Enhancing the sampling

How to save time, and time is money

Marcus Elstner, Daniel Holub and Tomáš Kubař

2019, July 4

Energy barriers in simulations

Energy landscapes in large (bio)molecular systems
– multitude of almost iso-energetic minima,
separated from each other by energy barriers of various heights

Each of these minima ≡ one particular structure (conformation);
neighboring minima correspond to similar structures

Structural transitions are barrier crossings, and
the transition rate is determined by the height of the barrier.

Using quotations by Helmut Grubmüller

Arrhenius' rate

Normal MD – only 100ns-or-so time scales are accessible,
so only the smallest barriers are overcome in simulations,
and only small structural changes occur.

$$k \propto \exp\left[-E_{\mathrm{A}}/kT\right]$$

Any larger barriers are traversed more rarely
(although the transition process itself may well be fast),
and thus are not observed in MD simulations.

Special techniques will be required to solve this problem.

Arrhenius' rate

How often does something happen in a simulation?

$k = A \times \exp\left[-E_A/kT\right]$, e.g. $A = 1 \times 10^9$ s^{-1}

E_A kcal/mol	k 1/s	$1/k$ μs
1	0.19×10^9	0.005
3	6.7×10^6	0.15
5	0.24×10^6	4.2
7	8.6×10^3	120

If the process has to overcome a barrier of 5 kcal/mol,
we have to simulate for 4 μs to see it happen once on average.

Overview Methods

- Conformational flooding – CF
- Metadynamics – METAD
- Replica-exchange molecular dynamics – REMD

Conformational flooding

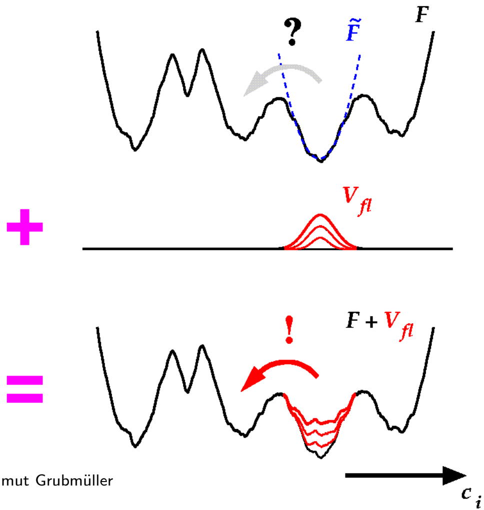
- to accelerate conformational transitions in MD simulations
  by several orders of magnitude
- makes it possible to simulate slow conformational transitions

1. generate a trajectory with a normal MD simulation
2. using the ensemble of structures from that trajectory,
   construct a localized artificial flooding potential V_{fl}:

- V_{fl} shall affect only the initial conformation
  and vanish everywhere outside of this region of conf. space
- V_{fl} shall be well-behaved (smooth)
  and 'flood' the entire initial potential-energy well

Grubmüller, Phys. Rev. E 1995

Conformational flooding



from the website of Helmut Grubmüller

Flooding potential

so, the simulation is performed with Hamiltonian

$$H = T + V + V_{\text{fl}}$$

a multivariate (n-dimensional) Gaussian function is good:

$$V_{\text{fl}} = E_{\text{fl}} \cdot \exp\left[-\frac{E_{\text{fl}}}{2k_{\text{B}}T} \cdot \sum_{i=1}^{n} q_i^2 \lambda_i\right]$$

E_{fl} – strength of the flooding potential (constant)
q_i – coordinates along the first n essential dynamics modes (PCA)

the first n PCA modes with eigenvalues λ_i will be flooded

The course of flooding simulation

The flooding potential is added to the force field,
and ‘flooding’ (biased) simulations are performed.

The energy minimum of the initial conformation is elevated
$\rightarrow$ the height of barriers is reduced
$\rightarrow$ the transitions are accelerated (TS theory)

Only the energy landscape within the minimum was modified $\rightarrow$

- the dynamics is already known there $\rightarrow$ uninteresting
- the barriers and all the other minima are unbiased
  – may be studied (are usually of interest)
- CF is expected to induce unbiased transitions
  – those that would occur without flooding (but slower)

Metadynamics

- a similar idea as flooding – discourage
  revisiting of states that have already been sampled
- ‘to reconstruct multidimensional ΔG of complex systems’
- artificial dynamics (metadynamics) performed
  in the space defined by a few collective variables S,
  assumed to give a coarse-grained description of the system
- history-dependent biasing potential constructed as a sum
  of Gaussians centered at points visited in the simulation

Laio & Parrinello, Proc. Natl. Acad. Sci. 2002

using quotations by Alessandro Laio

Metadynamics – how it works

- a new Gaussian is added at every time interval t_G
- the biasing potential at time t is given by

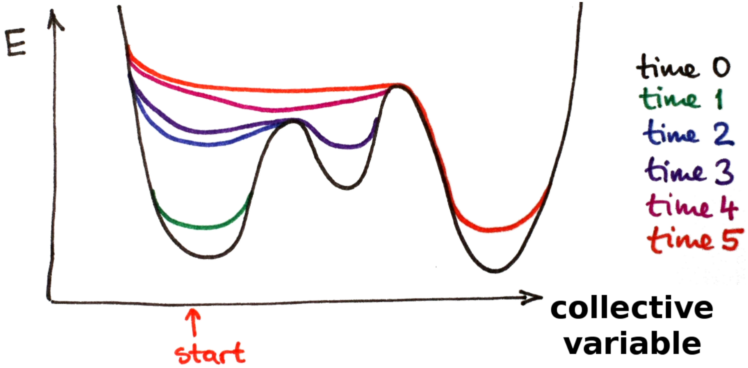
$$V_G(S(x), t) = \sum_{t'=t_G, 2t_G, 3t_G, \dots} w \cdot \exp\left[-\frac{(S(x) - s_{t'})^2}{2 \cdot \delta s^2}\right]$$

w and δs – height and width of the Gaussians (preset)
$s_t = S(x(t))$ – value of the collective variable at time t

- the simulation is performed with time-dependent Hamiltonian

$$H = T + V + V_G(S(x), t)$$

Metadynamics – what it looks like



https://www.youtube.com/watch?v=IzEBpQ0c8TA
https://www.youtube.com/watch?v=iu2GtQAyoj0

Metadynamics – how it works

- biasing potential is filling minima on the free energy surface
  that the system visits during the MD
- potential energy in metadynamics
  $\equiv$ genuine potential energy + sum of biasing Gaussians
  – is a function of collective variable(s) S
- resulting free energy in metadynamics
  – includes the sum of biasing potentials also
  – is becoming constant as simulation time is progressing
  – once this stage is reached:

  negative of sum of biasing Gaussians = unbiased free energy
- the MD has a kind of memory via the biasing potential

Properties of metadynamics

- explores new reaction pathways
- accelerate rare events
- estimates free energies efficiently
- the system escapes a local free energy minimum through the lowest free-energy saddle point.
- the free-energy profile is filled with the biasing Gaussians
- the sum of the Gaussians → (negative of) the free energy:

$$\lim_{t\to\infty} V_G(S,t) = -\Delta F(S) + const$$

(if the dynamics along the remaining degrees of freedom is much faster than the dynamics along S)

Collective variables

Crucial point: what collective variables S shall we consider?

Identify the variables that are of interest
and are difficult to sample because of barriers
that cannot be cleared in the available simulation time.

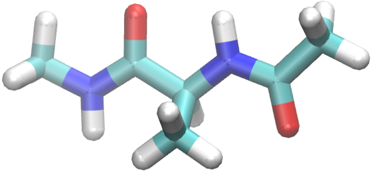
These variables $S(x)$ are functions of the coordinates of the system;
practical applications – up to 3 such variables,
and the choice depend on the process being studied.

Typical choices – principal modes of motion obtained with PCA.

Still, the choice of S may be far from trivial.

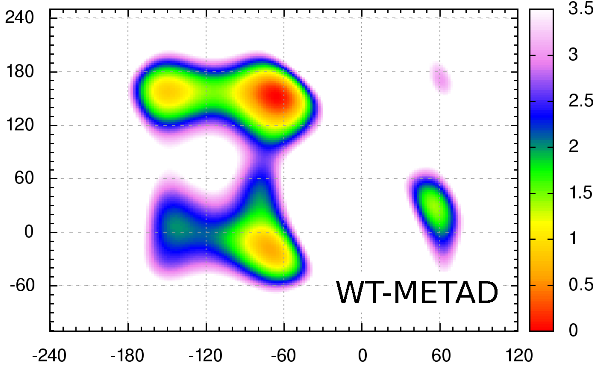
Metadynamics – example – alanine dipeptide

- 22 atoms, 1 pair of $\varphi - \psi$ angles



- one of the smallest molecules with peptide bonds
- sum of all biasing Gaussians during the simulation
  $\rightarrow$ estimate of free energy ΔG (in kcal/mol)
- whenever the current global minimum is populated further,
  the estimate of its ΔG decreases,
  i.e. ΔG everywhere else increases

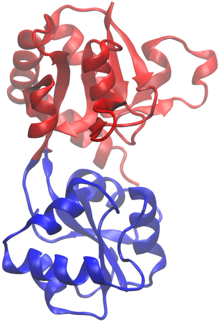
Metadynamics – example – alanine dipeptide

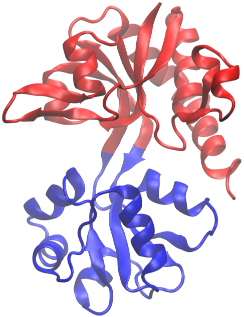


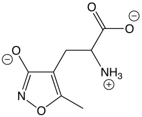
color coded: ΔG in kcal/mol

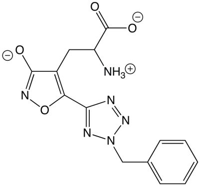
Metadynamics – example – glutamate receptor GluA2

- opening/closing of the ligand-binding domain (LBD)
- known ligand AMPA, novel ligand 2-BnTetAMPA (BTA)



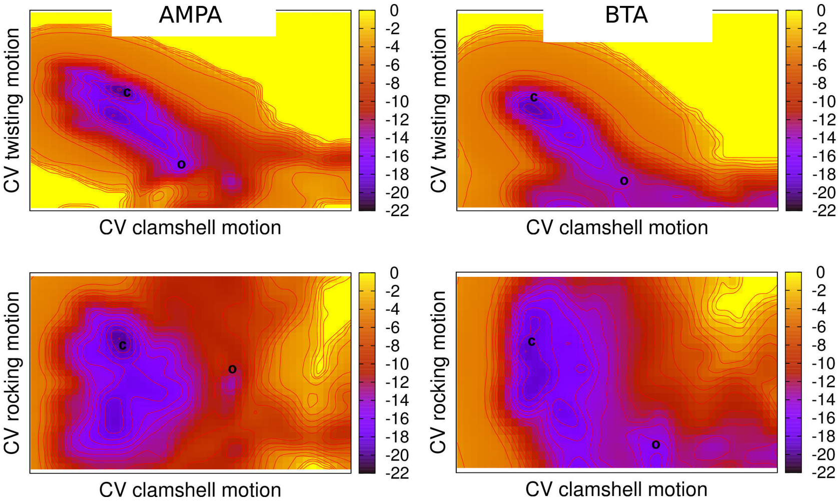






- collective variables: three dominant eigenvectors from PCA: clamshell motion, twisting motion and rocking motion
- 500 ns of metadynamics simulations of each complex
- two minima – open (O) and closed (C) state of LBD

Metadynamics – example – binding pocket of a protein



courtesy Tino Wolter

Replica-exchange molecular dynamics

REMD / parallel tempering

- method to accelerate the sampling of configuration space in case of high barriers between relevant configurations
- several (identical) replicas of the system are simulated simultaneously, at different temperatures
- coordinates+velocities of the replicas may be switched (exchanged) between two temperatures

Probability of replica exchange

- probability of exchange between $T_1 < T_2$
- determined in regular time intervals
- instantaneous potential energies U_1 and U_2
  in the two simulations needed

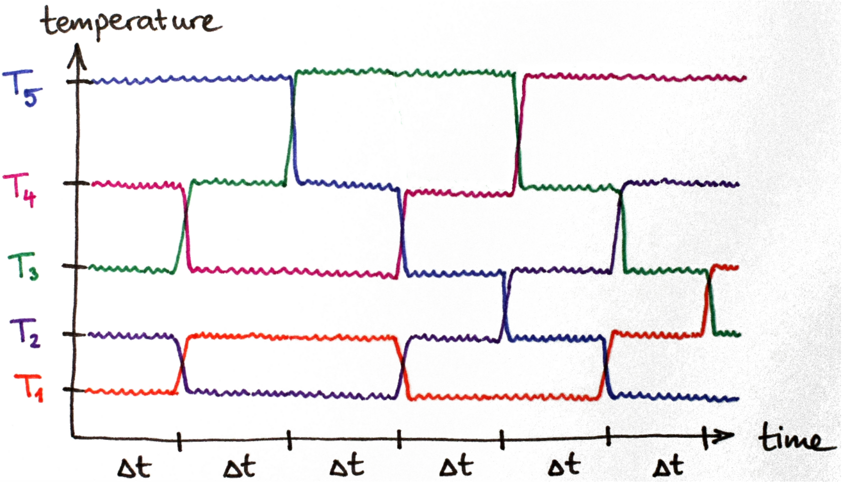
$$\mathcal{P}(1 \leftrightarrow 2) = \begin{cases} 1 & \text{if } U_2 < U_1, \\ \exp\left[\left(\frac{1}{k_B T_1} - \frac{1}{k_B T_2}\right) \cdot (U_1 - U_2)\right] & \text{otherwise.} \end{cases}$$

- if $\mathcal{P}(1 \leftrightarrow 2) >$ random number from $(0, 1)$,
  then replicas in simulations at T_1 and T_2 are exchanged
- a flavor of Metropolis' Monte Carlo

Setup of the simulation of replicas

- one replica at the temperature of interest ($T_1 = 300$ K)
- several others at higher temperatures ($T_1 < T_2 < T_3 < \ldots$)
- after 1 ps, attempt exchanges $1 \leftrightarrow 2$, $3 \leftrightarrow 4$ etc.
- after another 1 ps, do the same for $2 \leftrightarrow 3$, $4 \leftrightarrow 5$ etc.
- so, try to exchange replicas at "neighboring" temperatures

Setup of the simulation of replicas



Advantages of REMD

- due to the simulations at high temperatures:
- faster sampling and more frequent crossing of energy barriers
- correct sampling at all temperatures obtained,
  above all at the (lowest) temperature of interest
- increased computational cost (multiple simulations)
  pays off with largely accelerated sampling
- simulations running at different temperatures are independent
  except at attempted exchanges → easy parallelization
- first application – protein folding   Sugita & Okamoto, Chem. Phys. Lett. 1999

Choice of temperatures to simulate

Important – suitable choice of temperatures T_i – criteria:

- how frequent exchanges we wish (average prob. $\mathcal{P}(1 \leftrightarrow 2)$)
- the size of the system (degrees of freedom N_{dof})
- the number of temperatures/simulations

For protein/water systems with all bond lengths constrained:

- $N_{\text{dof}} \approx 2N$ (N – number of atoms)
- average probability is related to $T_2 - T_1 = \varepsilon T_1$ as

$$\overline{\mathcal{P}(1 \leftrightarrow 2)} \approx \exp\left[-2\varepsilon^2 N\right]$$

- set of temperatures may be designed to suit the problem

REMD generalized

- multiple different simulation parameters...
- different temperatures and different (e.g. biasing) potentials
- great flexibility

Simulations 1 and 2 performed

- at different temperatures T_1 and T_2
- with different potentials U_1 and U_2 (umbrella or other)

$$\Delta = \frac{1}{kT_1}\Big(U_1(q_2) - U_1(q_1)\Big) - \frac{1}{kT_2}\Big(U_2(q_1) - U_2(q_2)\Big)$$

$$\mathcal{P}(1 \leftrightarrow 2) = \begin{cases} 1 & \text{if } \Delta \leq 0, \\ \exp\left[-\Delta\right] & \text{otherwise.} \end{cases}$$

Extended sampling methods

Biasing potential methods – US, METAD

- required: a priori choice of reaction coordinate(s) to be biased
- problem – success depends on that choice, possibly non-trivial

REMD (parallel tempering)

- $+$ no such required, can be used rather blindly
- $-$ all of the system heated $\rightarrow$ may destroy something
- $-$ no knowledge of the system may be embedded
- $-$ poor efficiency for big systems: $\overline{\mathcal{P}(1 \leftrightarrow 2)} \approx \exp\left[-2\varepsilon^2 N\right]$
  $\rightarrow$ critical problem

Extended sampling methods

Hamiltonian replica exchange (HREX)

- in intermediate position between US/METAD and REMD/PT
- simpler to use than US/METAD
  - results depend not so strongly on the choices to be made
- efficiency does not depend on the overall system size
- many possibilities; our choice: REST2

REST1: Berne et al., Proc. Natl. Acad. Sci. USA 2005
modif: Ceulemans et al., J. Chem. Theory Comput. 2011
modif: Takada et al., J. Comput. Chem. 2011
REST2: Berne et al., J. Phys. Chem. B 2011
review and Gromacs implementation: Bussi, Mol. Phys. 2014

Replica-exchange with solute tempering

$$\mathcal{P} \propto \exp\left[-\frac{U}{kT}\right] = \exp\left[-\beta U\right]$$

- note: $\frac{1}{2}U$ would be the same as $2T$
- U is combined from terms that we can scale individually
  - is not possible for T
  - 'heating' of a portion of the system
  - a group of atoms, or just a group of interaction terms

REST2

- divide the system into two parts:
- hot – small, will be subject to extended sampling
- cold – all of the rest

Generate replicas with different $\lambda_m < 1$, modify parameters in hot:

- scale the charges by $\sqrt{\lambda_m}$
- scale the LJ depths ε by λ_m
- scale the amplitudes of dihedrals within hot by λ_m
- scale dihedrals partly within hot by $\sqrt{\lambda_m}$

Then, the 'effective' temperatures are

- inside hot: $T/\lambda_m > T$
- interactions between hot and cold: $T/\sqrt{\lambda_m}$
- inside cold: T is retained

REST2

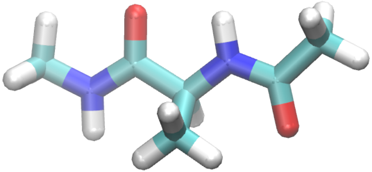
Meanings of temperature

- kinetic energy $\leftarrow$ velocities
  - does not change, is the same in hot and cold (300 K)
  - simulation settings need not be adjusted (time step!)
  - unlike in parallel tempering
- factor affecting the population of states
  - we play with this

REST2 – example

Solute tempering – dialanine

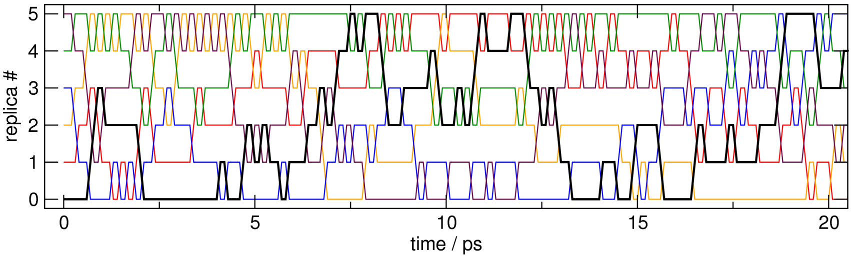
- alanine dipeptide – 22 atoms, 1 pair of $\varphi - \psi$
- force field: Amber99SB + TIP3P
- 5 replicas, $\lambda = 1 \ldots 0.18$ i.e. $T_m = 300 \ldots 1700$ K
- exchange every 0.1 ps, leading to $\overline{\mathcal{P}}$ =0.25–0.50



REST2 – example

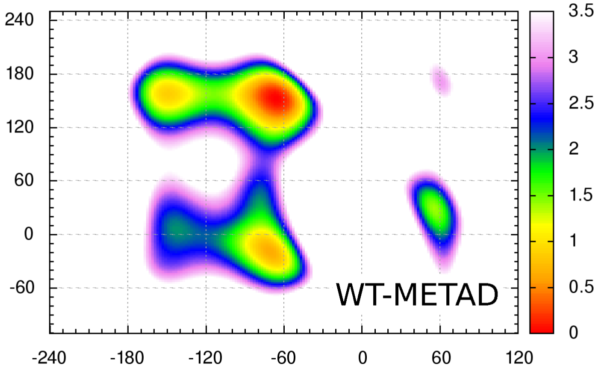
Solute tempering – dialanine

- alanine dipeptide – 22 atoms, 1 pair of $\varphi - \psi$
- force field: Amber99SB + TIP3P
- 5 replicas, $\lambda = 1 \ldots 0.18$ i.e. $T_m = 300 \ldots 1700$ K
- exchange every 0.1 ps, leading to $\overline{\mathcal{P}}$ =0.25–0.50



REST2 – example

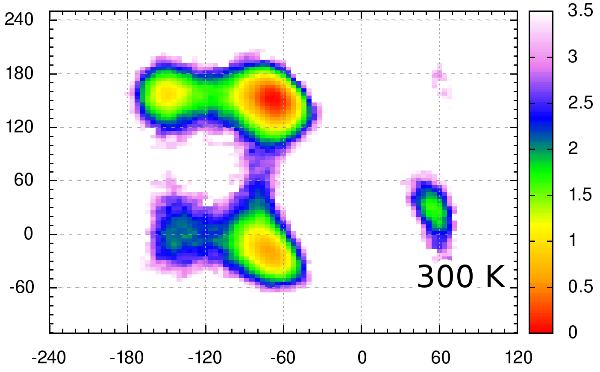
Solute tempering – dialanine – reference result from metadynamics



$\varphi - \psi$ in degrees, ΔF in kcal/mol

REST2 – example

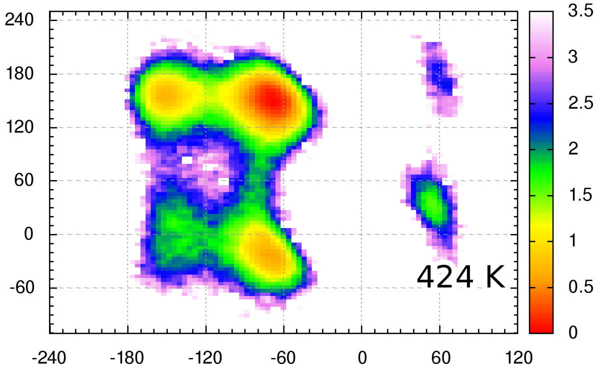
Solute tempering – dialanine – replica #0



$\varphi - \psi$ in degrees, ΔF in kcal/mol

REST2 – example

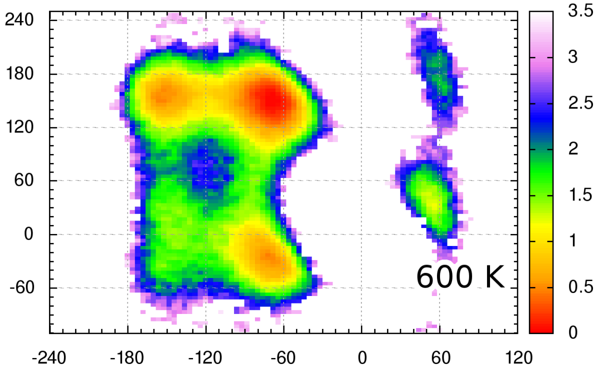
Solute tempering – dialanine – replica #1



$\varphi - \psi$ in degrees, ΔF in kcal/mol

REST2 – example

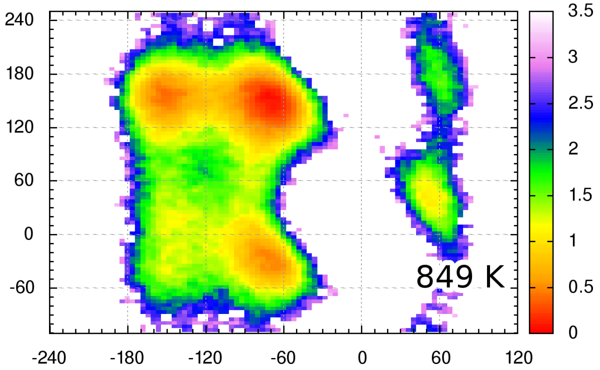
Solute tempering – dialanine – replica #2



$\varphi - \psi$ in degrees, ΔF in kcal/mol

REST2 – example

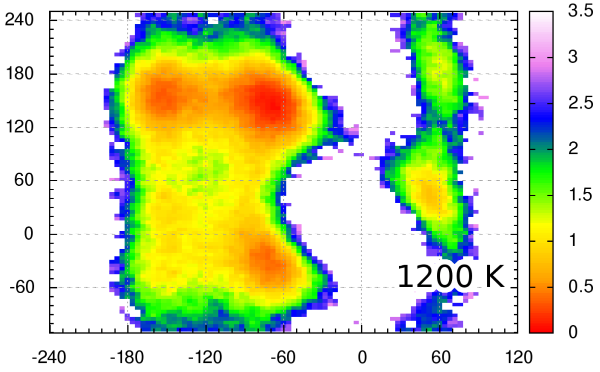
Solute tempering – dialanine – replica #3



$\varphi - \psi$ in degrees, ΔF in kcal/mol

REST2 – example

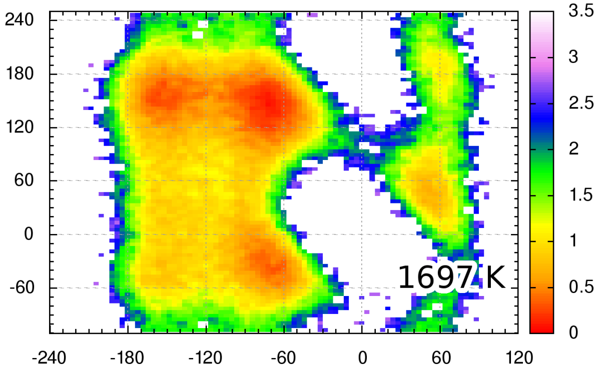
Solute tempering – dialanine – replica #4



$\varphi - \psi$ in degrees, ΔF in kcal/mol

REST2 – example

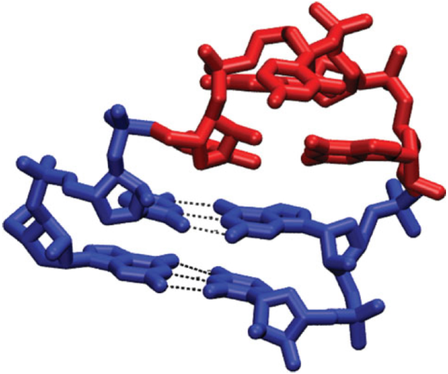
Solute tempering – dialanine – replica #5



$\varphi - \psi$ in degrees, ΔF in kcal/mol

REST2 – example

Partial tempering – RNA tetraloop



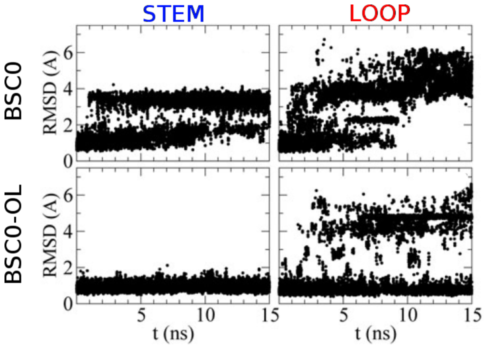
REST2 – example

Partial tempering – RNA tetraloop

- GC-UUCG-GC
- difficult – slow sampling, force field issues – Olomouc FF
- stem – WC HB restrained, kept 'cold'
- loop – 'hot', 16 replicas, $\lambda = 1 \ldots 0.3 \rightarrow \mathcal{P}$ =0.3–0.5
- 4600 TIP3P waters, 14 Na^{+}, 7 Cl^{-}

REST2 – example

Partial tempering – RNA tetraloop



deficiency of BSC0 manifests quickly: ladder-like structure of stem