE decreases in the order P > G > S**
- ☐ **addition of $-OCH_3$ groups enhances hydrophobic interactions**

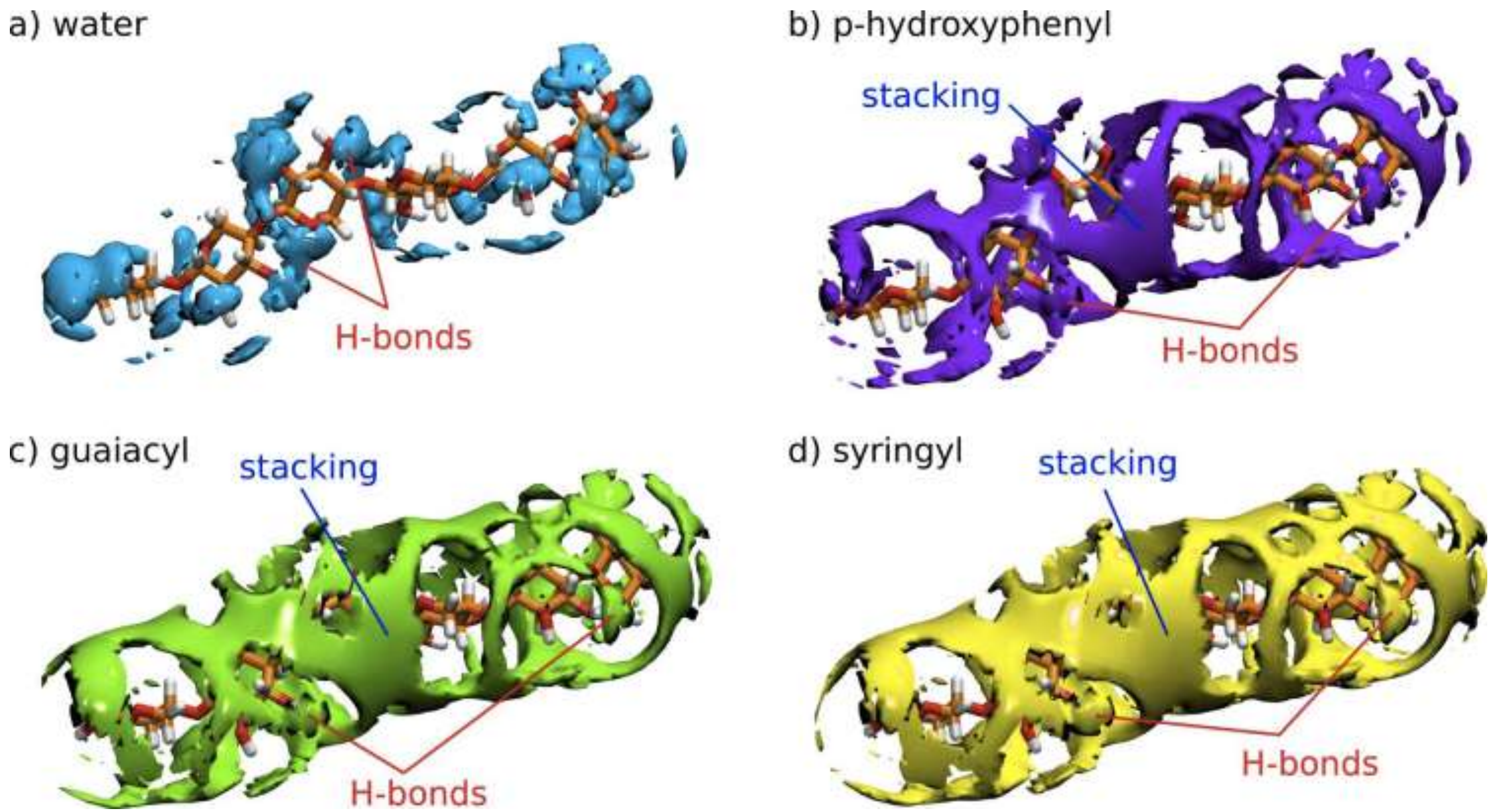
Lignin stabilizes the system indirectly, by expulsion of water from the hydrophobic surfaces of hemicellulose

Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Phys. Chem. Lett.* 2015, 6., 206

Hemicellulose in lignin hydrogel: Solvation Free Energy Density

hydrophobic interactions increase in the order P < G < S

Constant SFED isovalue

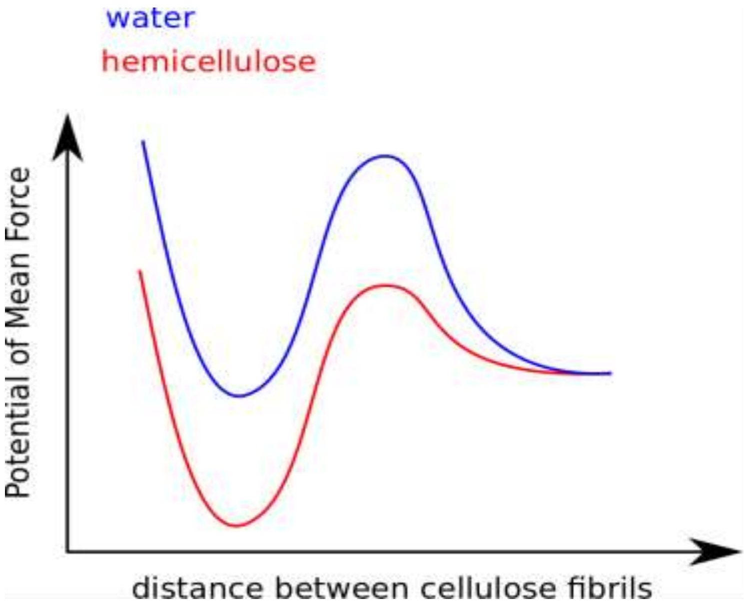


- For water, only specific H-bonding
- For lignin hydrogel, mainly hydrophobic stacking with some H-bonding

Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Phys. Chem. Lett.* (2015, 6, 206.

Summary on biomass conversion

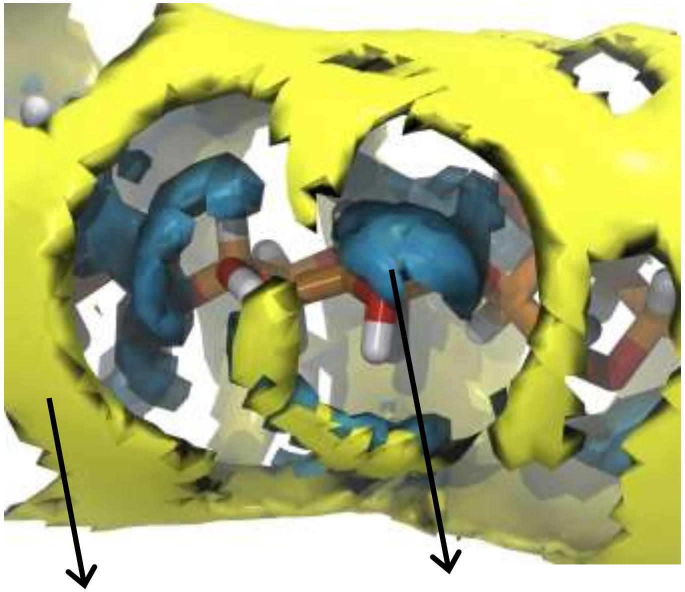
Hemicellulose-cellulose interactions favor the aggregation of cellulose microfibrils in the plant cell wall:



- **Basic groups of the glucuronate branches significantly contribute to cell wall stability**
- **The cellulose-hemicellulose interactions are highly site-specific and complementary**

Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; And Kovalenko, A. *J. Am. Chem. Soc*. 203, *135*, 19048

Lignin-hemicellulose interaction is mainly hydrophobic stacking:



Lignin – stacking

Water – H-bonding

- **Methoxy groups yield no H-bonding**
- **water – hydrophobic HC surface interactions are replaced by lignin – hydrophobic HC surface interactions**

Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A. *J. Phys. Chem. Lett.* 2015, *6*, 206.

CONCLUSIONS

Molecular theory of solvation – an essential part of multiscale modeling

- **Statistical-mechanical, 3D-RISM-KH molecular theory of solvation**

  Solvation structure and thermodynamics of liquid and solution systems
  of a given composition, thermodynamic conditions, and local environment:
  - **nonpolar and polar solvents, cosolvents, ionic liquids, electrolyte solutions**
  - **solid-liquid interfaces, macromolecules, biomolecules, nanoparticles, and nanoporous materials**
  - **polymeric solutions and melts, and soft matter systems**

- **Application to chemical nanosystems and nanobiomaterials, including:**
  - **Detection and placement of structural solvent in biomolecular systems**
  - **Protein-ligand mapping and fragment-based drug design**
  - **Cellulosic biomass and CNC based bionanomaterials**
  - **Adsorption on clays and other surfaces**
  - **Supramolecular systems in solution, including synthetic organic supramolecular nanoarchitectures (rosette nanotubes)**

Oh, Canada!..

Acknowledgments

Sergey Gusarov
Alex Kobryn
Nikolay Blinov
Maria Stepanova
Stanislav Stoyanov

Piotr Drabik
Kengo Ichiki
Qingbin Li
Lucian Livadaru
Dragan Nilkolic
Kyrylo Tabunshchyk
Takeshi Yamazaki

Lyudmyla Dorosh
Wenjuan Huang
Yansen Lauw
Xianjun Liu
Olga Liubimova
Radha K Motkuri
Nataraj S Pagadala
Bhalchandra S Pujari
Andrey Tokarev
Oleksandr Zelyak

Seigo Hayaki
Jakub Kaminski
Viktor Leontyev
Siriporn Pansri
Daisuke Yokogawa

Leonardo M da Costa
Jackeline Oliveira
Rodrigo L. Silveira
Mateus R. Lage

Yuri Yu Dmitriev (U St. Petersburg)
Ihor Omelyan (Inst. Cond. Matte r Phys,)
Andrij Trokhymchuk (Inst. Cond. Matter Phys,)

Fumio Hirata (Inst. Molec. Sci.; Ritsumeikan U)

Amber
Assisted
Model
Building with
Energy
Refinement

CHEMICAL COMPUTING GROUP

SCM Scientific Computing & Modelling

UFRJ

CAPES

MEXT MINISTRY OF EDUCATION, CULTURE, SPORTS, SCIENCE AND TECHNOLOGY-JAPAN

CNPq 60 ANOS

compute • calcul CANADA

WestGrid

UNICAMP

uff Universidade Federal Fluminense

INSTITUTE FOR OIL SANDS INNOVATION AT THE UNIVERSITY OF ALBERTA
Breakthrough Energy Research | Advanced Environmental Technologies
Imperial
Alberta Innovates Energy and Environment Solutions
FACULTY OF ENGINEERING

prion Alberta Prion Research Institute

Alberta Innovates Bio Solutions

Alberta Innovates Technology Futures

Canada NRC·CNRC

National Institute for Nanotechnology

UNIVERSITY OF ALBERTA

UBC THE UNIVERSITY OF BRITISH COLUMBIA