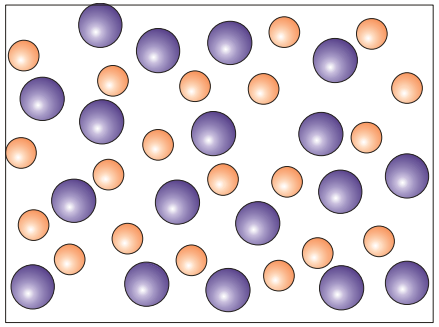
Steady state of a protein-ligand system in a temperature gradient investigated with Wang-Landau simulations

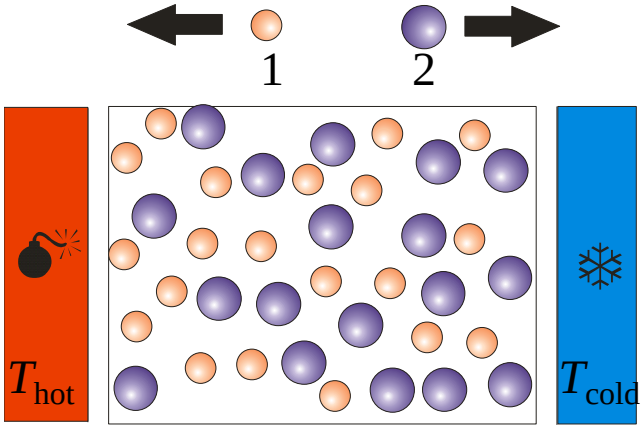
Jutta Luettmer-Strathmann
Departments of Physics and Chemistry, The University of Akron, Akron, OH

- Thermodiffusion – Ludwig-Soret effect
- Chemical contribution to the Soret coefficient
- Proteins – HP model
- Square-well potentials
- Wang-Landau algorithm for the density of states
- Chain collapse and ligand binding
- 2-d Wang-Landau for Soret effect

Thermodiffusion — Ludwig-Soret Effect

Fluid mixture with uniform temperature





under a temperature gradient $\delta T/L$

Mass flow due to temperature difference

$$J = -\rho D(\delta c - c(1-c)S_T \delta T)/L$$

In the steady state

$$J = 0 \quad \Rightarrow \quad \delta c = -c(1-c)S_T \delta T$$

T = temperature

c = mass fraction of component 2

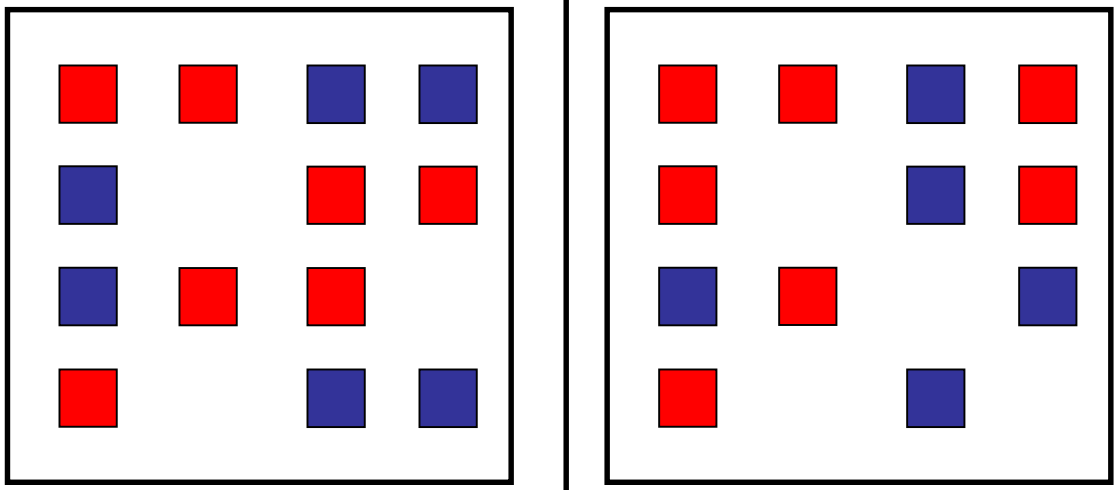
Soret coefficient of component 2:

$$S_T = -\frac{1}{c(1-c)}\frac{\delta c}{\delta T}$$

Chemical contribution to the Soret coefficient

Two-chamber lattice model

Chamber A, temperature T^A **Chamber B, temperature T^B**



$\Delta T = 10^{-4}$ K

$N^A = N^B = N/2$

Sum of states over all occupations $\{N_i^A\}$:

$$Q = \sum_{[N_i^A]} Z^A(\{N_i^A\}) Z^B(\{N_i - N_i^A\}), \quad i \in \{a, b, v\}$$

Calculate Q in exact enumeration by considering all possible occupations.

Average mass fractions of component a in chambers A and B

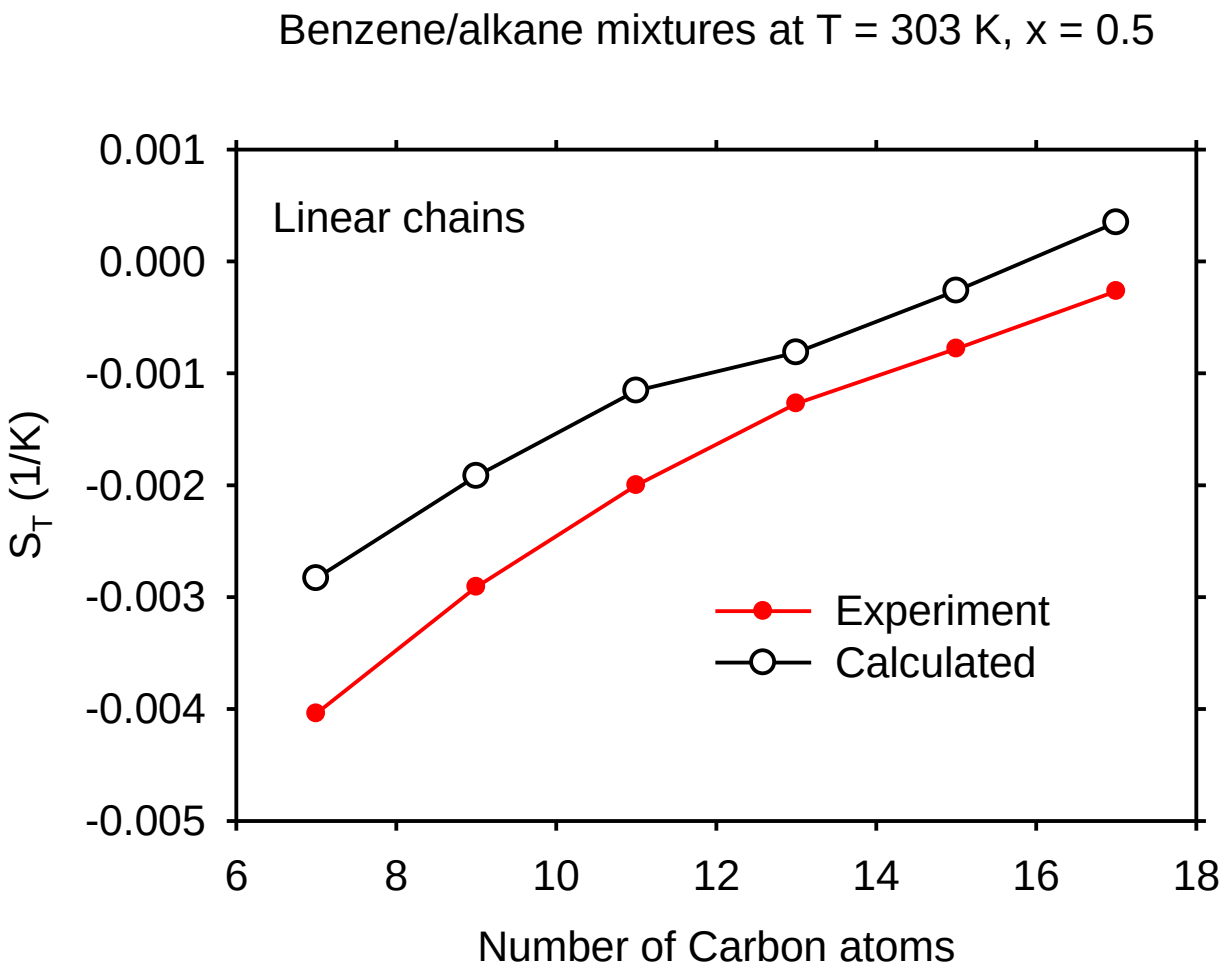
$$\left\langle c_a^A \right\rangle = \sum_{[N_i^A]} \frac{Z^A(\{N_i^A\}) Z^B(\{N_i - N_i^A\})}{Q} c_a^A(\{N_i^A\}), \quad i \in \{a, b, v\}$$

$$\left\langle c_a^B \right\rangle = \sum_{[N_i^A]} \frac{Z^A(\{N_i^A\}) Z^B(\{N_i - N_i^A\})}{Q} c_a^B(\{N_i^A\}), \quad i \in \{a, b, v\}$$

Soret coefficient of the alkane $S_T = -\frac{1}{c_a(1-c_a)} \frac{\langle c_a^A \rangle - \langle c_a^B \rangle}{T^A - T^B}$

- Determine occupation of a lattice with N = 5000 sites for the desired temperature, pressure, and composition.
- Perform exact enumeration for sum of states of two chambers with $N^A = N^B$ = 2500 and accumulate mass fractions of the chambers.
- Evaluate including orientation-dependent interactions of the solvent
- Calculate S_T.

Benzene with linear alkanes at fixed T and x, experiment and theory



P. Polyakov, J. Luettmer-Strathmann, and S. Wiegand, J. Phys. Chem. B **110**, 2006, 26215-26224

Chemical contribution to the Soret coefficient

Two-chamber lattice model

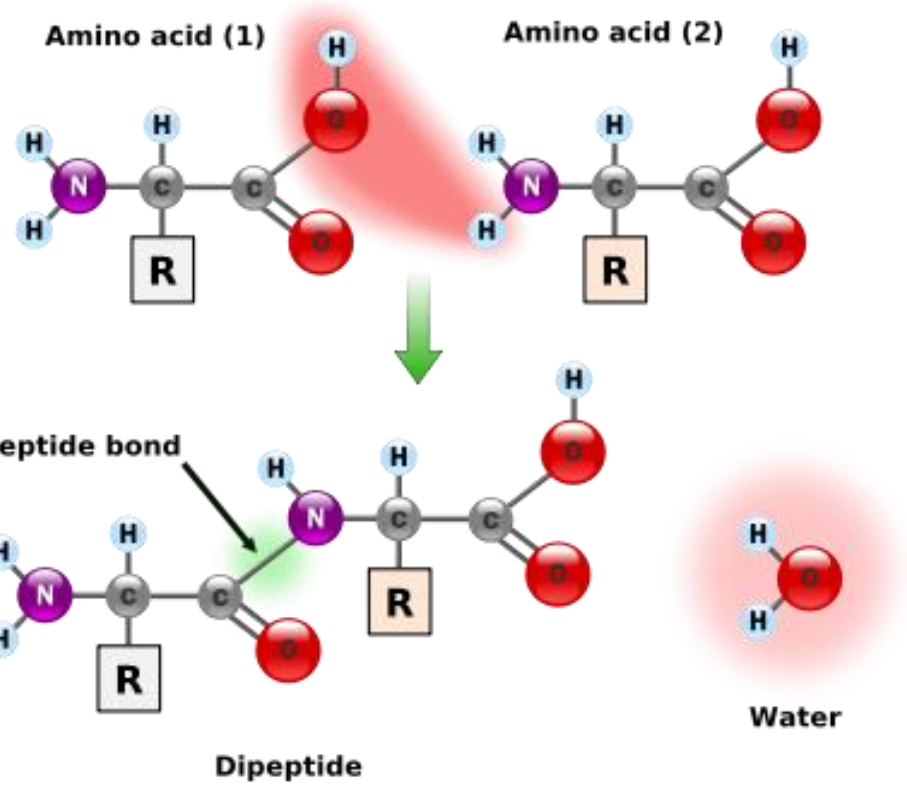
Advantages:

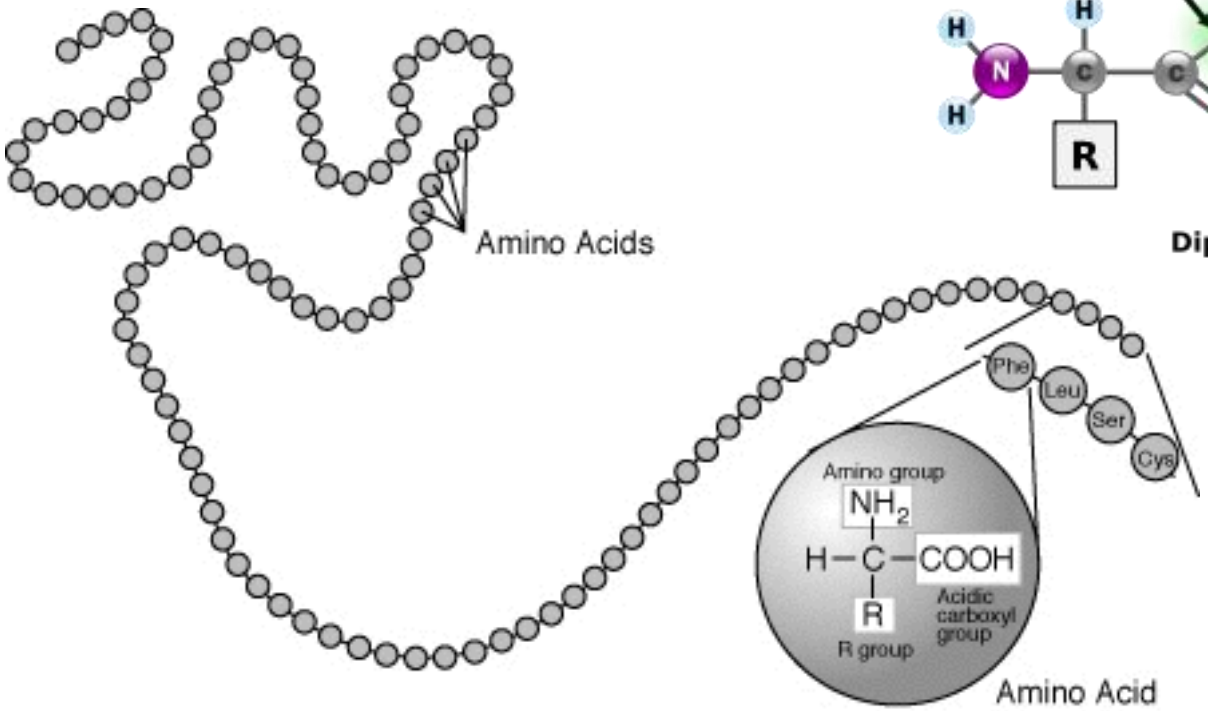
- Consistent equation of state
- Includes compressibility effects
- Works with very small temperature gradients
- Reasonable prediction of chemical contributions to Soret coefficients

Disadvantages:

- Lattice model equations of state are valid in limited temperature/density ranges
- The lattice structure limits the possible interactions
- Exact enumerations are limited to small systems
- No dynamics, even if a different method is used for the model.

Proteins are polymers of amino acids



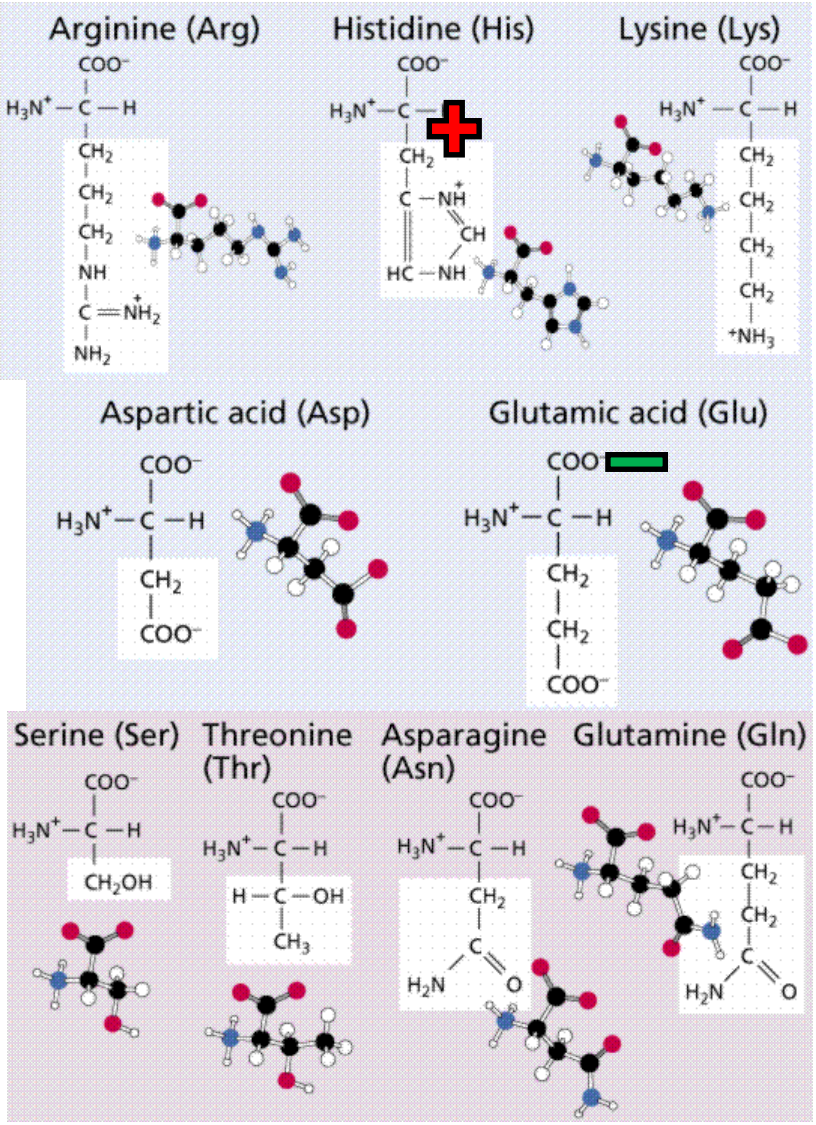


Primary structure:

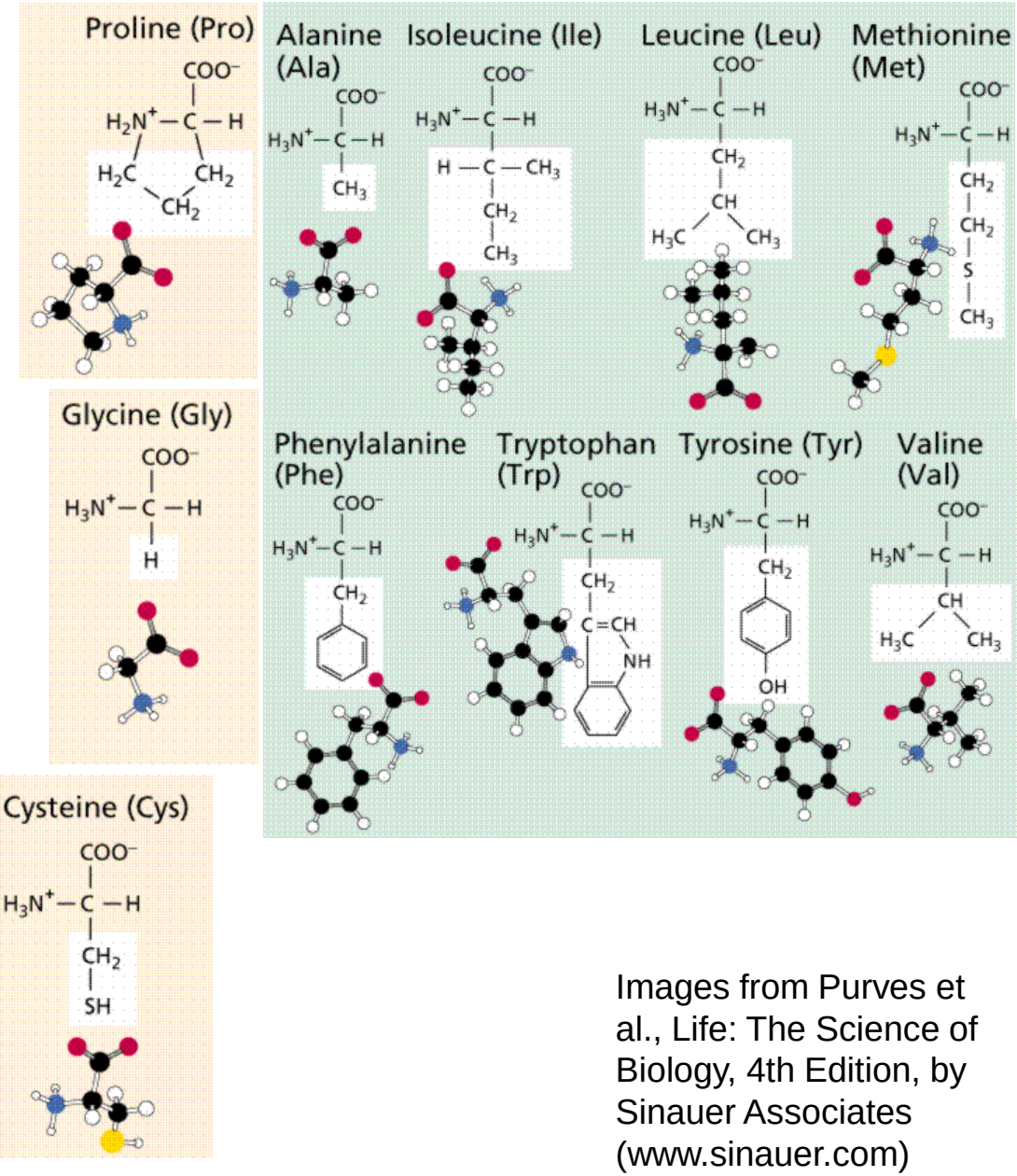
Sequence of amino acids along the chain

Source: Wikipedia

Polar residues (P)



Nonpolar residues (H)



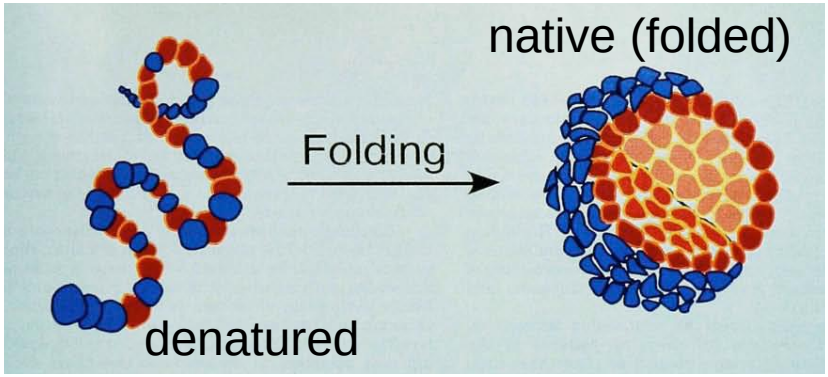
P. D. Thomas and K. A. Dill, *An iterative method for extracting energy-like quantities from protein structures*. PNAS **93**, 11628 – 11633 (1996)

Images from Purves et al., Life: The Science of Biology, 4th Edition, by Sinauer Associates (www.sinauer.com)

HP Model for Proteins

Key ideas:

- The native state of a globular protein is compact, but does not have the spatial order of a crystal. The denatured state has an open conformation.
- Amino acids may be classified as polar (P) and nonpolar (H).
- The folded state minimizes the interactions between hydrophobic (non-polar) residues and the polar solvent.

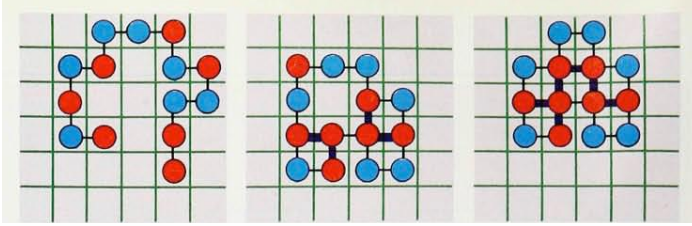


Blue: polar residues
Red: hydrophobic residues
Grey: aqueous (polar) solvent

HP model:

1. A protein is modeled as a chain of N amino acids.
2. Each amino acid is either polar (P) or nonpolar (H).
3. Each contact between non-bonded H residues contributes a (free) energy $\varepsilon < 0$; the energy of the system is $E = \varepsilon N_{\mathrm{HH}}$, where N_{HH} is the number of H – H contacts.

HP lattice model

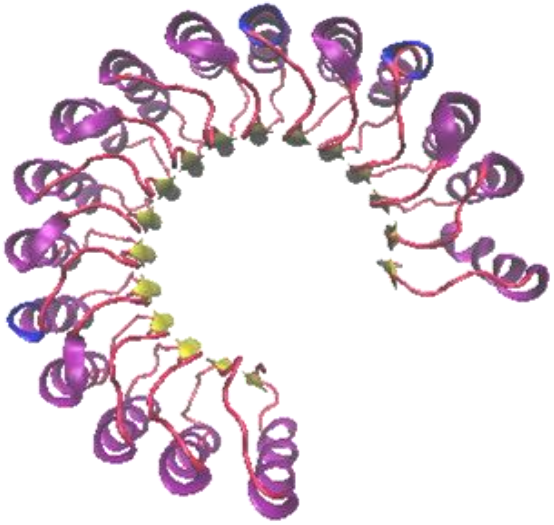


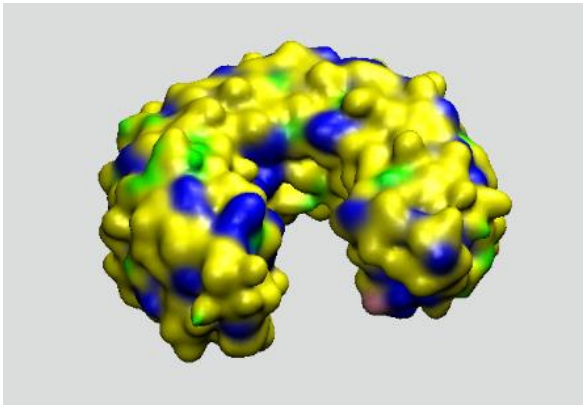
$N_{\mathrm{HH}} = 0$ $N_{\mathrm{HH}} = 4$ $N_{\mathrm{HH}} = 6$

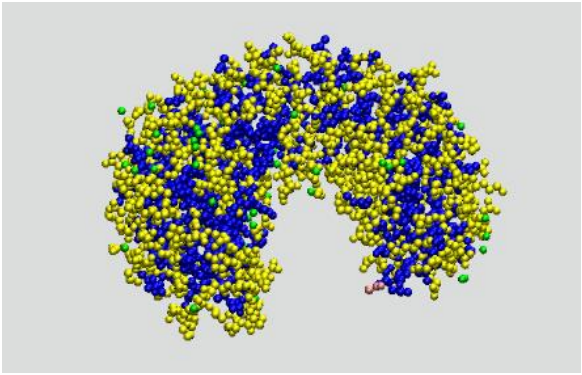
H. S. Chan and K. A. Dill, The protein folding problem, Physics Today, **46**, 24 - 32 (1993)
K. A. Dill, *Theory for the folding and stability of globular proteins*. Biochemistry **24**, 1501 – 1509 (1985).
K. F. Lau and K. A. Dill, *A lattice statistical mechanics model of the conformational and sequence spaces of proteins.* Macromolecules **22**, 3986 – 3997 (1989)

Example

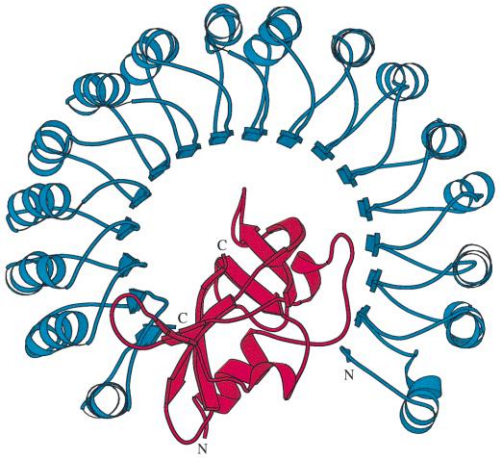
Robert Stewart







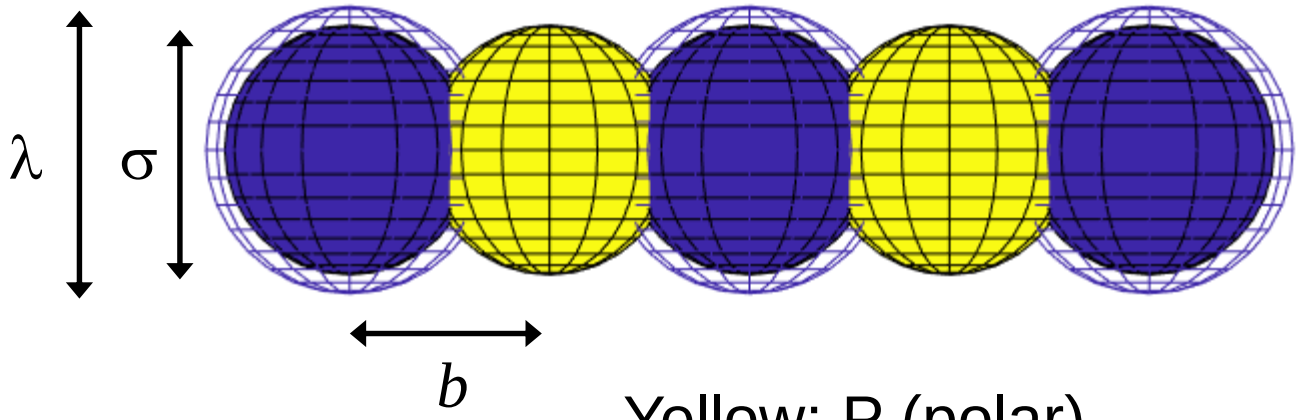
Porcine Ribonuclease Inhibitor (RI):
Image created by Robert Stewart from 2BNH. pbd file. Beta-sheet (yellow), alpha-helixes (purple), coils and turns (red).



Ribonuclease inhibitor (blue)
Ribonuclease (red)

Yellow: polar residues
Blue: hydrophobic residues

Square-well chain HP model



Yellow: P (polar)
Blue: H (hydrophobic)

hard core diameter: σ
well diameter: $\lambda = 1.15\,\sigma$
bond length: $b = 0.8\,\sigma$

H-H interactions:
square well potential

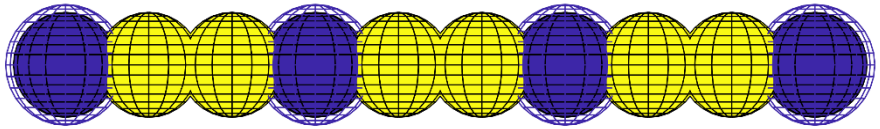
$$E_{ij} = \begin{cases} \infty \text{ for } r_{ij} < \sigma \\ \varepsilon \text{ for } \sigma \le r_{ij} < \lambda \\ 0 \text{ for } r_{ij} \ge \lambda \end{cases}$$

P-H and P-P interactions:
hard core potential

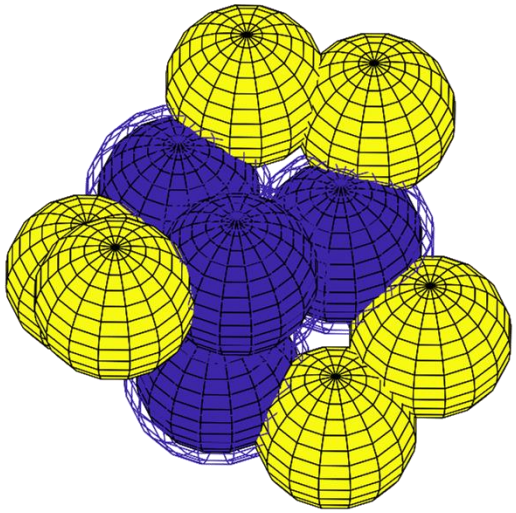
$$E_{ij} = \begin{cases} \infty \text{ for } r_{ij} < \sigma \\ 0 \text{ for } r_{ij} \ge \sigma \end{cases}$$

Examples

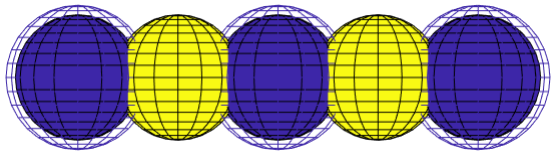
Sequence:
HPPHPPHPPH



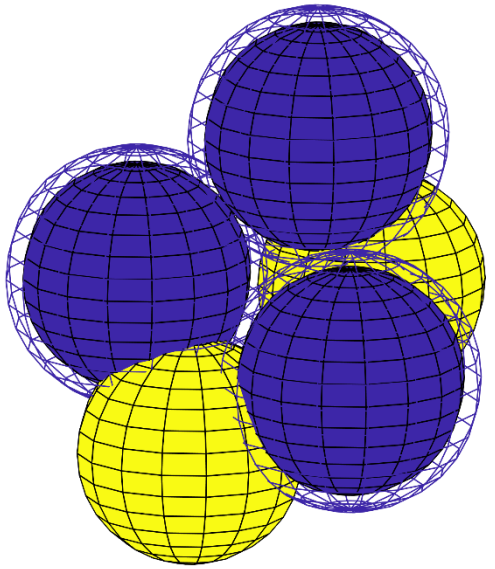
Ground state (folded)
energy 6ε, $\varepsilon < 0$



Sequence:
HPHPH

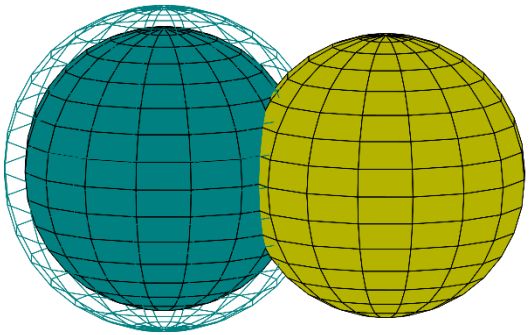


Ground state
energy 3ε, $\varepsilon < 0$

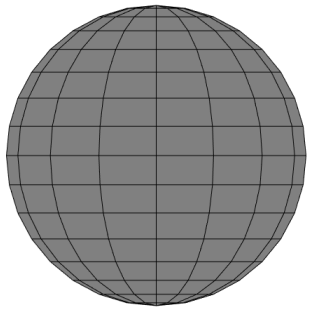


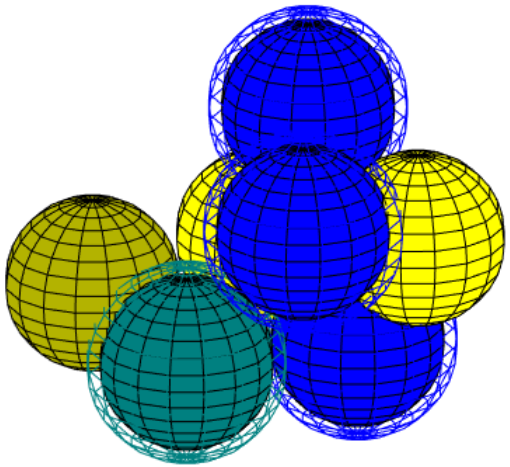
Ligand and solvent

Ligand Sequence: HP

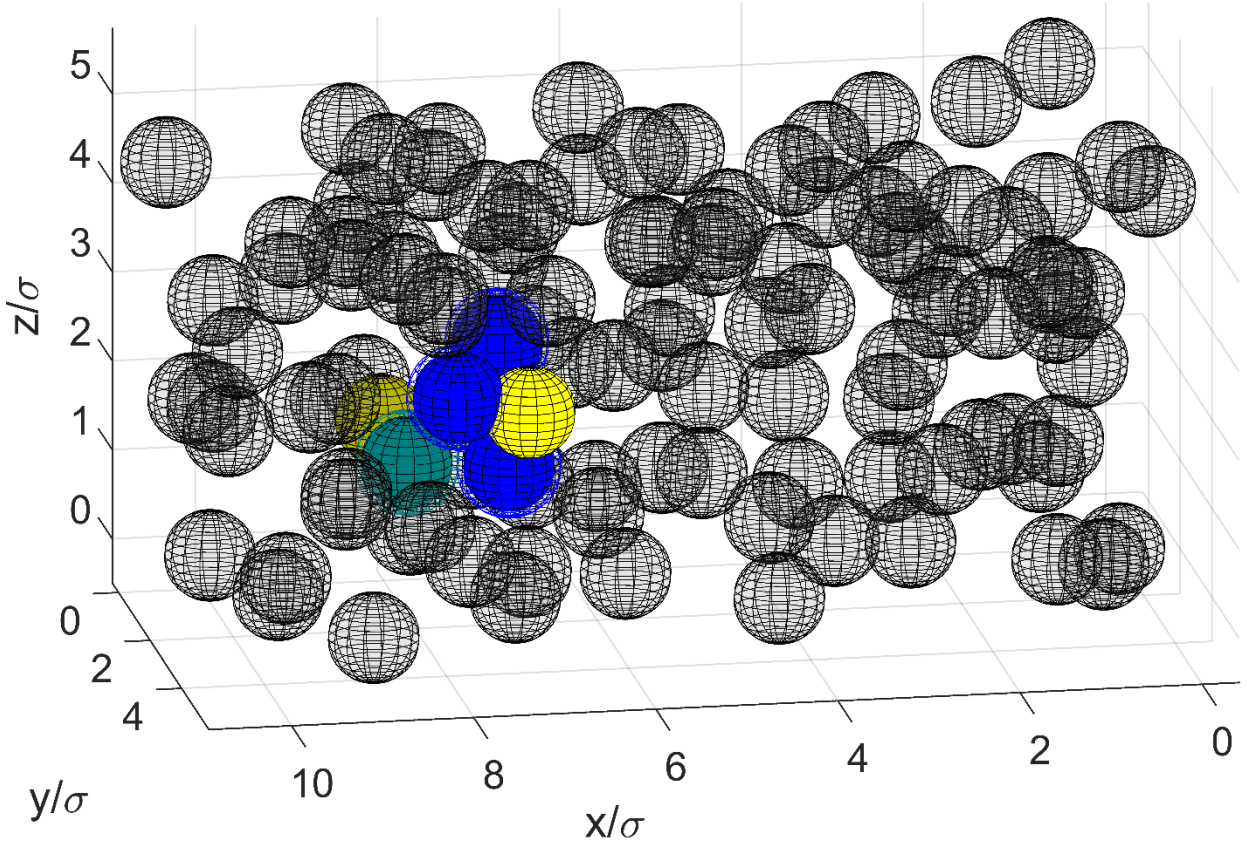


Solvent P Hard-sphere solvent



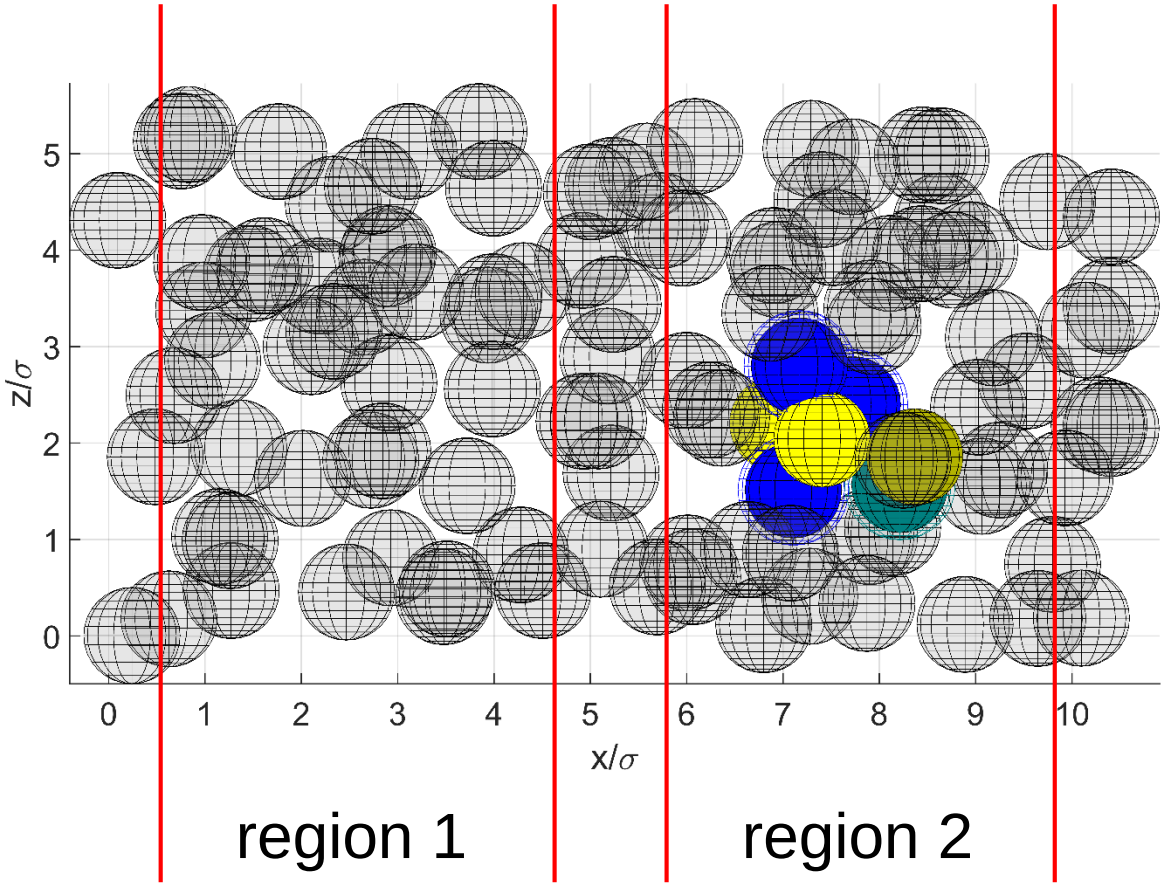


Periodic boundary conditions



Simulation Method:

- Monte Carlo simulations with Wang-Landau algorithm
- Goal: find the density of states $g \propto$ number of states for a given energy E and evaluate the entropy $S(E) = k_B \ln g(E)$ or evaluate the canonical partition function



$$Z(T) = \sum_E g(E) e^{-E/k_B T}$$

Here: divide the box into regions.

Simulation Method:

- Monte Carlo simulations with Wang-Landau algorithm
- Microstate: coordinates of all bead
- Macrostate: energy E_{1i} (side 1) and energy E_{2j} (side 2)
- Acceptance criterion:

$$p\left(\left(E_{1i}, E_{2j}\right) \rightarrow \left(E'_{1i}, E'_{2j}\right)\right) = \min\left(\frac{g\left(E_{1i}, E_{2j}\right)}{g\left(E'_{1i}, E'_{2j}\right)}, 1\right)$$

$g\left(E_i, V_j\right) =$ current estimate for the density of states

- After each elementary move, update the density of states by a refinement factor f and the histogram of visits by unity

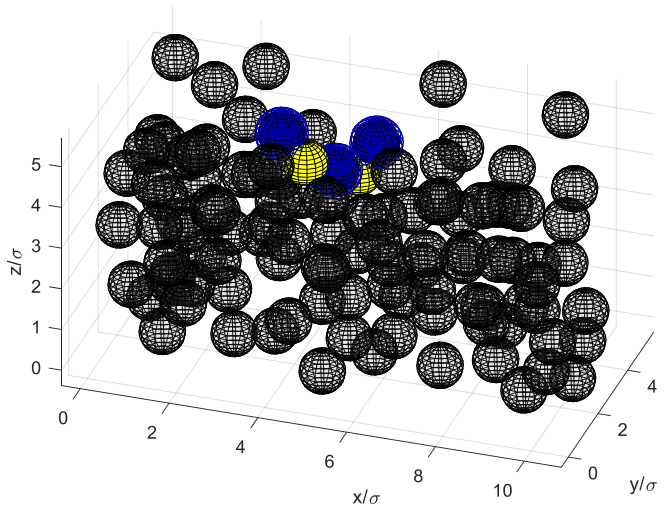
$$g\left(E'_{1i}, E'_{2j}\right) \rightarrow f \times g\left(E'_{1i}, E'_{2j}\right)$$

$$h\left(E'_{1i}, E'_{2j}\right) \rightarrow h\left(E'_{1i}, E'_{2j}\right) + 1$$

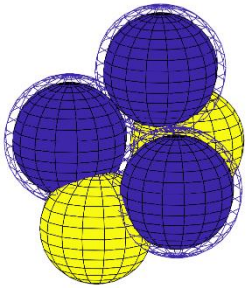
For accepted moves update g and h of the new state, otherwise update g and h of the original state.

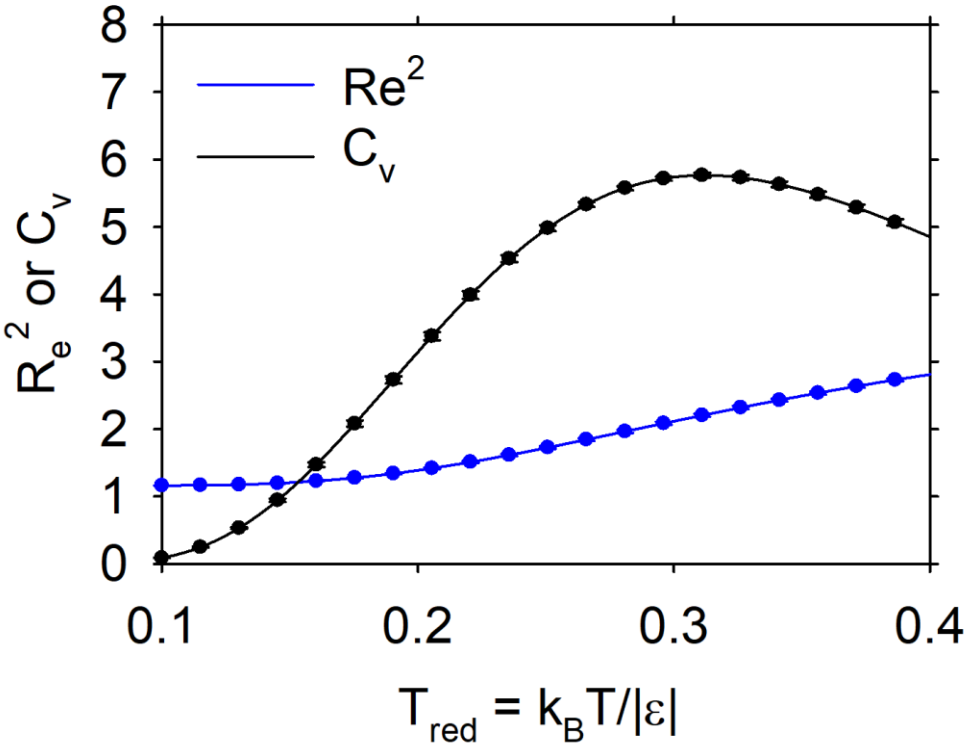
One (protein) chain in solvent

dos and chain dimensions evaluated for uniform temperature.

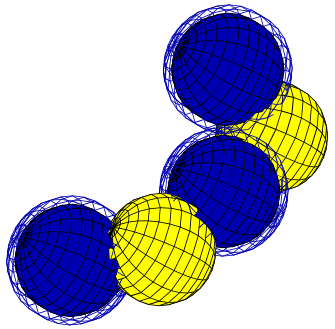


compact



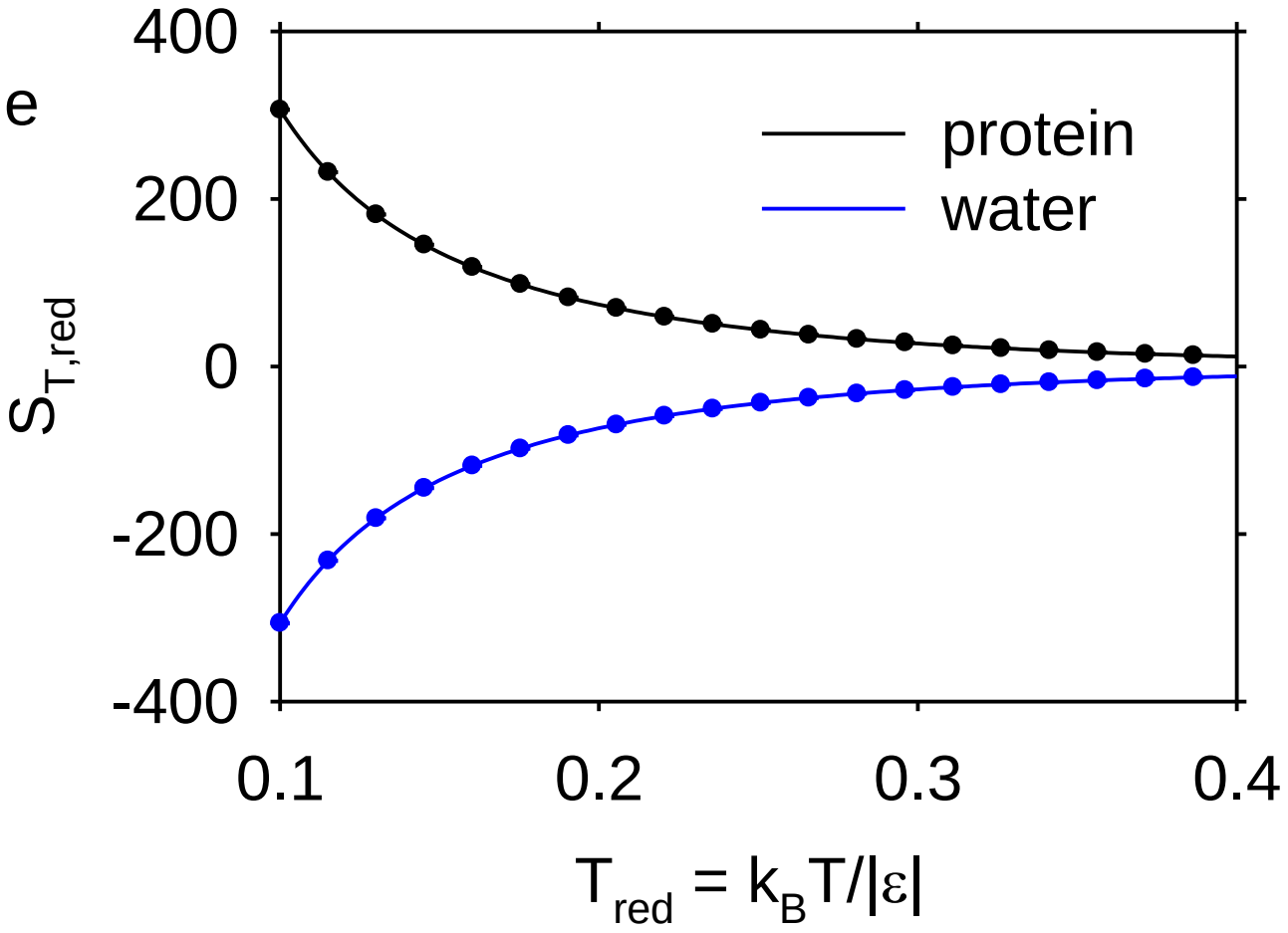


extended



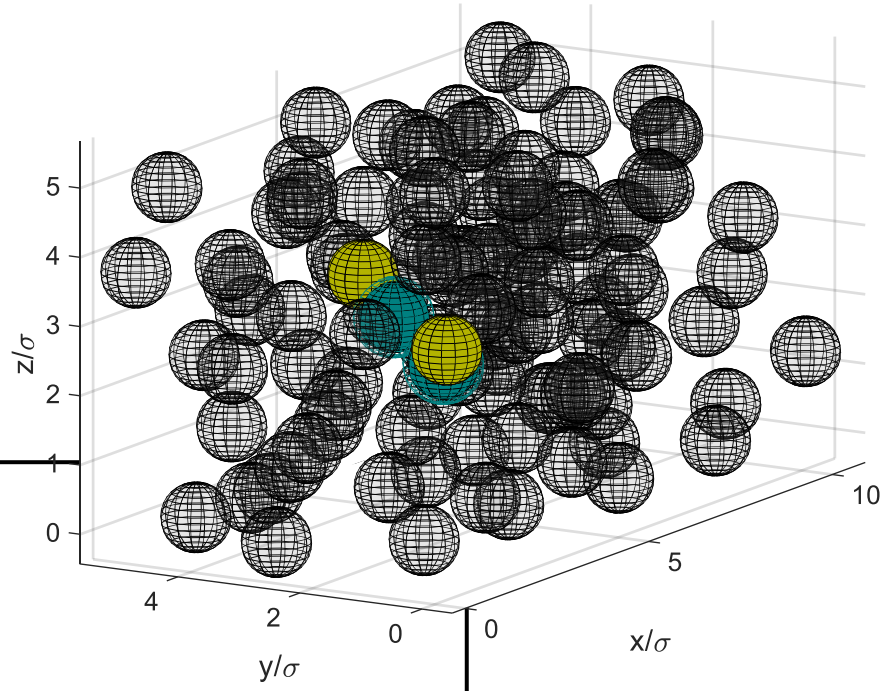
One chain in solvent, dos and particle numbers evaluated for a temperature difference of $\Delta T_{red} = 0.001$

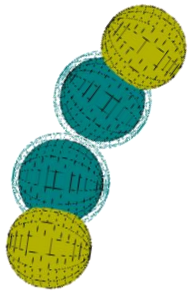
The protein is enriched on the cold side.



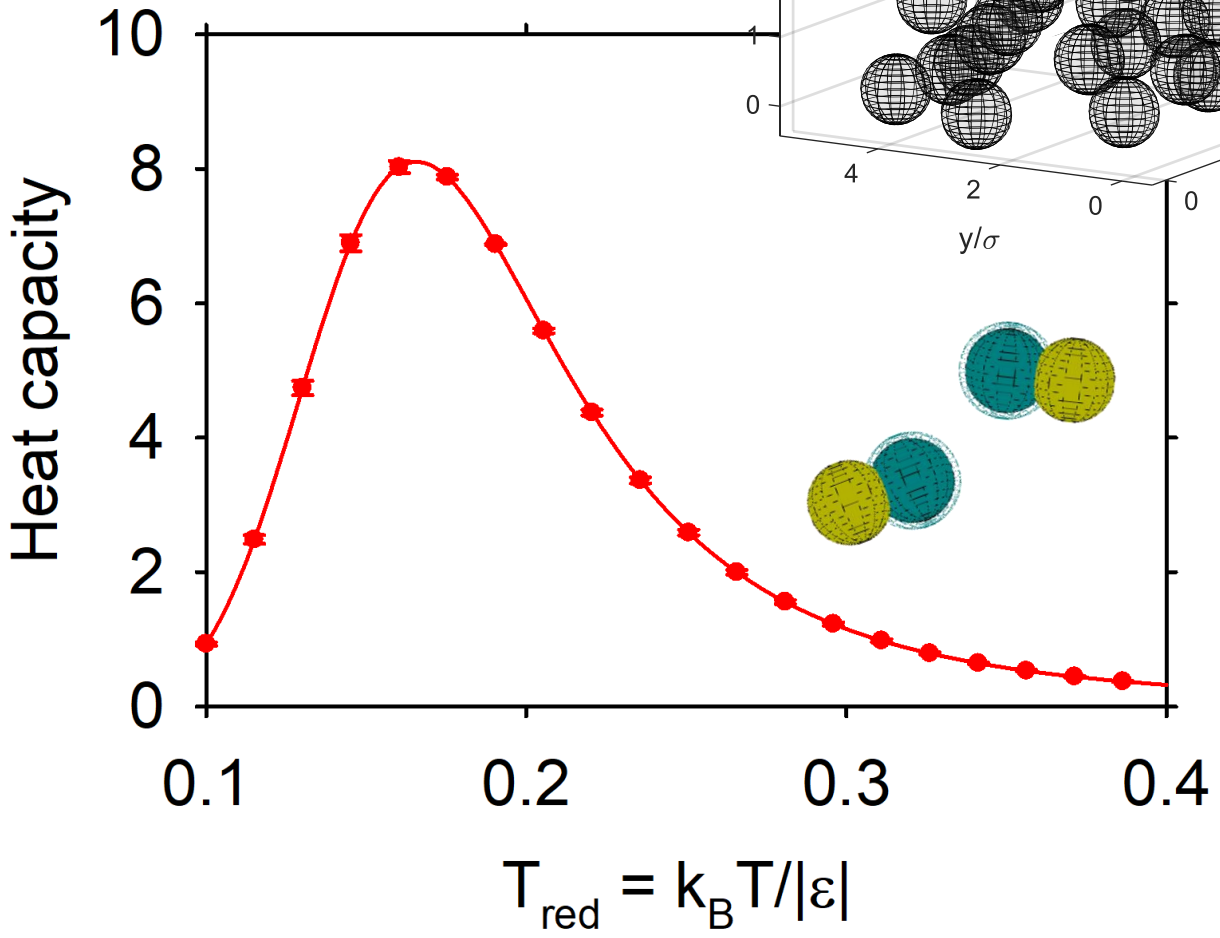
Two ligands in solvent

dos evaluated for uniform temperature



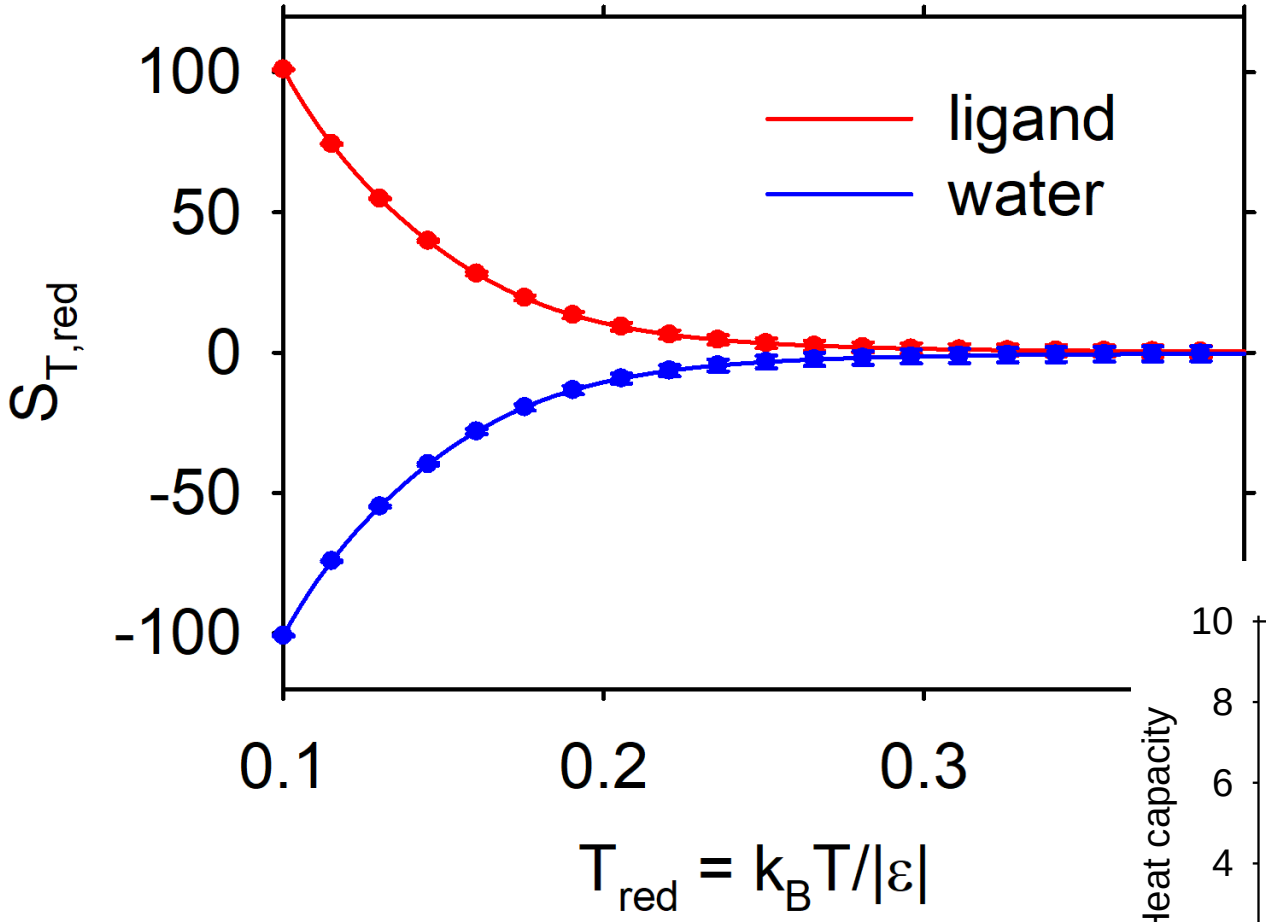


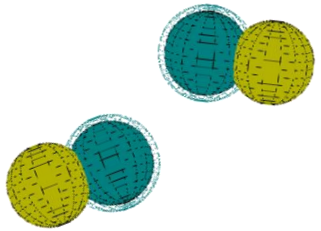
in contact



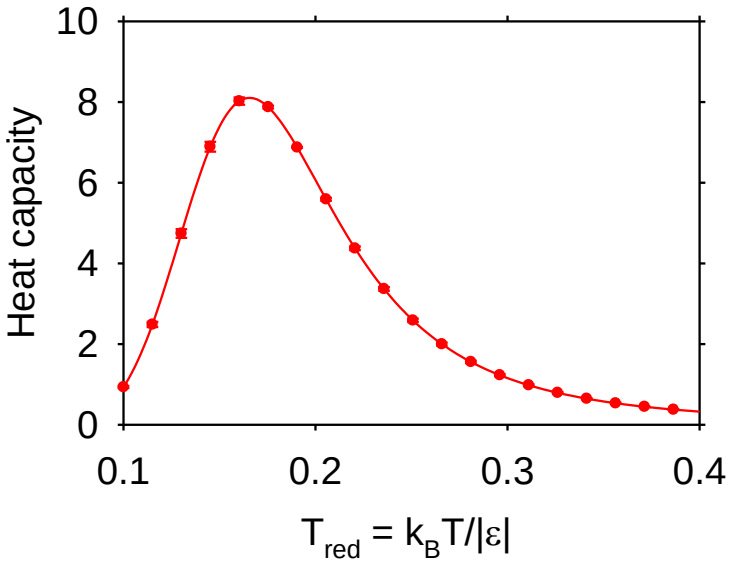
separate

Two ligands in solvent, dos and particle numbers evaluated for a temperature difference of $\Delta T_{red} = 0.001$





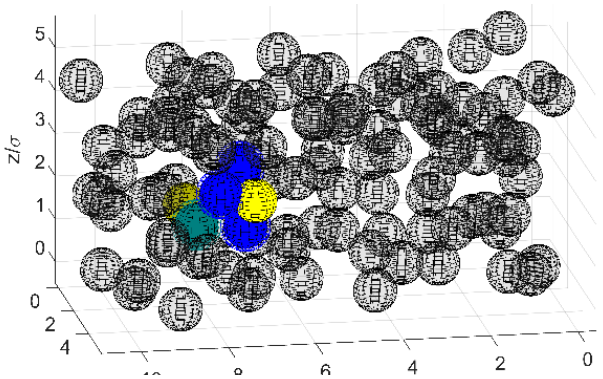
compare with

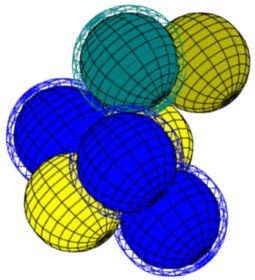


The ligands drift to the cold side once they start to aggregate.

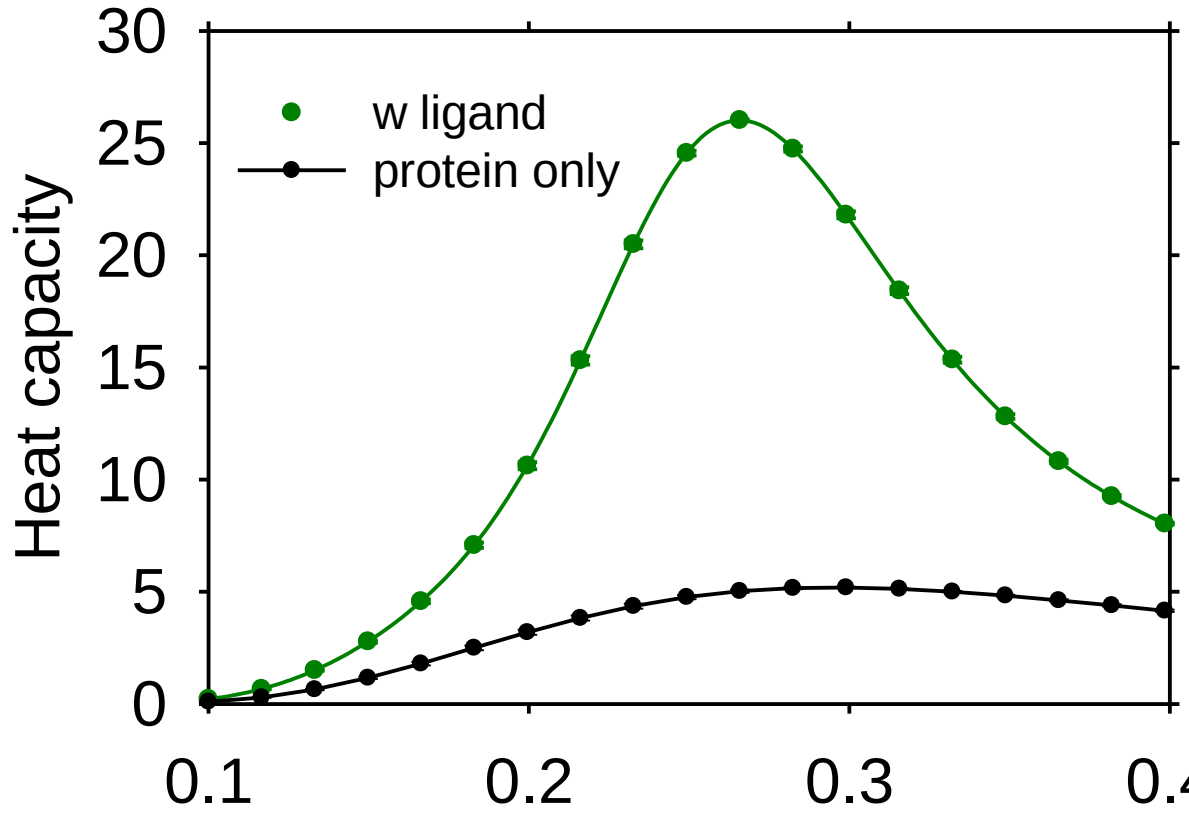
One (protein) chain and one ligand in solvent

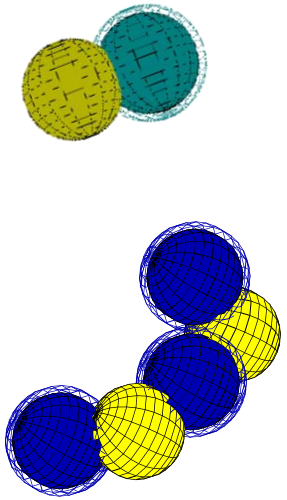
dos and chain dimensions evaluated for uniform temperature.



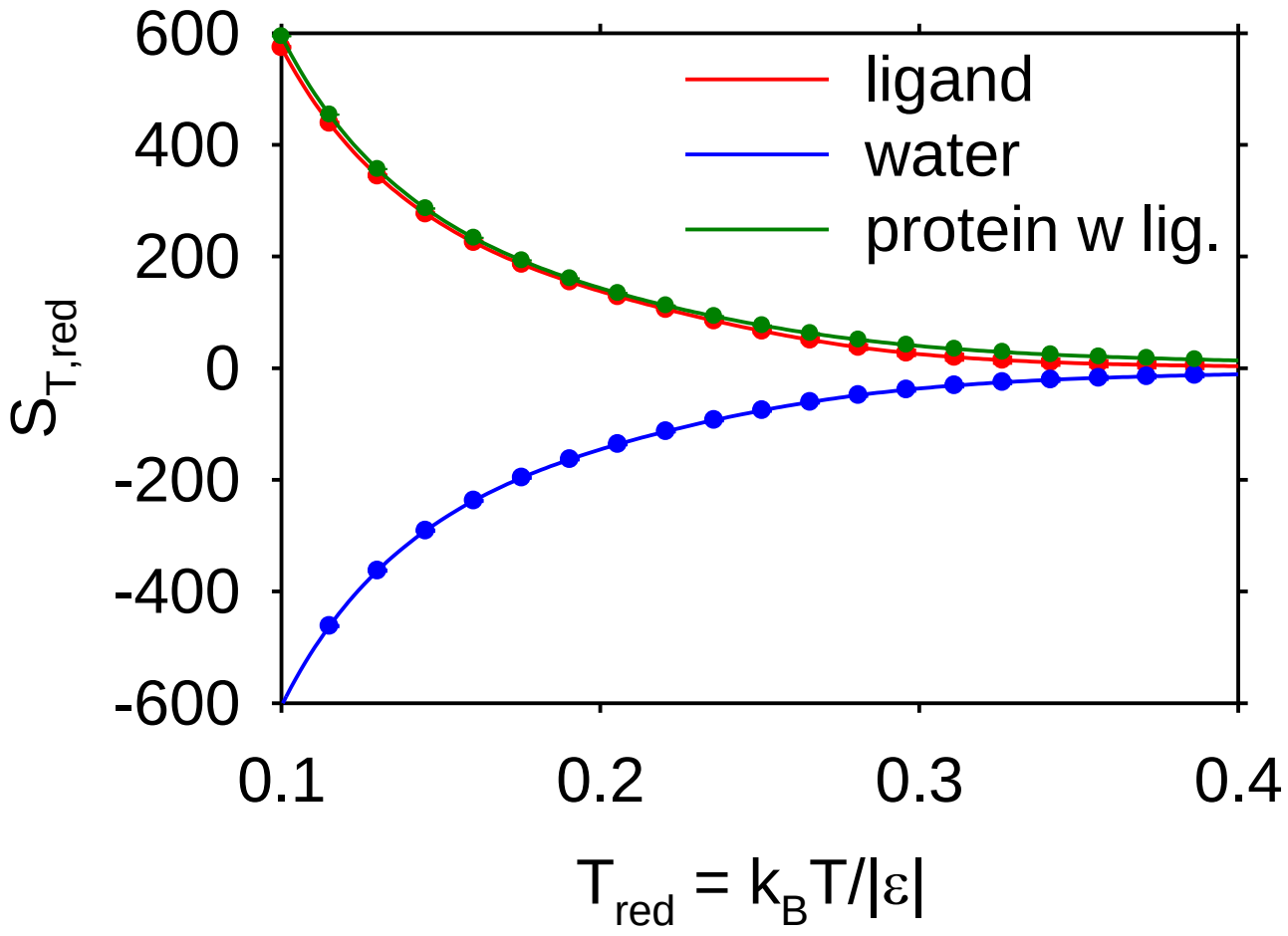


bound



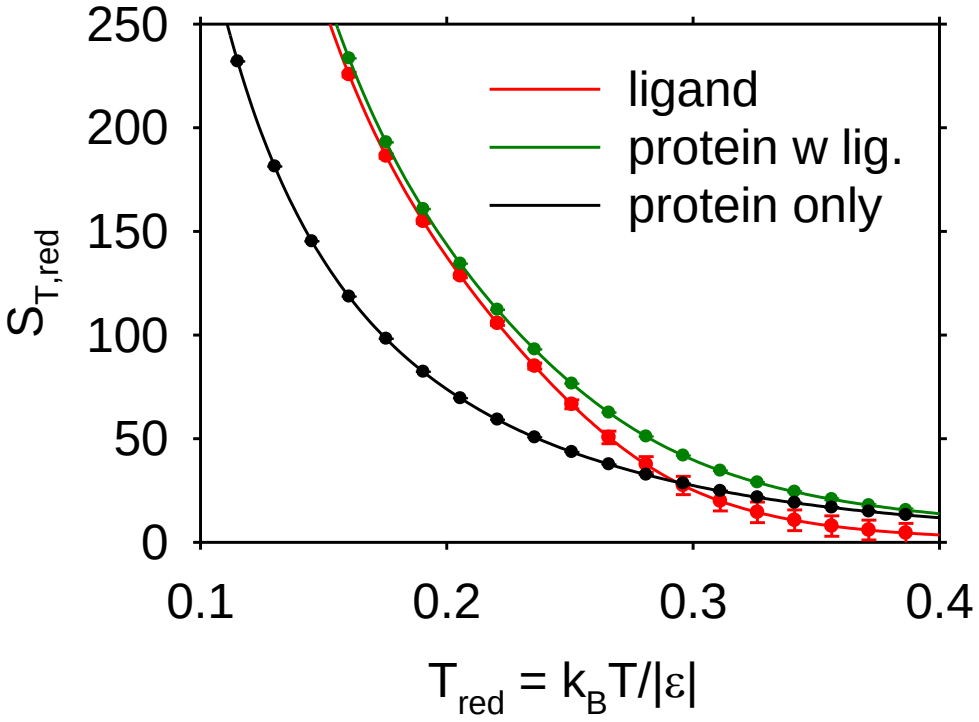


One (protein) chain, one ligand in solvent, dos evaluated with a temperature difference of $\Delta T_{red} = 0.001$

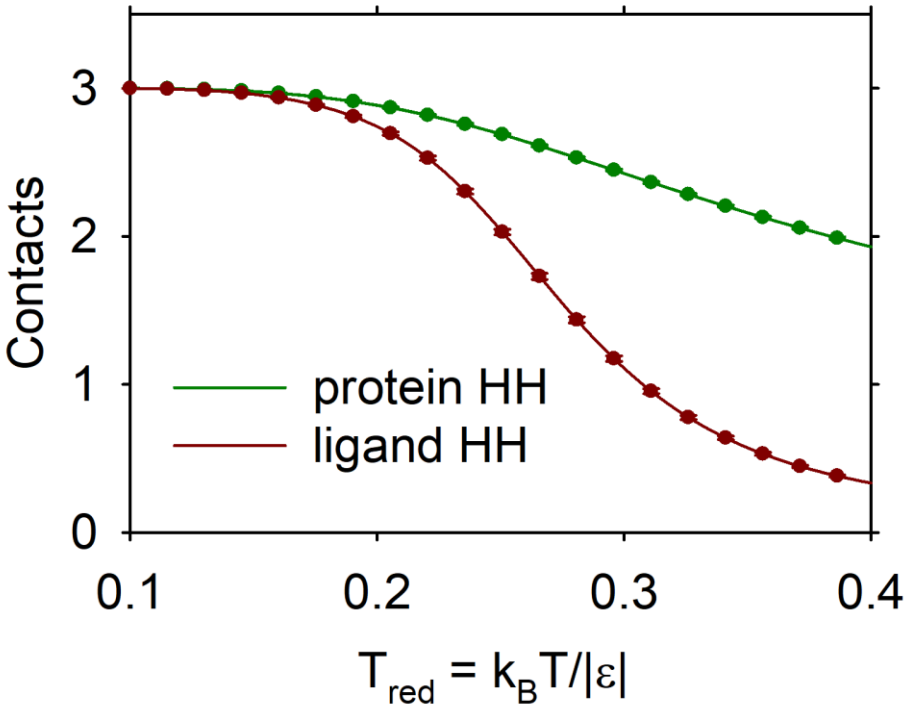


The Soret coefficients of the protein and the ligand are both positive (thermophobic) and very similar for most temperatures.

One (protein) chain, one ligand in solvent, dos evaluated with a temperature difference of $\Delta T_{red} = 0.001$



The Soret coefficient of the protein changes significantly once the ligand binds.



"MicrcscaleThermophoresis" allows quantitate ligand binding analysis.

M. Jerabek-Willemsen et *al*. *J. Molec. Struct.* **1077**, 101 (2014)

Discussion

Density-of-states simulation:

- provide access to thermodynamics over a wide temperature range from one simulation
- require simple (coarse-grained) systems
- if performed in a divided box, give an estimate of the chemical contribution to the Soret coefficient. The solvent needs to be treated explicitly.

Square-well HP model

- Simple, off-lattice with discrete energy level
- Reproduces some characteristics of proteins
- Hydrogen bond interactions can be modeled through orientation dependent interactions.
- More sophisticated SW-type models have been parametrized for proteins and can be simulated with molecular dynamics simulations, which gives access to dynamics.

Thanks!

Matt Murrow

Maxx Swoger

Mark Taylor
Hiram College

Simone Wiegand, FZ Juelich, and her group.