UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

GROMACS *inside* ™

- Electrostatics
- Force fields
- Polarizability
- Parallellization
- Future technology

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Electrostatics

- Coulomb's law gives the electrostatic interaction V_{ij} between particles i and j at distance r_{ij}:

$$V_{ij} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

UPPSALA
UNIVERSITET

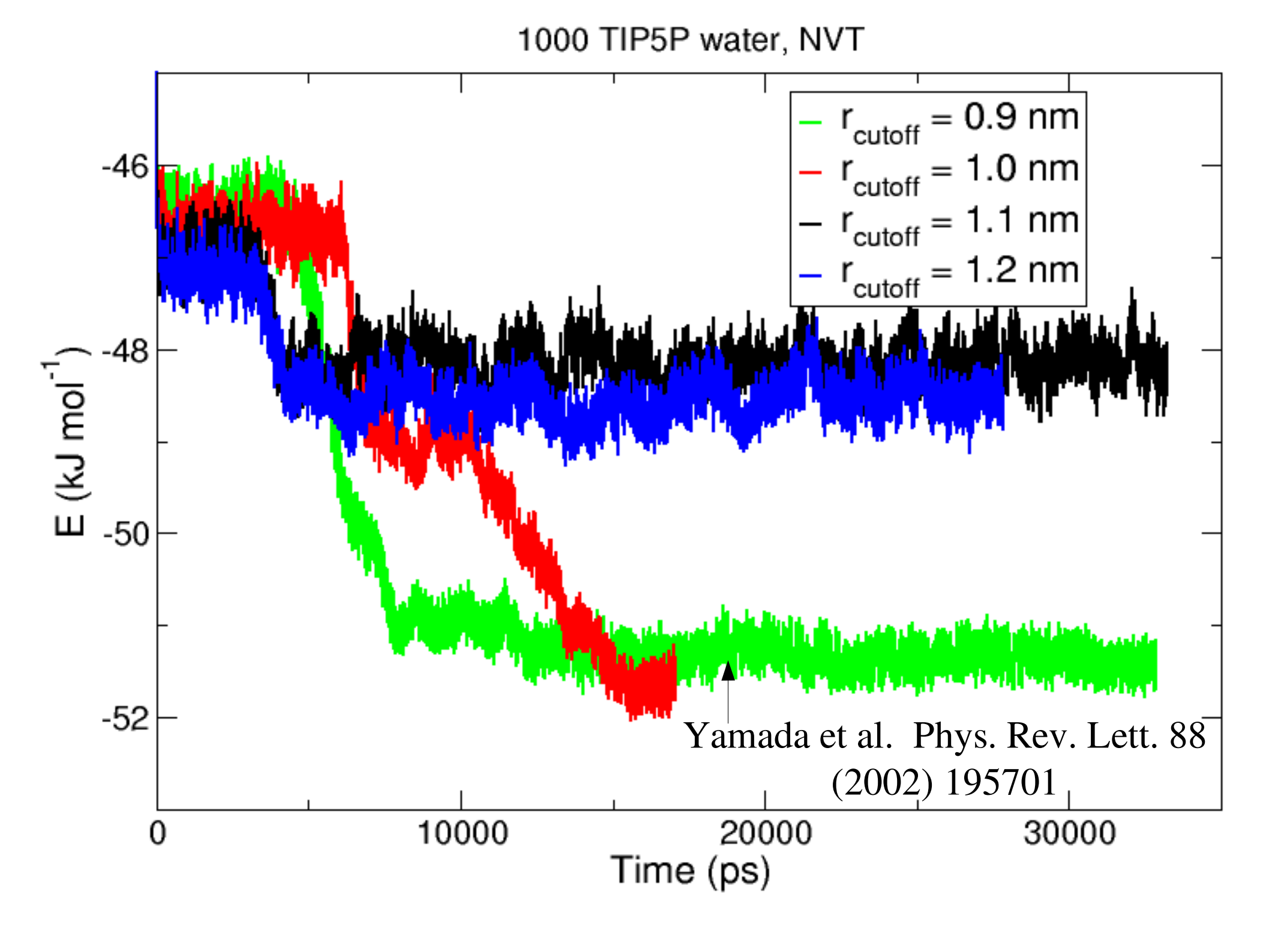
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Electrostatics

- Long ranged, in particular with full charges, but even with large dipoles like in water.
- Cut-off methods give huge artefacts


1000 TIP5P water, NVT
r cutoff = 0.9 nm
r cutoff = 1.0 nm
r cutoff = 1.1 nm
r cutoff = 1.2 nm
-46
-48
-50
-52
E (kJ mol-1)
0
10000
20000
30000
Time (ps)
Yamada et al. Phys. Rev. Lett. 88
(2002) 195701

UPPSALA
UNIVERSITET

David van der Spoel

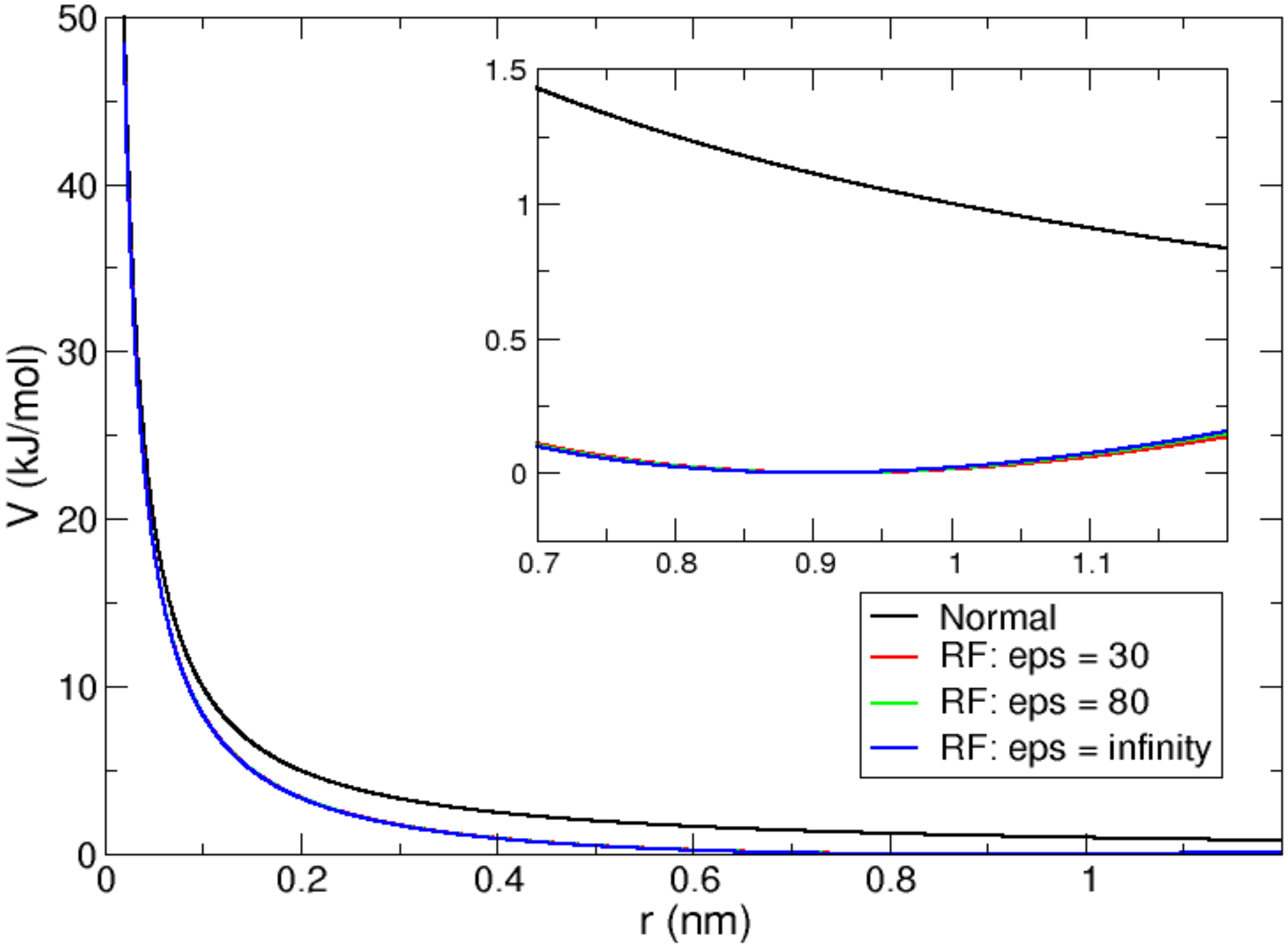
Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Electrostatics - Reaction field

- Assume there is a dielectric continuum ε_{RF} beyond the cut-off r_c that modulates the interaction between two charges.
- Interaction can then be written as:

$$V_{ij}^{RF} = V_{ij}\left(1 + \frac{\epsilon_{RF} - 1}{2\epsilon_{RF} + 1}\frac{r_{ij}^3}{r_c^3}\right)$$


50
40
30
20
10
0
V (kJ/mol)
0
0.2
0.4
0.6
0.8
1
r (nm)
1.5
1
0.5
0
0.7
0.8
0.9
1
1.1
Normal
RF: eps = 30
RF: eps = 80
RF: eps = infinity

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Electrostatics – Reaction field

- Cheap to compute with short r_c
- Reasonable results for dipolar liquids
- Problematic for charged groups
- Makes molecules more “slippery” (i.e., higher mobility)

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

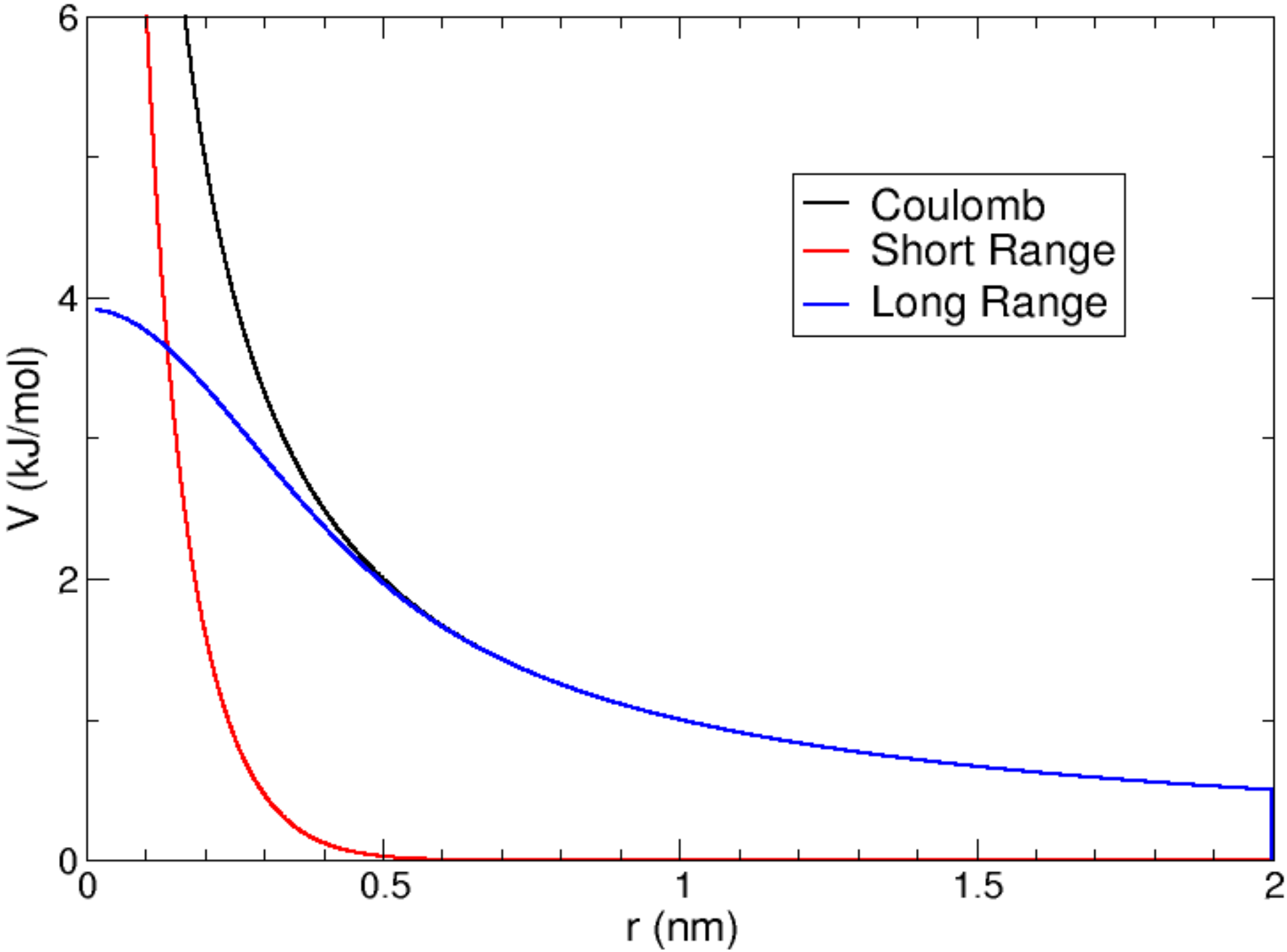
spoel at xray.bmc.uu.se

Electrostatics – Ewald

- Split the Coulomb interaction into two parts, a short range and a long range.

$$V_{ij}^{SR} = V_{ij}\,\mathrm{erfc}\left(\frac{r_{ij}}{\beta}\right)$$

$$V_{ij}^{LR} = V_{ij}\left[1 - \mathrm{erfc}\left(\frac{r_{ij}}{\beta}\right)\right]$$


Coulomb
Short Range
Long Range
V (kJ/mol)
r (nm)
0
2
4
6
0
0.5
1
1.5
2

UPPSALA UNIVERSITET

David van der Spoel

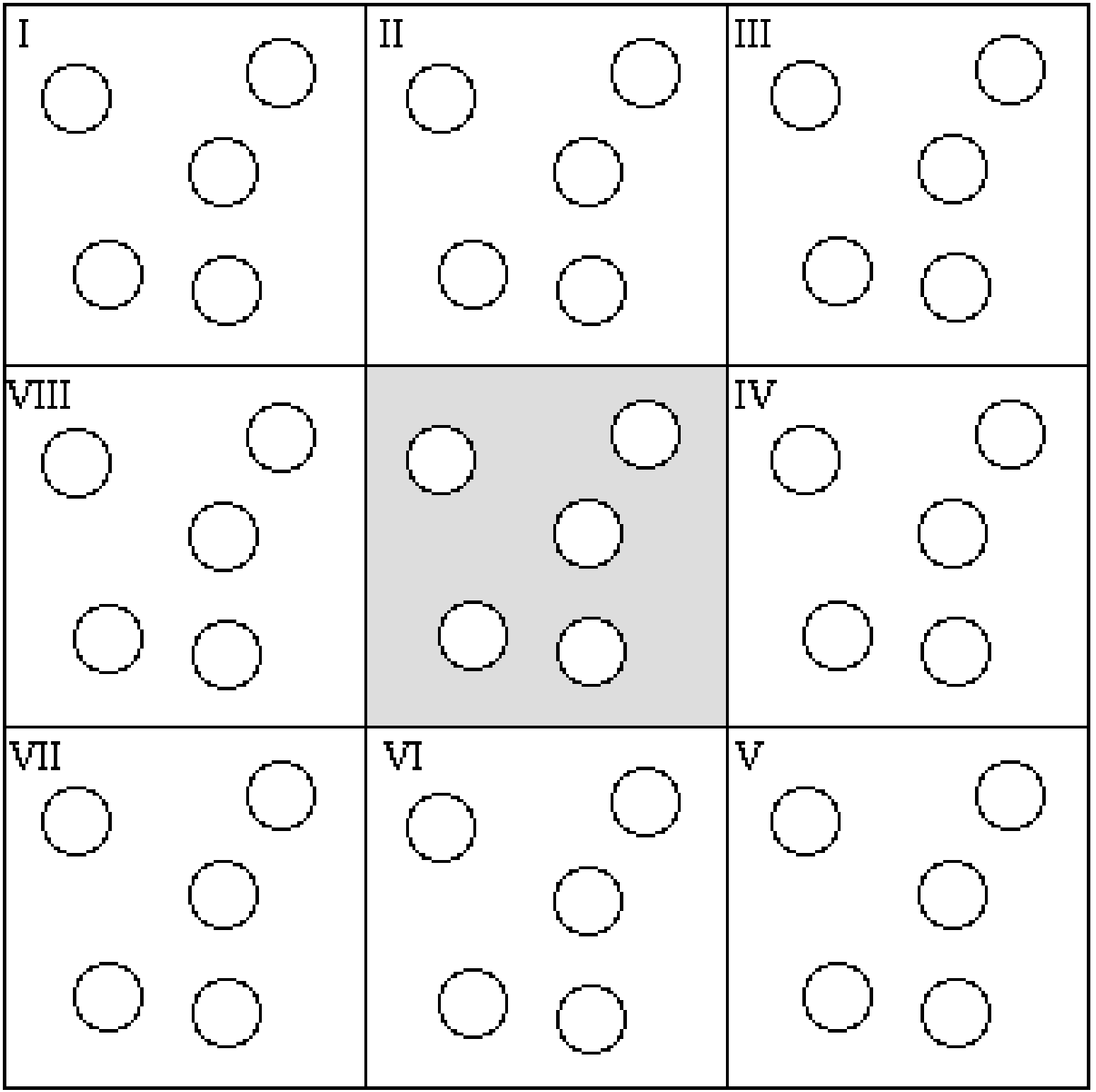
Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Electrostatics – Ewald

- Compute the short range part of the potential directly
- Compute the long range part by solving the modified Poisson equation in Fourier space (for an infinitely periodic sample):

$$\Delta\phi(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0}$$


I
II
III
VIII
IV
VII
VI
V

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Electrostatics – Ewald

- Long range part can be done analytically (classical Ewald summation)
- Long range on a grid can be done using fast Fourier transform

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

Electrostatics – Ewald

- Practical considerations:
- Grid spacing
- Cut-off distance
- Distribution of charges on a grid using splines (need to do the inverse when computing the forces).

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Force field status

- GROMOS and OPLS supported
- Amber underway
- CHARMM unknown?

UPPSALA
UNIVERSITET

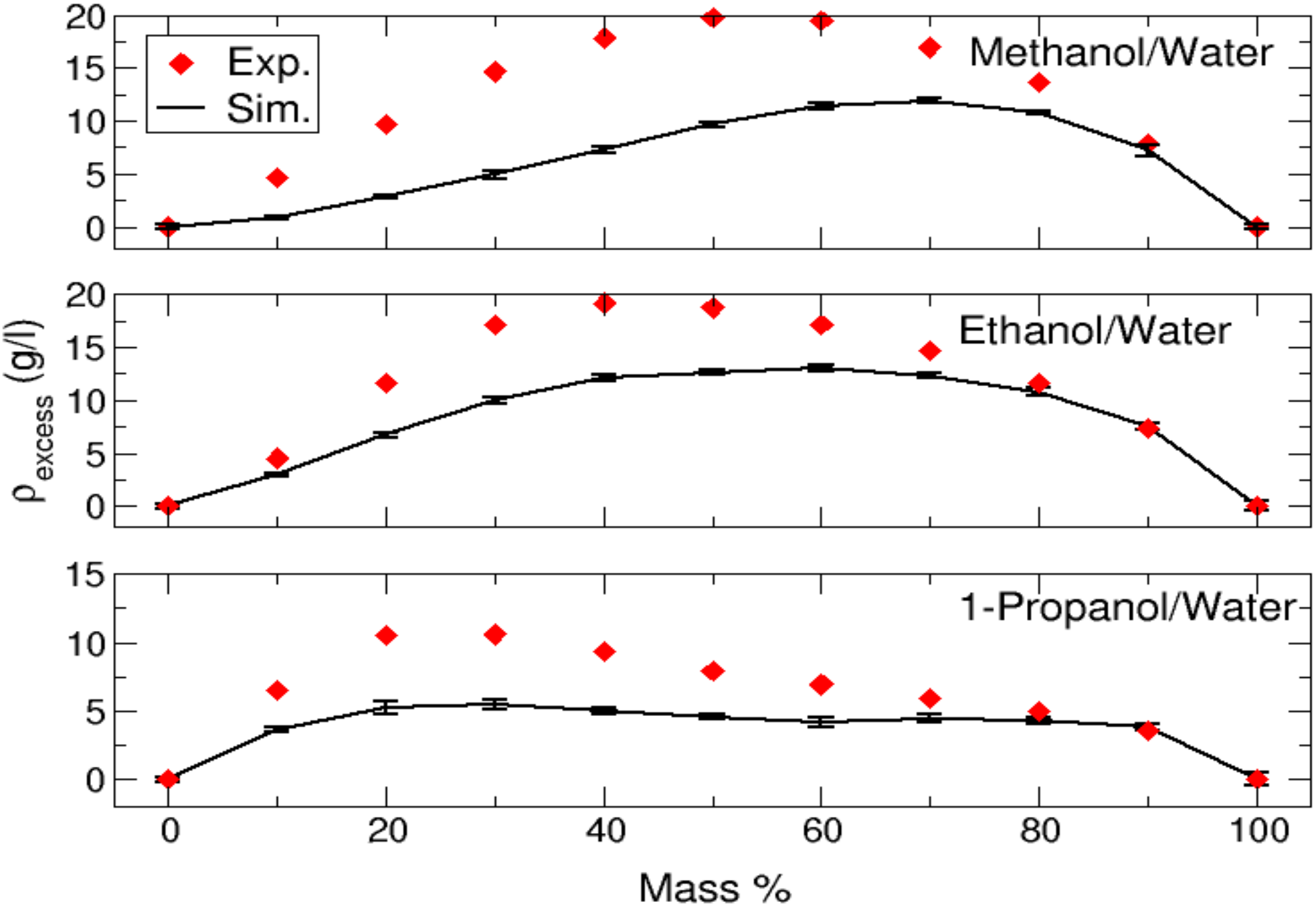
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

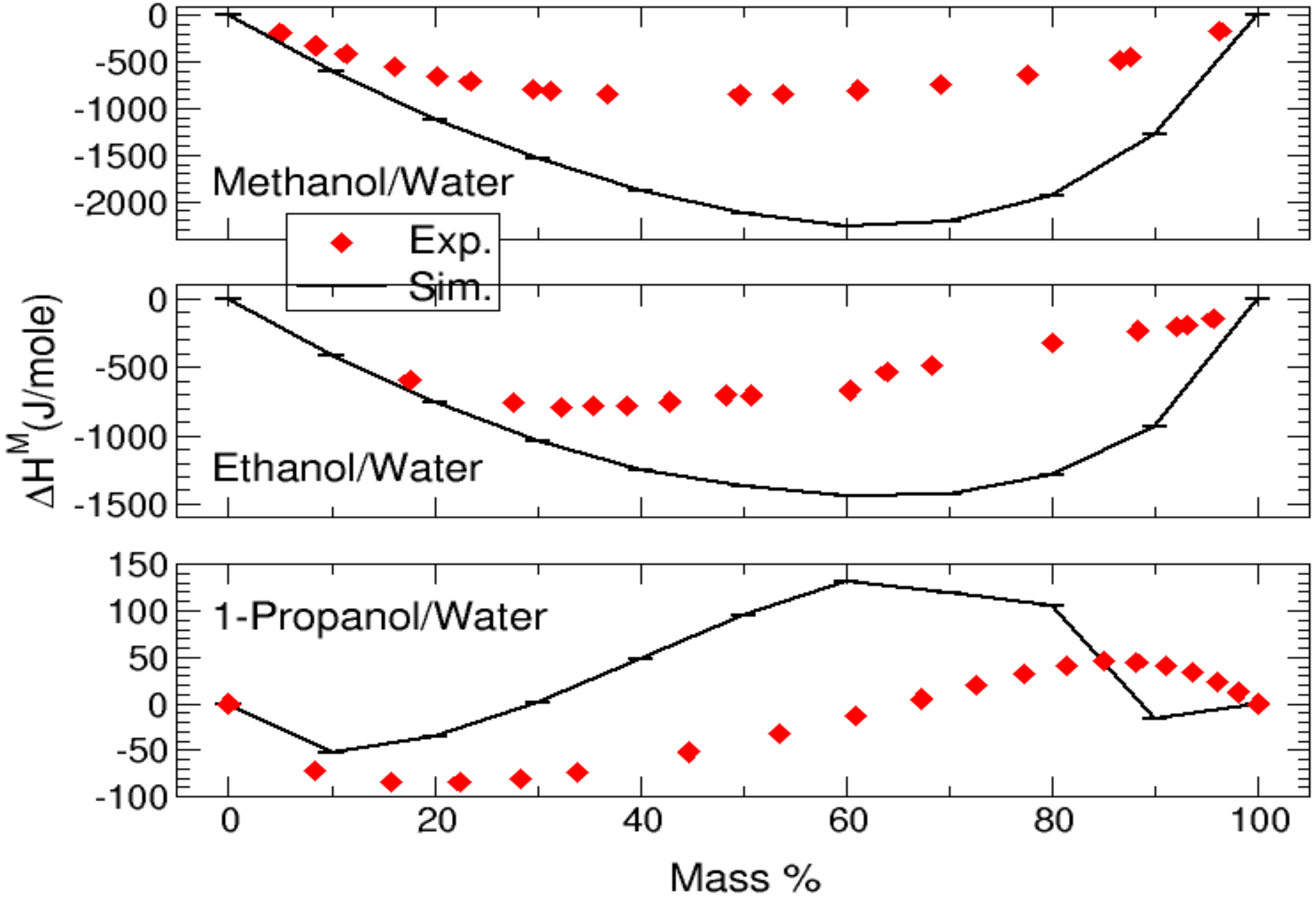
spoel at xray.bmc.uu.se

Force field status

- How to test force fields?
- Organic molecules



Wensink et al. J. Chem. Phys. 119 (2003) 7308



Wensink et al. J. Chem. Phys. 119 (2003) 7308

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Force field status

- Organic molecule simulations teach us that force fields are qualitatively correct but not quantitatively yet.
- Combination of too large excess enthalpy and too low excess density may mean that potentials are too attractive overall ***and*** too repulsive at short distance
- Solution: get rid of the 12 term in Lennard Jones potential.

UPPSALA
UNIVERSITET

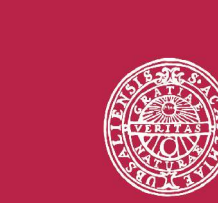
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Force field status

- How to test force fields?
- Protein simulations (Van der Spoel & Lindahl, J. Phys. Chem. B 107 (2003) 11178-11187)



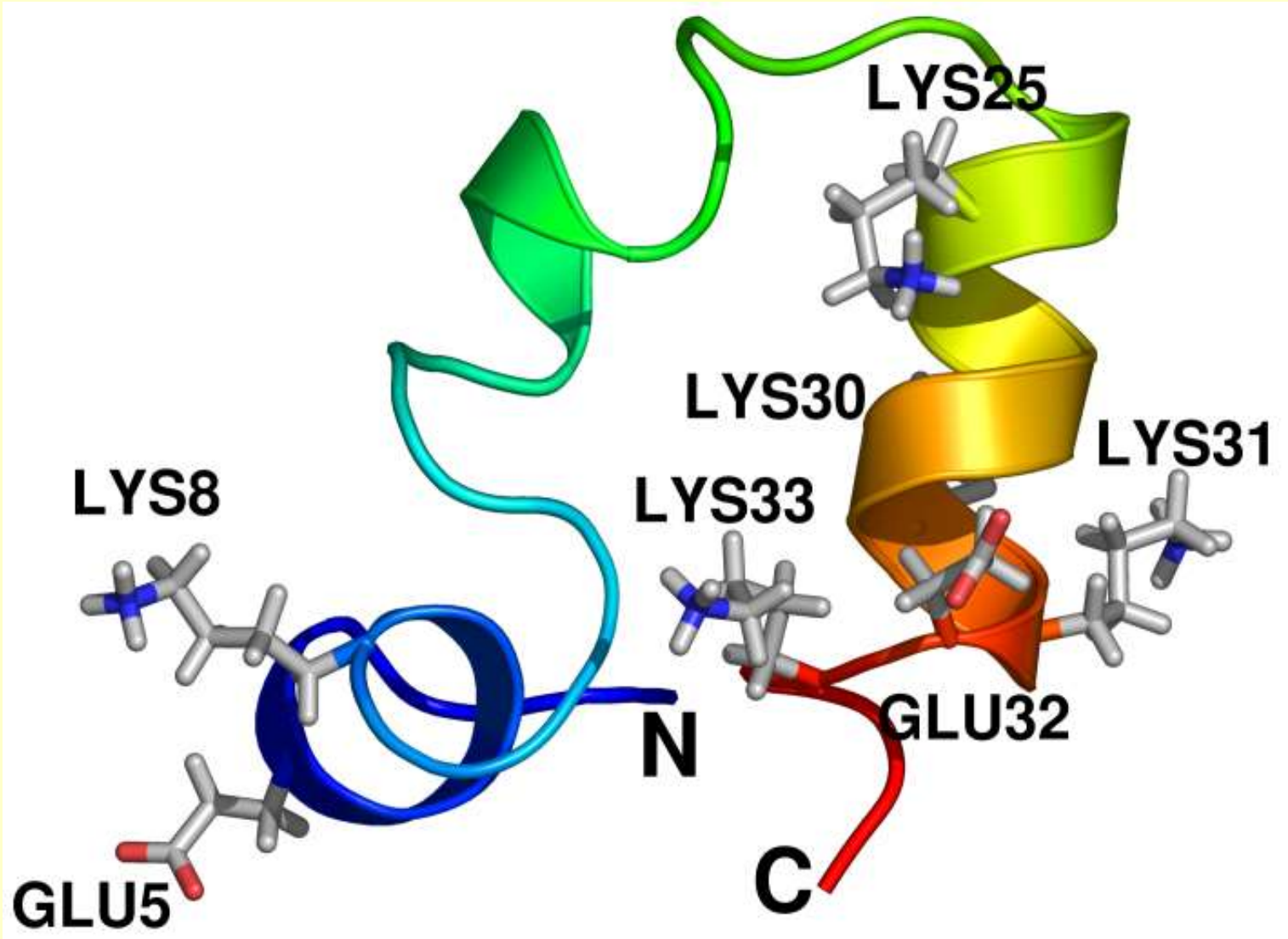
UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Villin headpiece subdomain



UPPSALA
UNIVERSITET

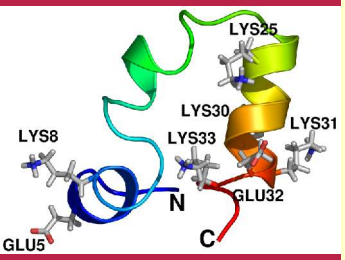
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

Villin headpiece subdomain

- 36 residue mini protein
- NMR Structure known
- Thermostable
- Hydrophobic core
- Protein stable under single mutations of the core
- Studied by by Duan and Kollman in a single microsecond simulation



UPPSALA
UNIVERSITET

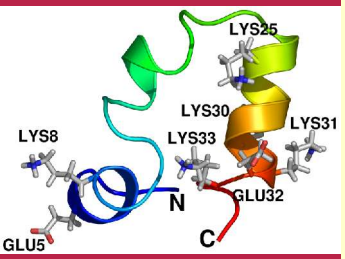
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

Correspondence with NMR

- Compute distances derived from NOE data
- Compare simulated to experimental results
- The *violation* of experimental distances (largest violation, number of violations and the sum of violations) determines the quality of the simulation



UPPSALA UNIVERSITET

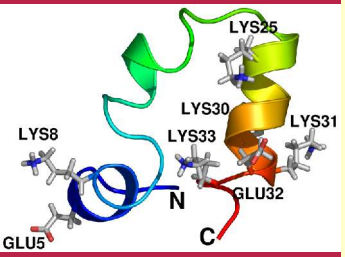
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

NMR distance restraints

Sim.	$N_{viol}(325)$	Max_{viol}(nm)	σ_{viol}(nm)
NMR	42	0.17	1.7
GROMOS	31	0.15	1.2
OPLS	42	0.22	2.0

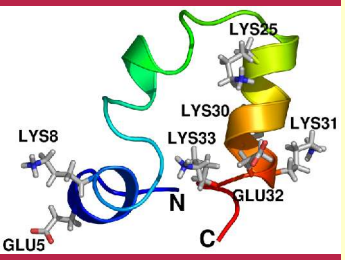


UPPSALA
UNIVERSITET

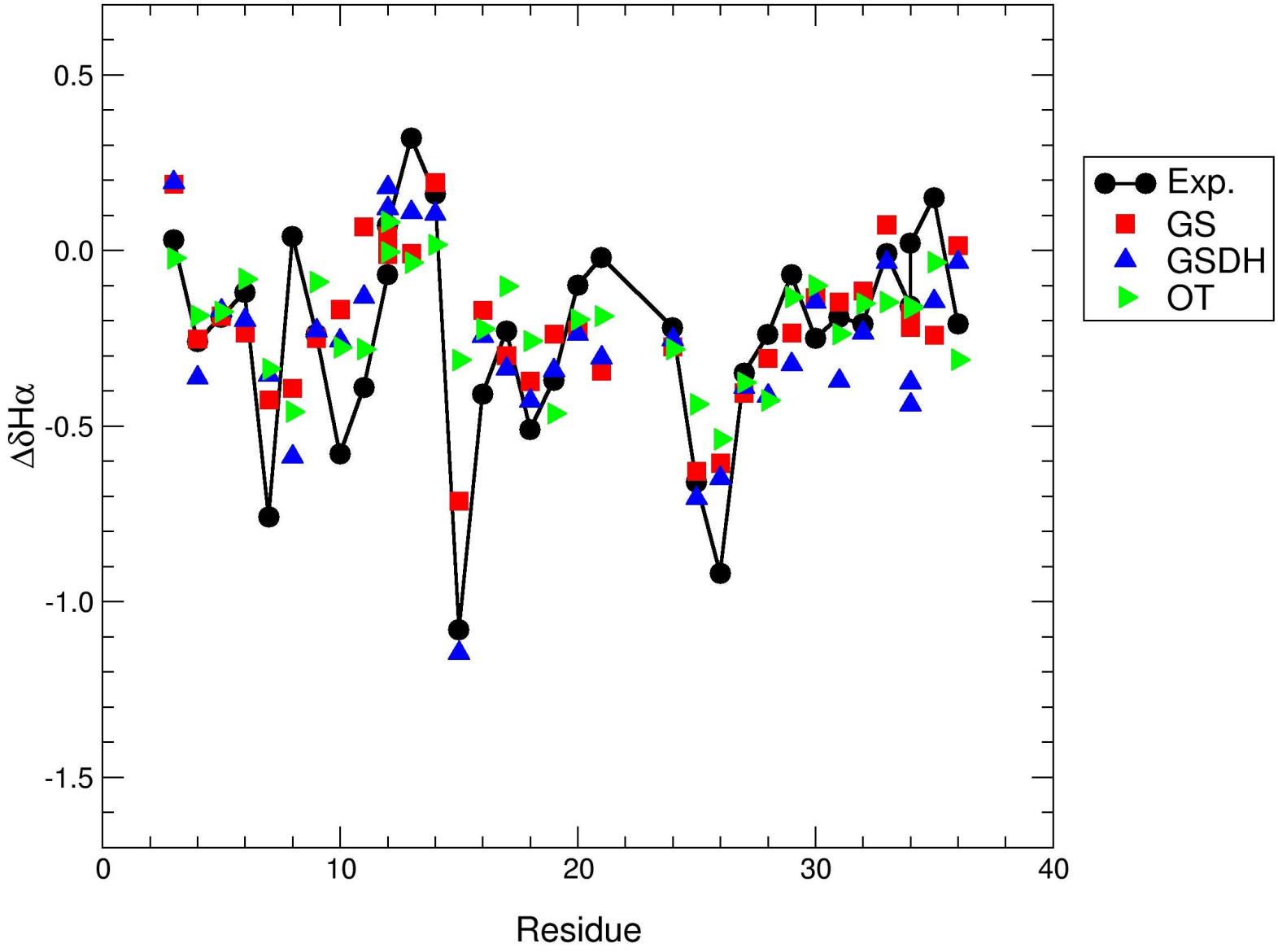
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se



NMR Chemical shifts



UPPSALA UNIVERSITET

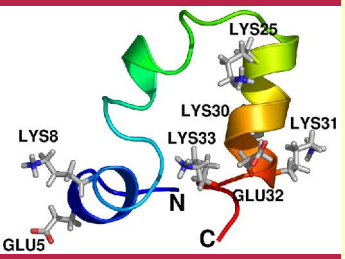
David van der Spoel

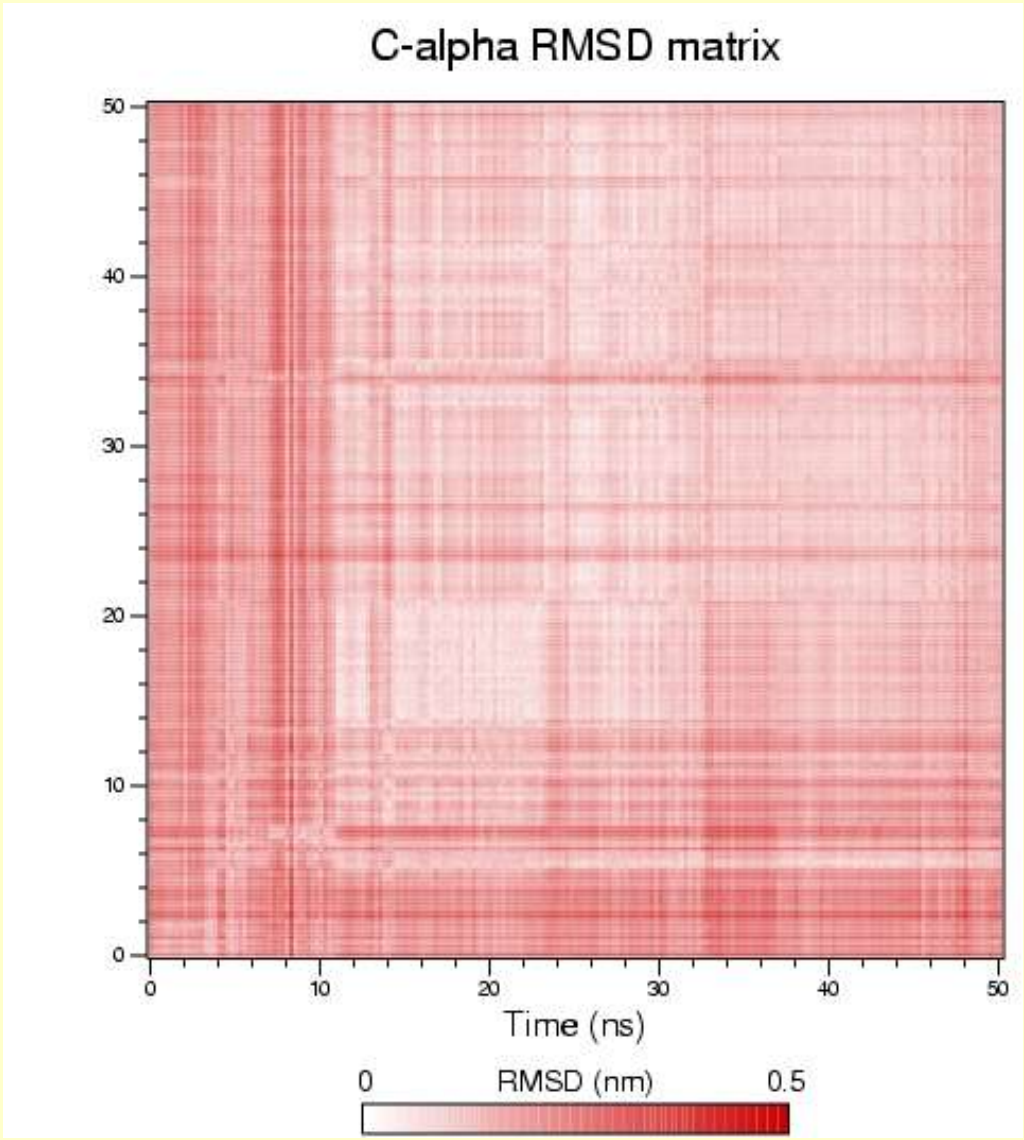
Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Effect of starting structure

- OPLS
- Max RMSD
- 50 ns





UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Force field status

- Simulation ensemble comparable to NMR ensemble, but not perfect compared to NMR data
- GROMOS slightly better than OPLS for Villin
- For Trp-cage it was the other way around (not shown)

UPPSALA
UNIVERSITET

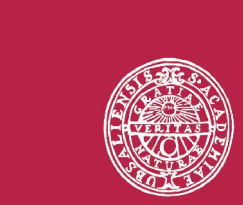
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Polarizability

- Phase transferability
- Interfaces
- Inhomogeneous systems
- Protein-metal interactions



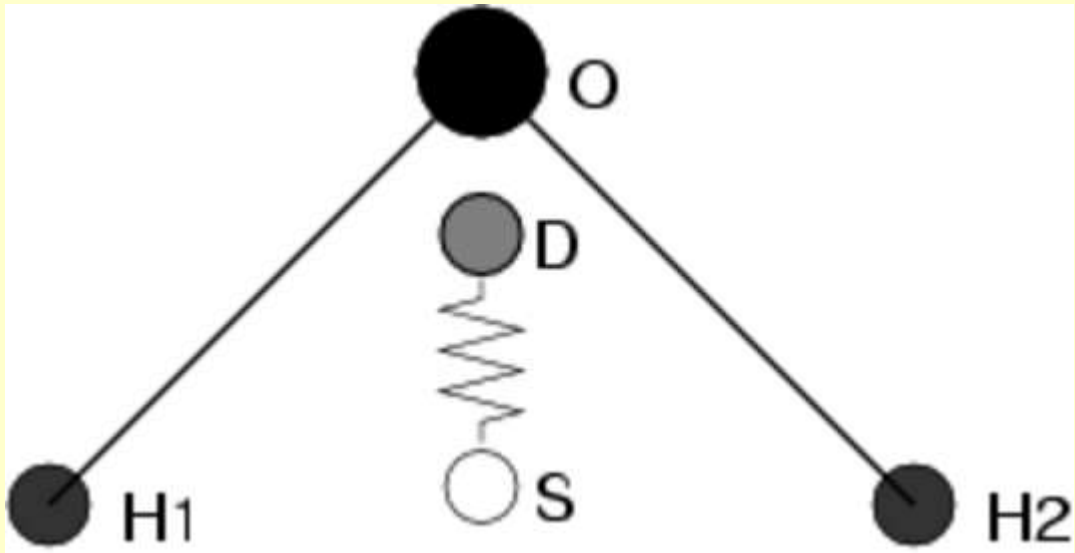
UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

Polarizability



Shell Water (SW)
Van Maaren & Van der Spoel,
J. Phys. Chem. B 105 (2001) 2618

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Polarizability

- $k\, r_{SD} = q_s\, E$
- $\mu_{ind} = \alpha E$
- $k = q_s^2/\alpha$

UPPSALA UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Polarizability – water dimer

Model	R_{OO}	E_{pot}	μ
SW	0.295	-21.78	2.59
Dang	0.287	-19.62	
MCDHO	0.292	-20.9	2.68
TIP4P	0.274	-26.35	2.70
SPC/E	0.274	-30.1	3.76
Ab Initio	0.291	-21.0	
Exper.	0.295	-22.6(3)	2.60

UPPSALA
UNIVERSITET

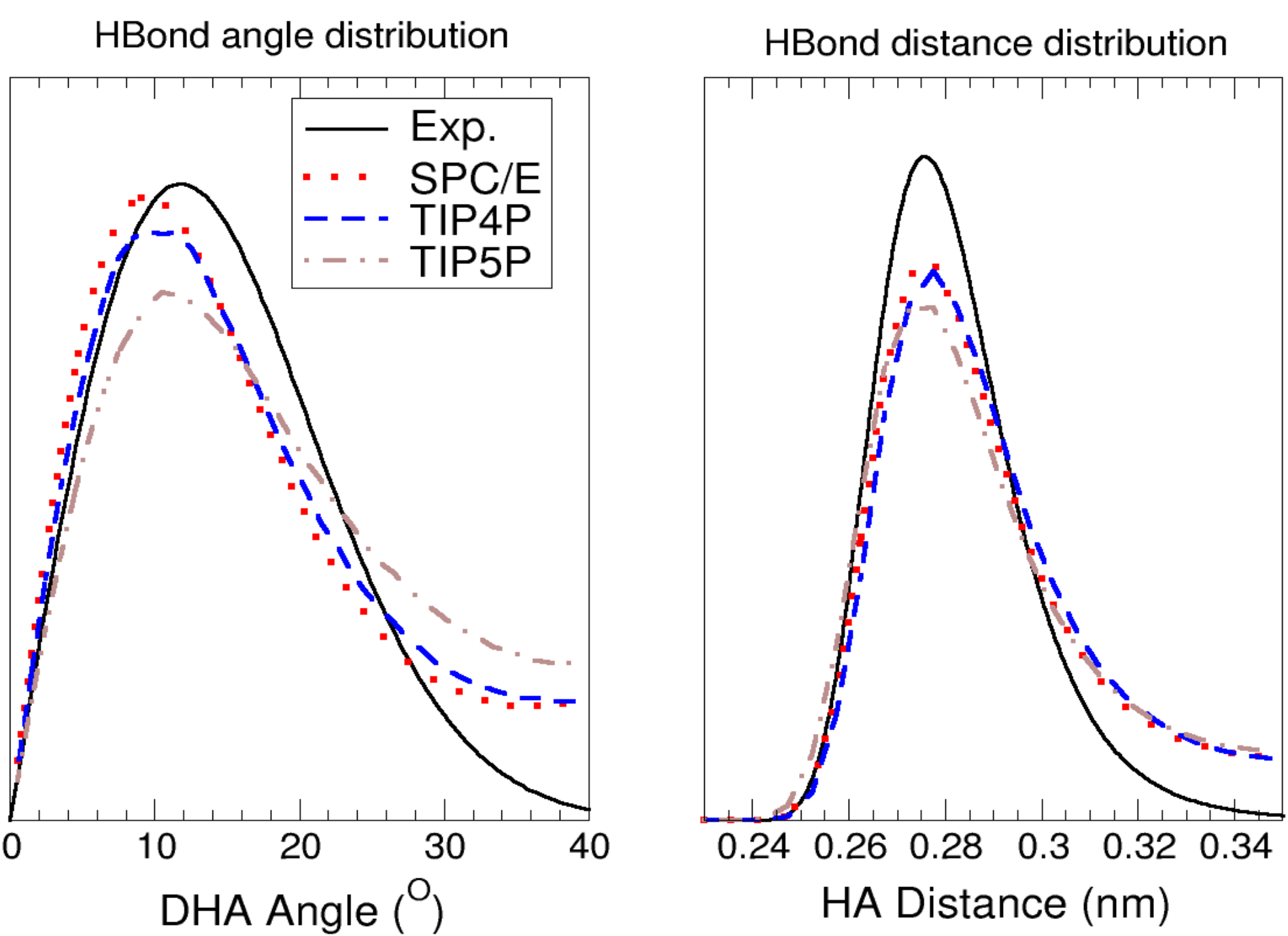
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Polarizability

- Further work in progress, based on ab initio calculations for HF, Water, NH3, CH4.
- Automatic parameterisation tools available in GROMACS



Liquid water data: Modig et al Phys. Rev. Lett. 90 (2003) 075502

UPPSALA
UNIVERSITET

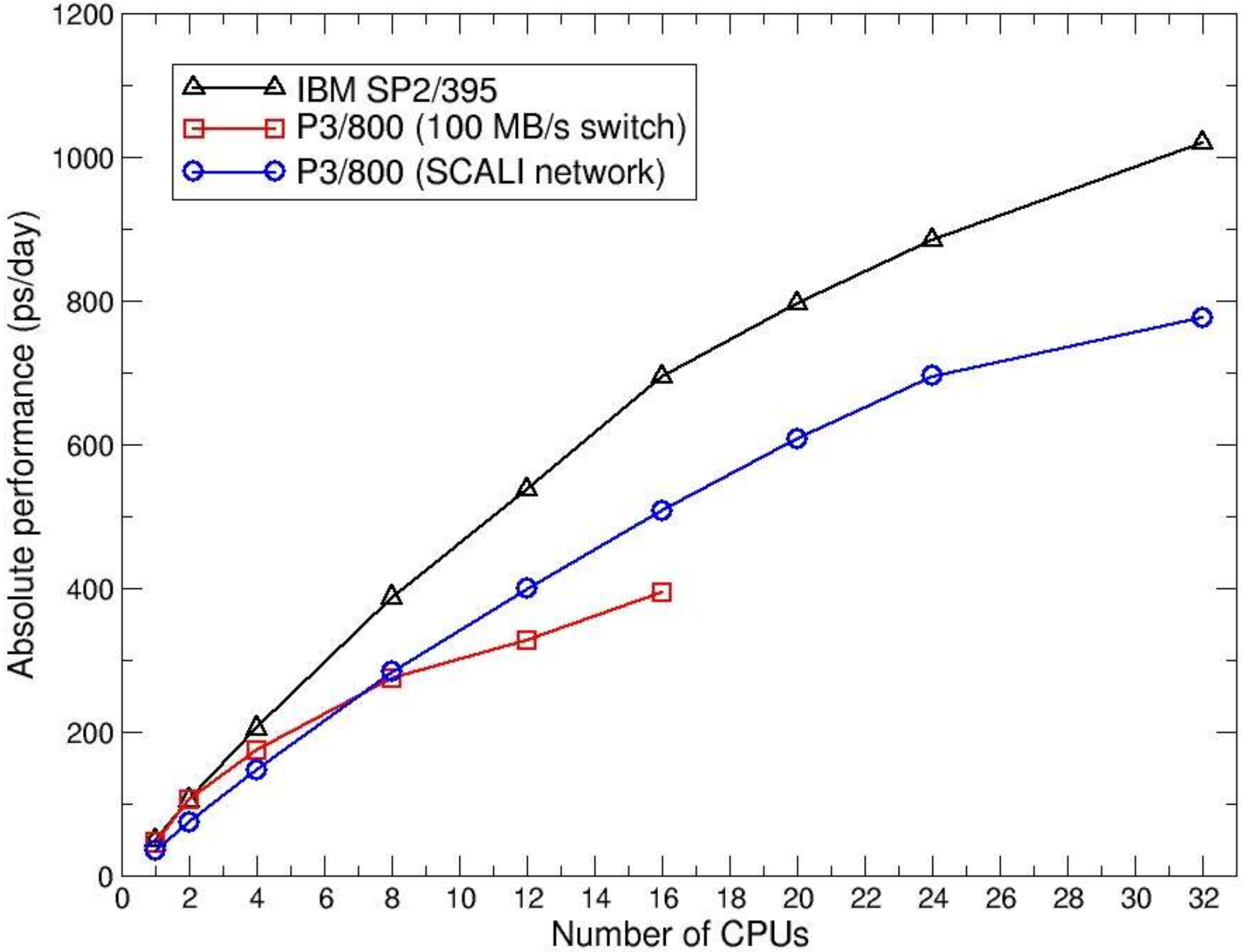
David van der Spoel

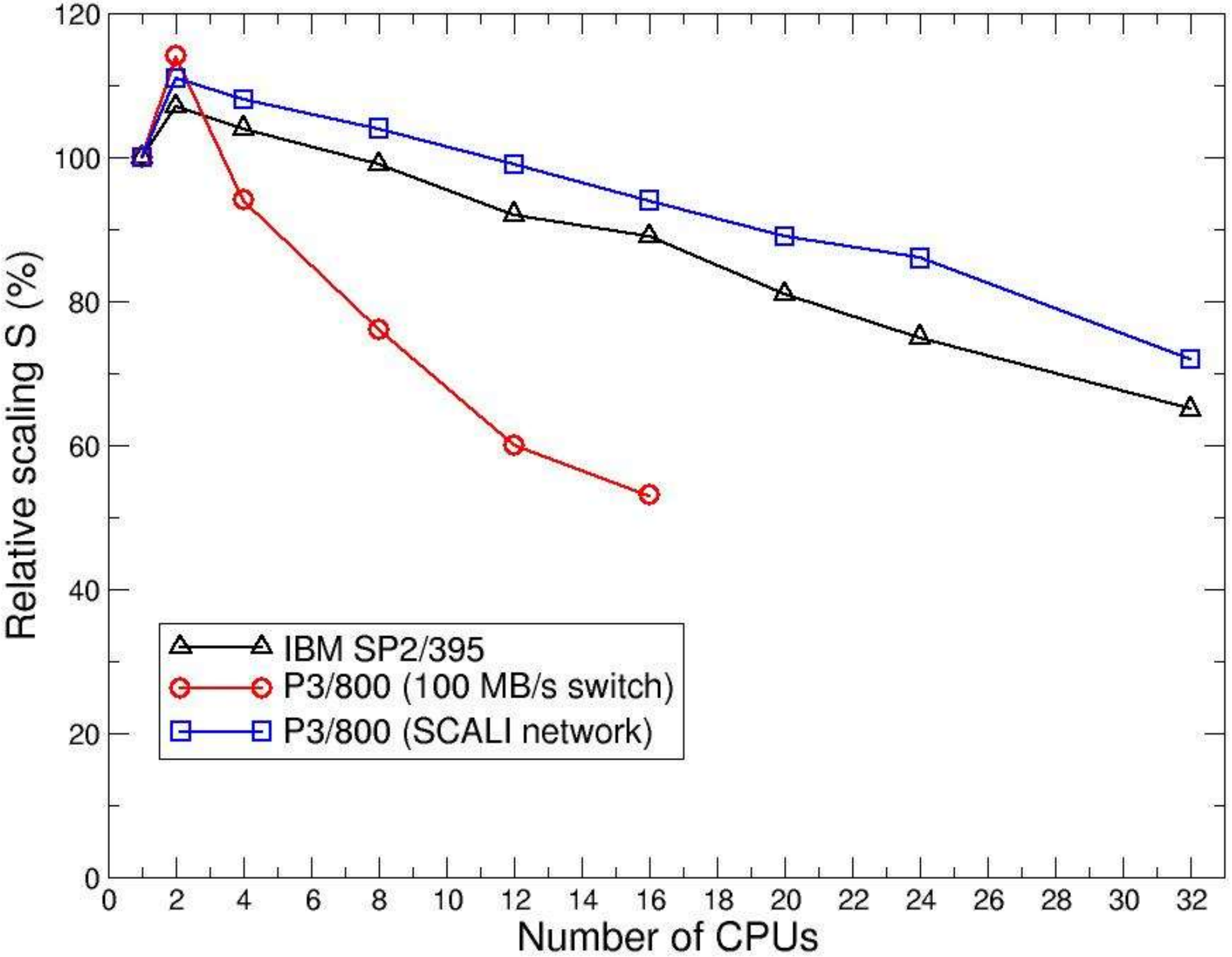
Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

Parallellization

- Particle decomposition (particles are allocated to processors, which is simple algorithmically)
- In principle SIMD
- Gromacs runs quite efficient with a cut-off, if the system is large enough
- Not with PME


IBM SP2/395
P3/800 (100 MB/s switch)
P3/800 (SCALI network)
Absolute performance (ps/day)
Number of CPUs
0
200
400
600
800
1000
1200
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32


Relative scaling S (%)
Number of CPUs
IBM SP2/395
P3/800 (100 MB/s switch)
P3/800 (SCALI network)
0
20
40
60
80
100
120
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Future parallellization

- Domain decomposition (chunks of space are given to each processors, more complex to program)
- PME on a few dedicated processors: real MIMD

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

Future technology

- XML for file formats

UPPSALA
UNIVERSITET

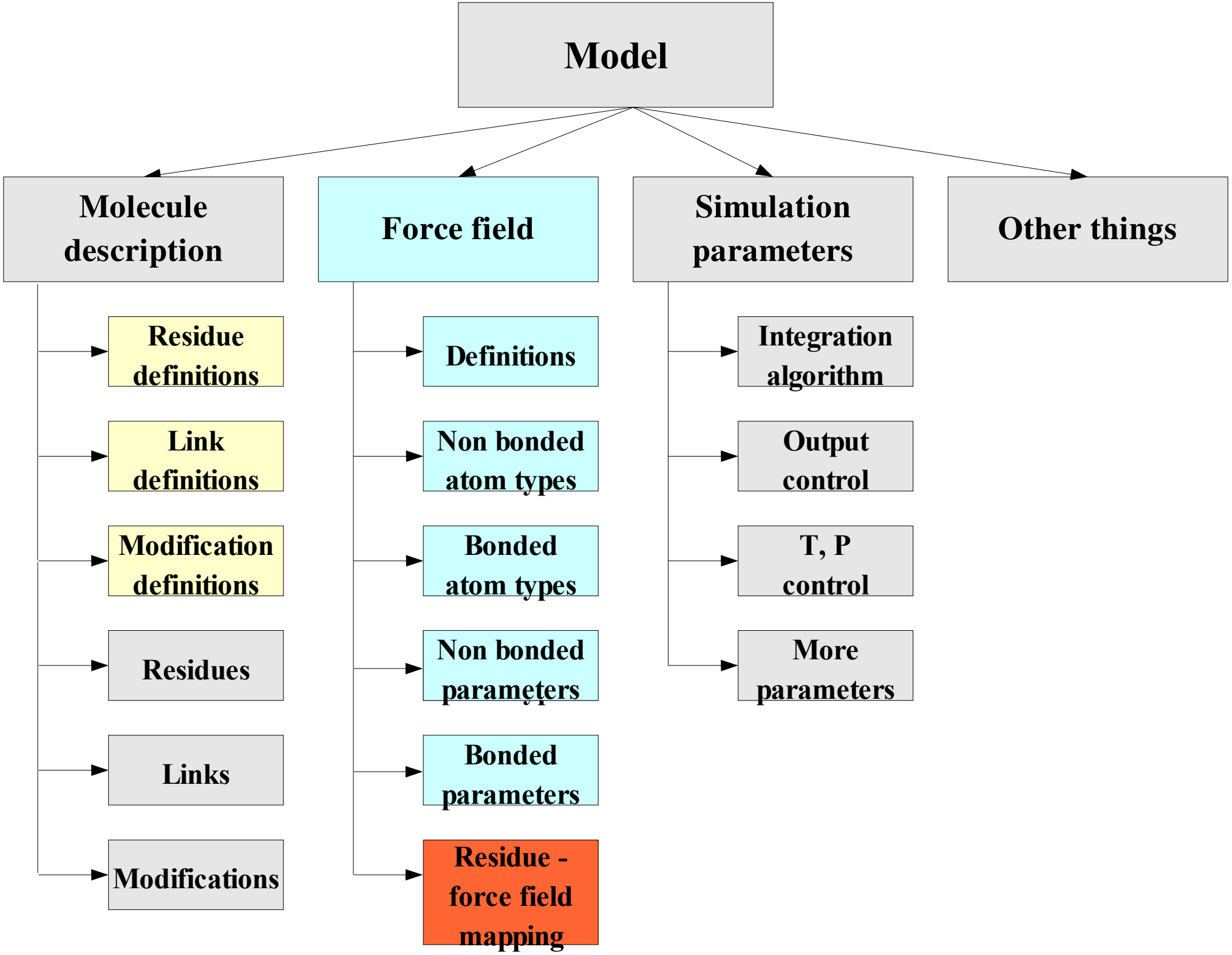
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Why use XML?

- Growing complexity of input files
- Self contained file format which can/could be validated
- Separate data from algorithms (anti object oriented!)
- Extensible?
- Machine readable (using e.g. libxml2)


Model
Molecule description
Force field
Simulation parameters
Other things
Residue definitions
Link definitions
Modification definitions
Residues
Links
Modifications
Definitions
Non bonded atom types
Bonded atom types
Non bonded parameters
Bonded parameters
Residue - force field mapping
Integration algorithm
Output control
T, P control
More parameters

UPPSALA
UNIVERSITET

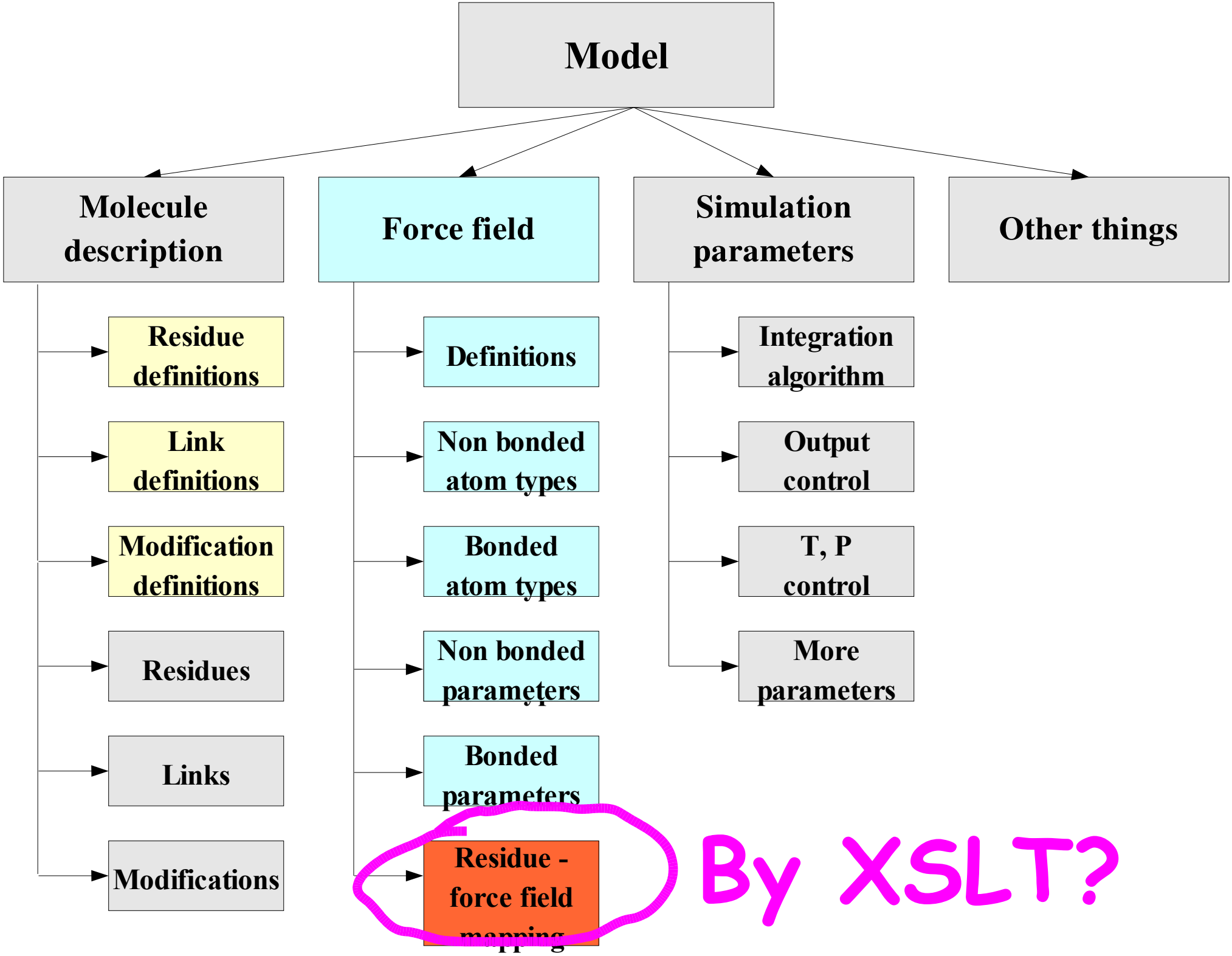
David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

Is this the right approach?

- DTD versus Schema
- Storing information in attributes
- Transformations using XSLT?
- Use CML?


Model
Molecule description
Force field
Simulation parameters
Other things
Residue definitions
Link definitions
Modification definitions
Residues
Links
Modifications
Definitions
Non bonded atom types
Bonded atom types
Non bonded parameters
Bonded parameters
Residue - force field mapping
Integration algorithm
Output control
T, P control
More parameters
By XSLT?

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

Why use Python?

- Easy to program
- Portable
- Can be used for GUI design using several underlying libraries (wxWindows, FLTK, PyGTK, Qt-Python etc.)
- Can be used as a generic scripting language (rather than a dedicated scripting language)

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Python in GROMACS...

- In the first hand for developing a GUI
- Using tools/libraries in a simple manner (NetCDF/XML)
- Interfacing with other programs (MMTK, pymol, VMD)

UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Python in GROMACS...

- Use the GROMACS library function `parse_common_args` to generate python scripts
- A simple generic dialog box routine using Tkinter

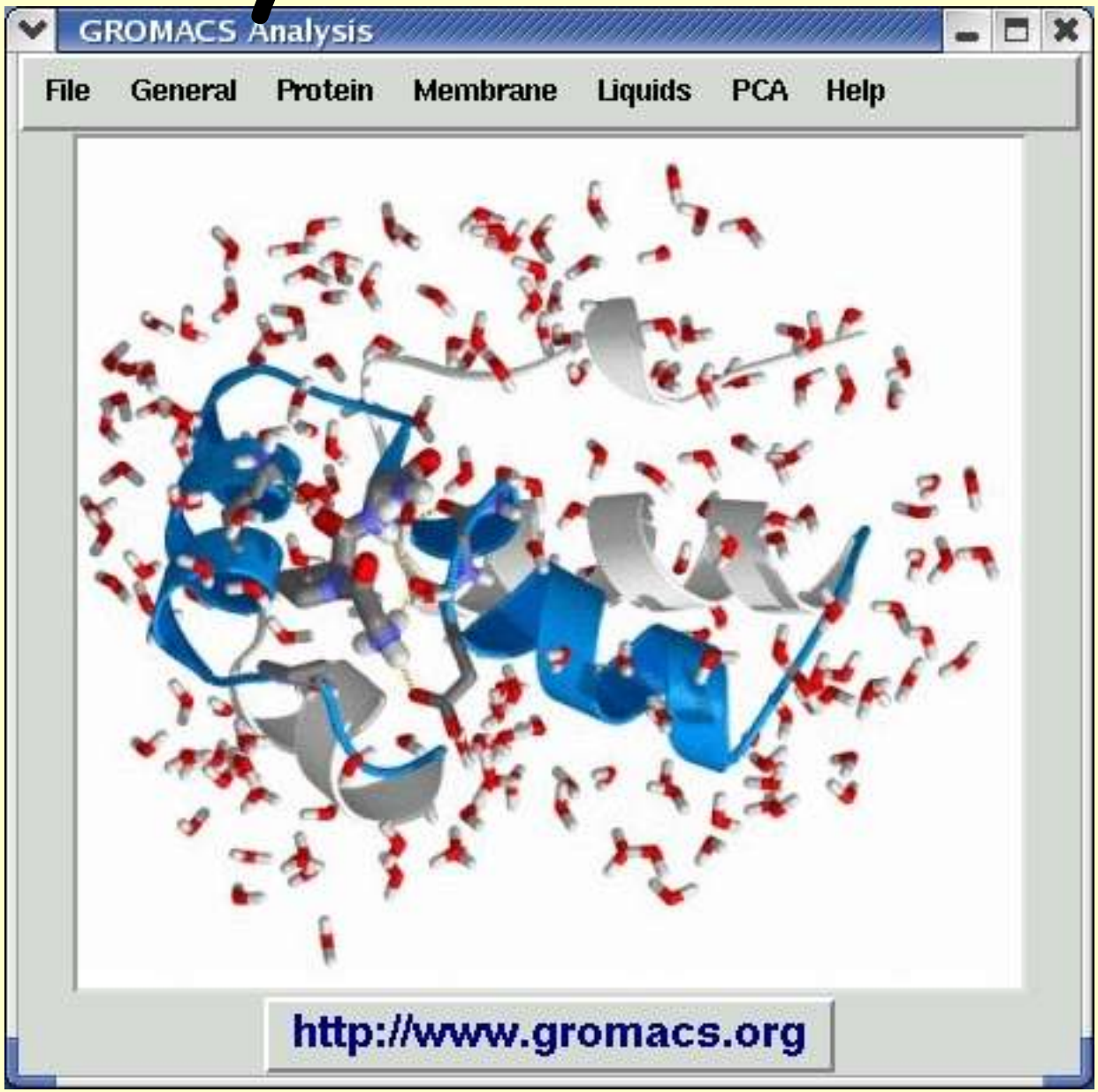
UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at xray.bmc.uu.se

Analysis front-end



UPPSALA
UNIVERSITET

David van der Spoel

Molecular Biophysics Group, Department of Cell and Molecular Biology

spoel at
xray.bmc.uu.se

Acknowledgements

- Berk Hess (Groningen)
- Erik Lindahl (Stanford)
- Anton Feenstra (Amsterdam)
- Herman Berendsen (Groningen)
- Konrad Pywowarczik (Krakow)
- Alexandra Patriksson, Michiel van Lun (Uppsala)