



Chris Oostenbrink
 Institute of molecular modeling and simulation
 University of Natural Resources and Life Sciences



Institute of molecular
 modeling and simulation

Structure refinement using molecular dynamics simulations (NMR observables)



Schedule

Time	Mon. 20/11/17	Tue. 21/11/17	Wed. 22/11/17	Thur. 23/11/17	Fri. 24/11/17	Mon. 27/11/17	Tue. 28/11/17	Wed. 29/11/17	Thur. 30/11/17	Fri. 01/12/17
9:00	Free Time	Welcome <small>JAG</small>	Free Time	Free Time	Free Time	Free Time	Free Time	Free Time	Free Time	Free Time
9:30		Lecture 1: Overview and Introduction <small>JAG</small>	Lecture 4: Thermodynamics <small>JAG</small>	Lecture 7: Classical Mechanics I <small>JAG</small>	Lecture 10: Ensembles I <small>CO</small>	Lecture 13: Free energies: reaction coordinates <small>JAG</small>	Lecture 16: Calculating properties from simulations <small>JAG</small>	Lecture 19: Electrostatics <small>WvG</small>	Lecture 22: On the ethics of the academic endeavour: where do we go? <small>WvG</small>	Lecture 24: Left-overs/ questions and future perspectives <small>WvG</small>
10:15		Break <small>JAG</small>	Break <small>JAG</small>	Break <small>JAG</small>	Break <small>CO</small>	Break <small>JAG</small>	Break <small>JAG</small>	Break <small>WvG</small>	Break <small>WvG</small>	Break <small>WvG</small>
10:30		Lecture 2: Molecular Simulations MD/SD/MC <small>CO</small>	Lecture 5: Force-Field Development <small>CO</small>	Lecture 8: Classical Mechanics II <small>JAG</small>	Lecture 11: Ensembles II <small>JAG</small>	Lecture 14: Boundary Conditions I <small>WvG</small>	Lecture 17: Comparison with Experiments <small>WvG</small>	Lecture 20: Polarization <small>WvG</small>	Lecture 23: QM/MM <small>WvG</small>	Lecture 25: Students plans I <small>WvG</small>
11:15		Coffee Break <small>CO</small>	Coffee Break <small>CO</small>	Coffee Break <small>CO</small>	Coffee Break <small>CO</small>	Coffee Break <small>WvG</small>	Coffee Break <small>WvG</small>	Coffee Break <small>WvG</small>	Coffee Break <small>WvG</small>	Coffee Break <small>WvG</small>
11:45		Lecture 3: How to simulate using GROMOS <small>CO</small>	Lecture 6: Structure Refinement <small>CO</small>	Lecture 9: Analyzing with GROMOS <small>CO</small>	Lecture 12: Free energies: alchemy <small>CO</small>	Lecture 15: Boundary Conditions II <small>WvG</small>	Lecture 18: Searching & Enhanced Sampling <small>WvG</small>	Lecture 21: Multi -resolution simulations <small>WvG</small>	Lecture 23: QM/MM <small>WvG</small>	Lecture 26: Students plans II <small>WvG</small>
12:30		Break for lunch, self-study, discussion.* <small>CO</small>	Break for lunch, self-study, discussion.* <small>CO</small>	Break for lunch, self-study, discussion.* <small>CO</small>	Break for lunch, self-study, discussion.* <small>CO</small>	Break for lunch, self-study, discussion.* <small>WvG</small>	Break for lunch, self-study, discussion.* <small>WvG</small>	Break for lunch, self-study, discussion.* <small>WvG</small>	Break for lunch, self-study, discussion.* <small>WvG</small>	Tutorial 10: **
14:00	Registration	Tutorial 2: Running MD GROMOS Tutorial	Tutorial 3: Running MD Students Plans	Tutorial 4: Running MD Students Plans	Tutorial 5: Running MD Students Plans	Tutorial 6: Statistical Mechanics exercises	Tutorial 7: Analyzing MD: GROMOS Tutorial	Tutorial 8: Analyzing MD: Students Plans	Tutorial 9: Analyzing MD: Students Plans	How to prepare a barbecue
16:00	Coffee Break	Coffee Break	Coffee Break	Coffee Break	Coffee Break	Coffee Break	Coffee Break	Coffee Break	Coffee Break	
17:30	End of session	End of session	End of session	End of session	End of session	End of session	End of session	End of session	End of session	
20:30	Free time	Free time	Free time	Free time	Beer, Science & Friendship Good Stock Bar	Free time	Free time	Free time	Free time	Farewell

Observed vs derived quantities

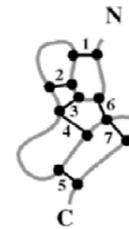
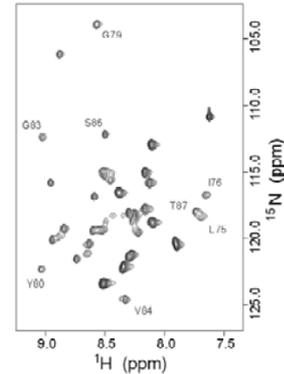
- Are we using **primary** or **secondary** experimental data?

- **primary**: “observable”, the quantity that is directly measured in the experiment

e.g. peak location and intensity in X-ray diffraction or NMR spectroscopy

- **secondary**: derived from primary data using some relationship that involves assumptions and approximations

e.g. structural parameters such as inter-atomic distances, dihedral angles



Don't compare to derived data, e.g. structures

W.F. van Gunsteren, J. Dolenc, & A.E. Mark, Curr. Opin. Struct. Biol., 18 (2008) 149-153

The Molecular Simulation Approach

- How to calculate a quantity or observable $Q(\mathbf{r})$?

Choose:

1. **(essential) degrees of freedom \mathbf{r}**
for $Q(\mathbf{r})$ *electronic*
atomic
solvent
2. **interaction function $V^{phys}(\mathbf{r})$**
between degrees of freedom (force field)
3. **equations of motion or sampling method**
generate a Boltzmann-weighted ensemble of conformers:
$$P(\mathbf{r}) = e^{-V^{phys}(\mathbf{r})/k_B T} / \int e^{-V^{phys}(\mathbf{r})/k_B T} d\mathbf{r}$$
4. **function $Q(\mathbf{r})$**
contains approximations and assumptions

Ensemble averages

$$\langle Q \rangle_r \equiv \int Q(\mathbf{r}) P(\mathbf{r}) d\mathbf{r}$$

are to be compared:

$$\langle Q \rangle_{sim}$$

should be compared to

$$\langle Q \rangle_{exp} \equiv Q^{exp}$$

If

- $V^{phys}(\mathbf{r})$ and $Q(\mathbf{r})$ are correct
- Sampling is infinite

problem solved, otherwise:

Make other choices
and try again

The structure (ensemble) determination problem

- Measurement of $Q(\mathbf{r})$ which depends on molecular conformation \mathbf{r}

$$Q^{\text{exp}} = \left\langle Q(\mathbf{r}) \right\rangle_{\substack{\text{molecules} \\ \text{time}}}$$

- Can we derive $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ from Q^{exp} ?

- Generally: insufficient Q^{exp} values to determine \mathbf{r}
accuracy and consistency of Q^{exp} ?

- How to deal with averaging $\langle \dots \rangle$?
Statistical mechanics --> Boltzmann weighting of conformers
Inversion of averaging is impossible

- Do we know the dependence of Q on \mathbf{r} , i.e. $Q(\mathbf{r})$?
How accurate is it? Can we invert it into $\mathbf{r}(Q)$?
 $\mathbf{r}(Q)$ may be multi-valued

- How do we bias the sampling to obtain \mathbf{r} such that $\langle Q(\mathbf{r}) \rangle_{\text{sim}} = Q^{\text{exp}}$?

Simulate with restraints

- Bias the coordinates \mathbf{r} to make $\langle Q(\mathbf{r}) \rangle_{\text{sim}}$ approach Q^{exp}
- Add an extra potential energy term to the force field

$$V^{\text{tot}}(\mathbf{r}) = V^{\text{phys}}(\mathbf{r}) + V^{\text{restr}}(\mathbf{r})$$

- Functional form**

- Full or half harmonic at short range

$$V_{\text{harm}}^{\text{restr}}(\mathbf{r}) = \frac{1}{2} K^{Qr} \left[\langle Q(\mathbf{r}) \rangle - Q^{\text{exp}} \right]^2 \quad \text{for } \langle Q(\mathbf{r}) \rangle < Q^{\text{exp}} + \Delta Q$$

- Bounded gradient (force) at long range

$$V_{\text{lin}}^{\text{restr}}(\mathbf{r}) = K^{Qr} \left[\langle Q(\mathbf{r}) \rangle - Q^{\text{exp}} - \frac{1}{2} \Delta Q \right] \Delta Q \quad \text{for } \langle Q(\mathbf{r}) \rangle \geq Q^{\text{exp}} + \Delta Q$$

- Continuous, continuous derivative

- Include**

- averaging over time (instantaneous vs. time-averaged)
- averaging over molecules if Q depends on multiple molecules
- enhanced sampling techniques if $r(Q)$ is multi-valued

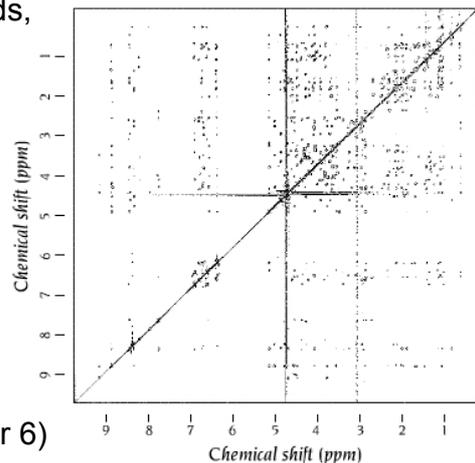
The structure (ensemble) determination problem

Examples of (observable) quantities $Q(\mathbf{r})$:

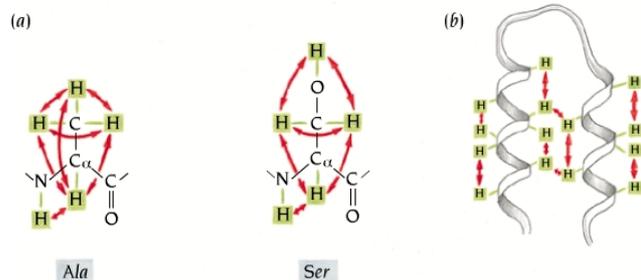
- NOE intensities or distances $\langle I_{i,j} \rangle \sim \langle r_{ij}^{-p} \rangle$ with $p = 3$ or 6
- 3J -coupling constants $\langle ^3J_{i,j} \rangle = a \langle \cos^2 \theta_{i,j} \rangle + b \langle \cos \theta_{i,j} \rangle + c$
- Residual dipolar couplings $\langle D_{i,j} \rangle = a \langle \cos^2 \theta_{ij} \rangle + b$
- Chemical shifts $\langle \sigma \rangle$
- Structure factors (amplitudes) $\langle F_{hkl} \rangle$
- FRET Efficiencies $\langle E(R, \theta) \rangle$
- CD spectra $\langle I(\lambda) \rangle$

NOE analysis

- Nuclei influence each other through bonds, or through space
- **COSY:**
Correlated spectroscopy
- **TOCSY:**
Total correlation spectroscopy
 - crosspeaks through bonds
- **NOESY:**
Nuclear Overhauser Effect spectroscopy
 - Crosspeaks through space
 - Intensity proportional to $\langle r_{ij}^{-p} \rangle$ ($p=3$ or 6)
 - Generally: upper bound to the distance



© 1999 GARLAND PUBLISHING INC.
A member of the Taylor & Francis Group



Atom-atom distance restraining

- Distance restraint:
 d_0 involving atoms i and j , derived from experiment

$$d(t) \text{ actual distance at time } t \quad d(t) = \sqrt{[\mathbf{r}_j(t) - \mathbf{r}_i(t)]^2}$$

- Instantaneous distance restraint:**

$$V^{dr} = 0 \quad \text{if } d(t) < d_0$$

$$V^{dr} = \frac{1}{2} K^{dr} [d(t) - d_0]^2 \quad \text{if } d(t) > d_0$$

$$f_{ix} = 0 \quad \text{if } d(t) < d_0$$

$$f_{ic} = -K^{dr} [d(t) - d_0] \frac{x_{ij}(t)}{r_{ij}(t)} \quad \text{if } d(t) > d_0$$

Atom-atom distance restraining

- Include time averaging

$$\bar{d}(t) = \left[\frac{1}{t} \int_0^t d(t')^{-3} dt' \right]^{-1/3}$$

$$\bar{d}(t) = \left[\frac{1}{N_t} \sum_{n=1}^{N_t} d(t_n)^{-3} \right]^{-1/3}$$

Use 1/3 averaging
because $\langle |NOE| \rangle$ is
proportional to $\langle d^{-3} \rangle$

- Time-average restraint:**

$$V^{dr} = 0 \quad \text{if } \bar{d}(t) < d_0$$

$$V^{dr} = \frac{1}{2} K^{dr} [\bar{d}(t) - d_0]^2 \quad \text{if } \bar{d}(t) > d_0$$

$$f_{ix} = 0 \quad \text{if } \bar{d}(t) < d_0$$

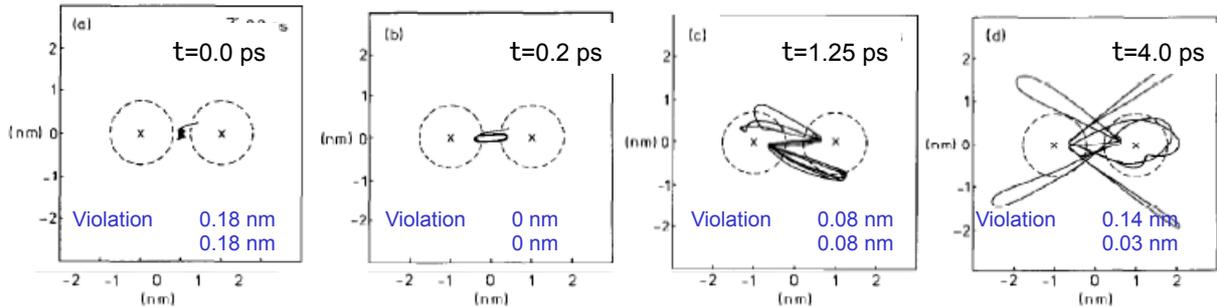
$$f_{ic} = -K^{dr} [\bar{d}(t) - d_0] \frac{1}{N_t} \left[-\frac{1}{3} \bar{d}(t)^4 \right] \left[-3d(t)^{-4} \right] \frac{x_{ij}(t)}{r_{ij}(t)} \quad \text{if } \bar{d}(t) > d_0$$

- Force becomes smaller with growing N_t , so damp the memory
- Time-averaging with a memory relaxation time τ :

$$\bar{d}(t) = \left[\frac{1}{t} \frac{1}{1 - e^{-t/\tau}} \int_0^t e^{-(t-t')/\tau} d(t-t')^{-3} dt' \right]^{-1/3}$$

Time-averaging: example

- 3 particles: 2 particles (x) fixed at 2 nm distance from each other
1 particle freely moving with 2 distance restraints:
both 0.8 nm length to the fixed particles



- Refinement applied with

$$\bar{d}(t) = \left[\frac{1}{t} \frac{1}{1 - e^{-t/\tau}} \int_0^t e^{-(t-t')/\tau} d(t-t')^{-3} dt' \right]^{-1/3}$$

- Average violations calculated from

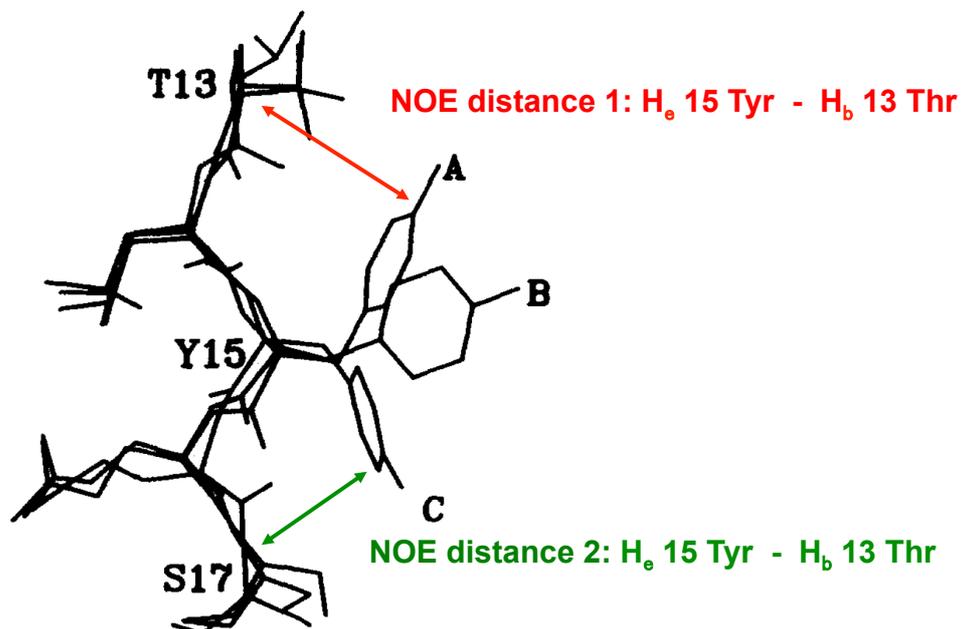
$$\bar{d}(t) = \left[\frac{1}{t} \int_0^t d(t-t')^{-3} dt' \right]^{-1/3}$$



A.E. Torda et al., Chem. Phys. Lett, **157** (1989) 289-294

Example Tendamistat

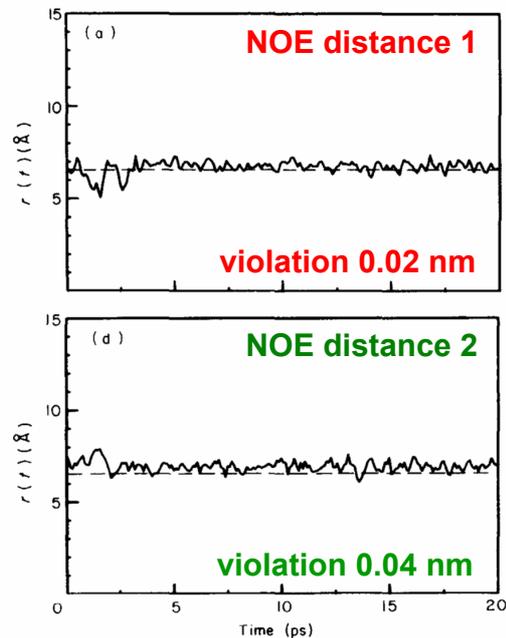
- small, 74 residue protein, 842 NOE distances from NMR
- conflicting NOE distances from the experiment
 - no single structure found that had no violations





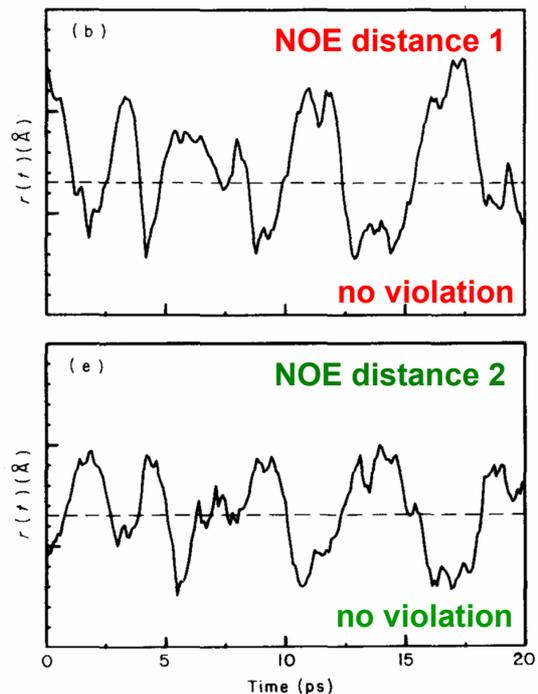
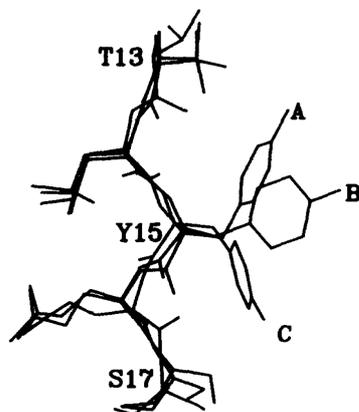
Instantaneous restraints

- Applying *instantaneous* distance restraints
- puts an extra force on the atoms which pulls them to the experimental NOE distance
- Atoms do get close, but quite a lot of strain is present in the system
- **Small distance fluctuations**
- **Yet, *small* bound violations remain present**



Time-averaged restraints

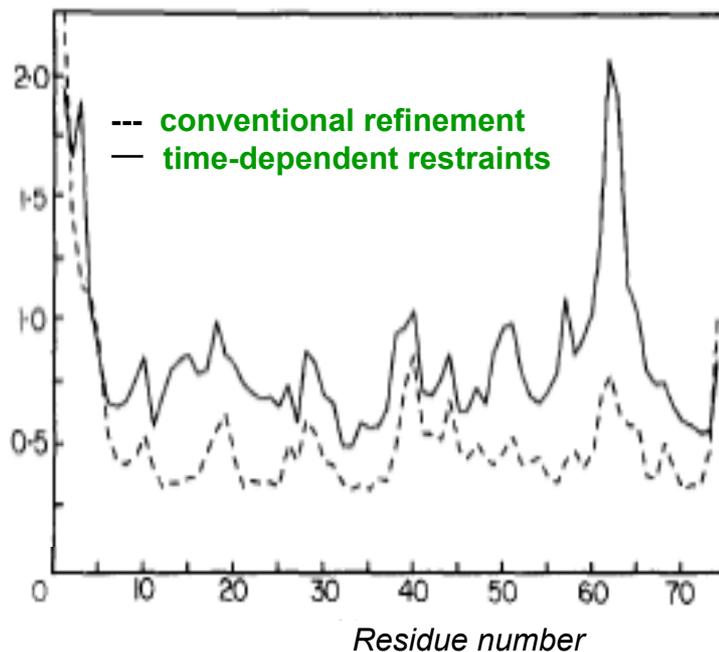
- *Time-averaged* distance restraints
- Extra forces on the atoms to enforce that the NOE distance is fulfilled on average
- Tyr13 is flipping back and forth



- **Large distance fluctuations**
- **Yet, *no* bound violations are present**

Fluctuations: effect of restraints

Root-mean-square atom-positional fluctuation(\AA) of C_{α} atoms



Conclusion:

conventional refinement restricts atomic motion too much (instantaneous restraints)

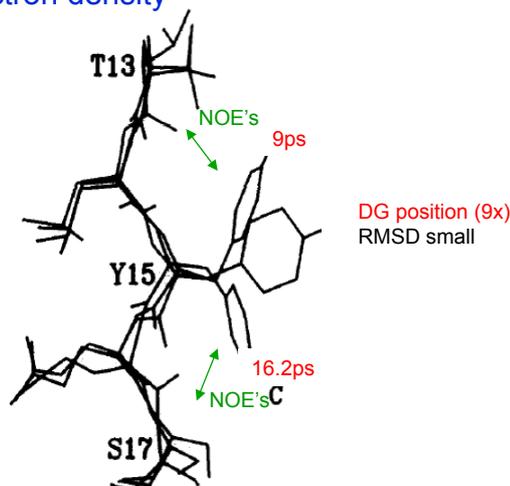


Comparison to other structures

DG: one position (violations) in all 9 DG (distance geometry) structures

MD: many positions (no violations)

X-ray: no electron density



J. Mol. Biol. **214** (1990) 223-235

Conclusion:

- Convergence to **one** structure does **not** indicate that only one structure fits the experimental data!
- The experimental data are compatible with more mobility than is suggested by **static** modeling



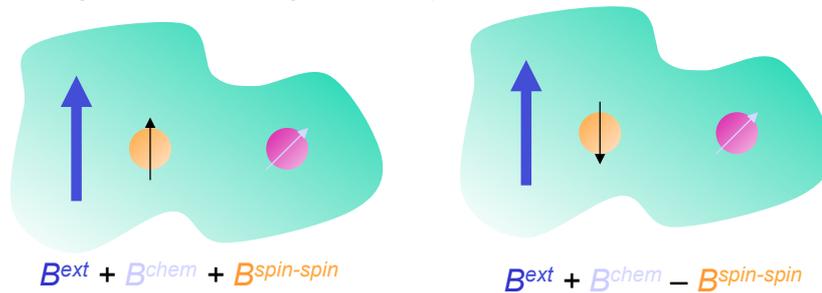
The structure (ensemble) determination problem

Examples of (*observable*) quantities $Q(\mathbf{r})$:

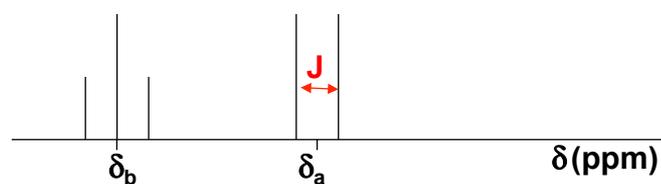
- NOE intensities or distances $\langle I_{i,j} \rangle \sim \langle r_{ij}^{-p} \rangle$ with $p = 3$ or 6
- **3J -coupling constants** $\langle ^3J_{i,j} \rangle = a \langle \cos^2 \theta_{i,j} \rangle + b \langle \cos \theta_{i,j} \rangle + c$
- Residual dipolar couplings $\langle D_{i,j} \rangle = a \langle \cos^2 \theta_{ij} \rangle + b$
- Chemical shifts $\langle \sigma \rangle$
- Structure factors (amplitudes) $\langle F_{hkl} \rangle$
- CD spectra $\langle I(\lambda) \rangle$

Spin-spin coupling

- Nuclei close to each other also influence each others signal (but not if they are chemically identical)



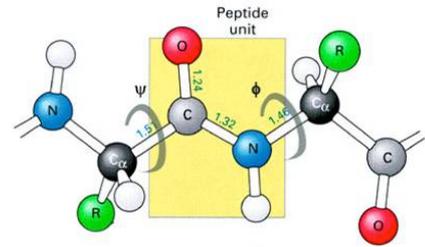
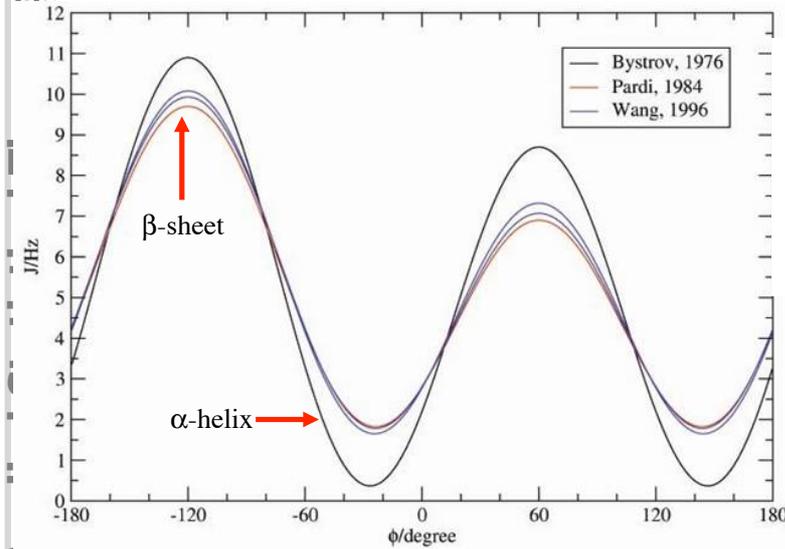
- Signal splits up with a spin-coupling constant J



- Amount of peaks depends on the number of neighbouring spins

Karplus relation

$${}^3J(\phi) = a \cos^2(\phi + \delta) + b \cos(\phi + \delta) + c; \quad \delta = -60^\circ$$



3J coupling constant between HN and H α depends on the angle ϕ



$\phi({}^3J)$ is a multiple-valued function of 3J
Accuracy of ${}^3J(\phi)$ is about 1 Hz

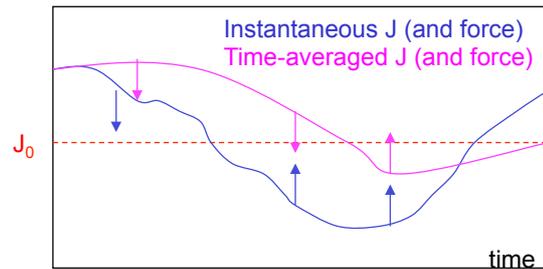
Applying restraints

- Instantaneous J-value restraining

$$V^{Jr} = \frac{1}{2} K^{Jr} [J(\phi(\mathbf{r})) - J_0]^2$$

- Time-average J-value restraining

$$V^{Jr} = \frac{1}{2} K^{Jr} [\bar{J}(\phi(\mathbf{r})) - J_0]^2$$



Problem with time averaging:

The force remains if the instantaneous value is already at the target

- For distances: -Van der Waals repulsion counteracts it
- Only attractive forces used
- r^{-3} averaging favours short distances

For J-values: time-averaging leads to large fluctuations

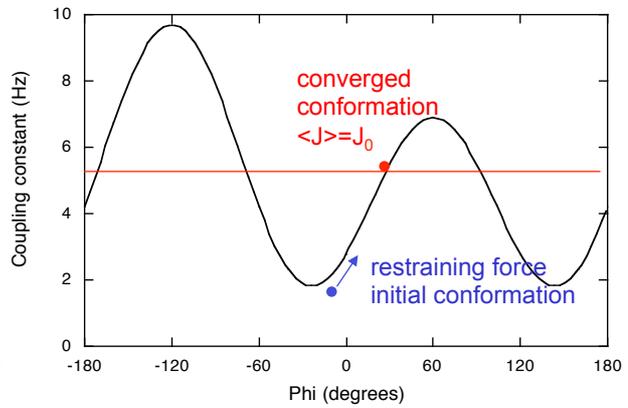
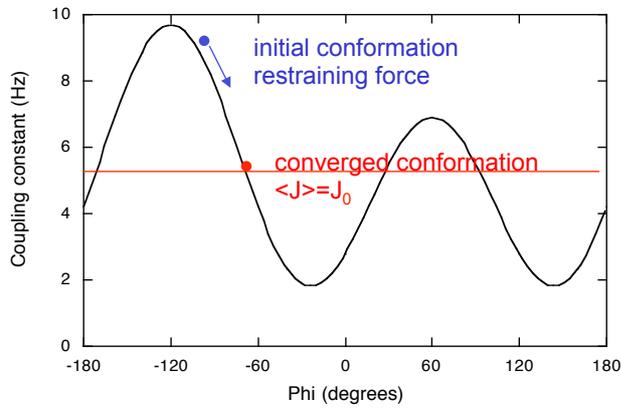
- Double harmonic J-value restraining

$$V^{Jr} = \frac{1}{2} K_{double}^{Jr} [J(\phi(\mathbf{r})) - J_0]^2 [\bar{J}(\phi(\mathbf{r})) - J_0]^2$$



Applying restraints

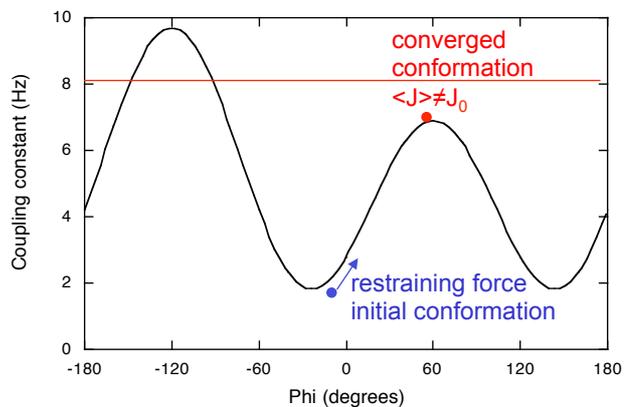
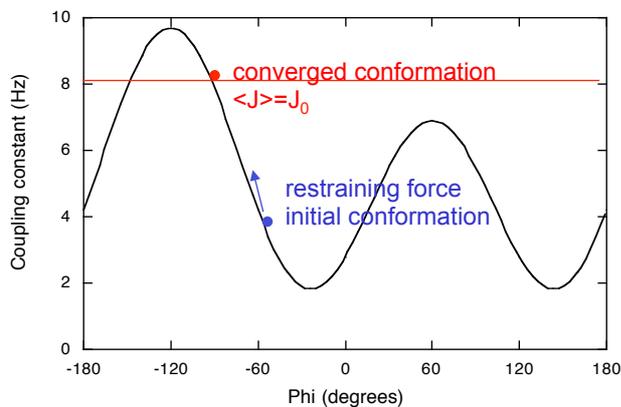
- Effect of multi-valued function $\phi(J)$



- The restraint will typically converged to a reasonable solution
- But another solution or multiple solutions may be relevant

Applying restraints

- Effect of multi-valued function $\phi(J)$



- The restraint may get locked in a (wrong) local minimum
- **The restraining potential should go over barriers to find all relevant solutions**



Search conformational space

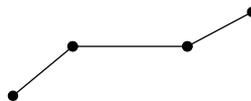
Idea: Include information obtained so far during the simulation into the search scheme: **memory function**

A. Characterize molecular conformations using:

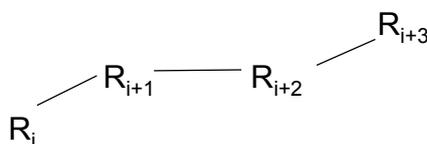
- cartesian coordinates

too many

- *torsional angles* φ, ψ, χ



- *dihedral angles* spanning residues:



M. Christen & W.F. van Gunsteren, J. Comput. Chem. 29 (2007) 157 - 166



B. Penalize the visited conformations by changing the energy function V as function of time

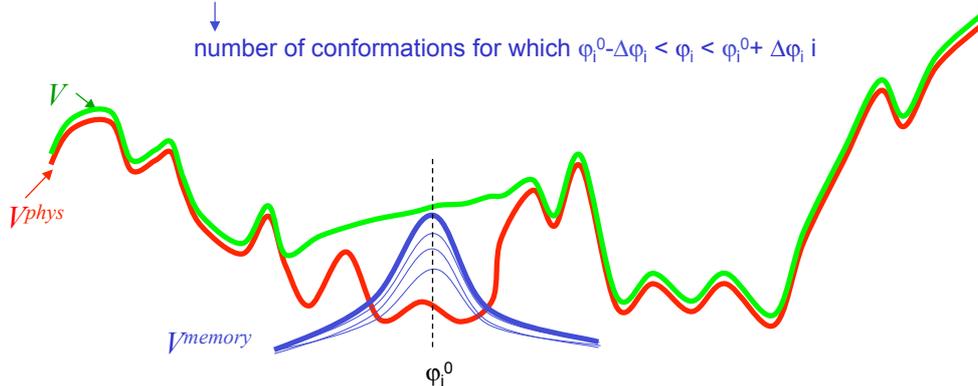
$$V(\mathbf{r}) = V^{phys}(\mathbf{r}) + V^{memory}(\varphi)$$

- potential energy term that pushes molecule out of the current conformation φ

$$V^{memory}(\varphi) = kN_{\varphi_i = \varphi_i^0} e^{-\sum_i (\varphi_i - \varphi_i^0)^2 / 2\omega^2}$$

number of conformations for which $\varphi_i^0 - \Delta\varphi_i < \varphi_i < \varphi_i^0 + \Delta\varphi_i$

Local elevation search
(in 2002 called *meta-dynamics*)

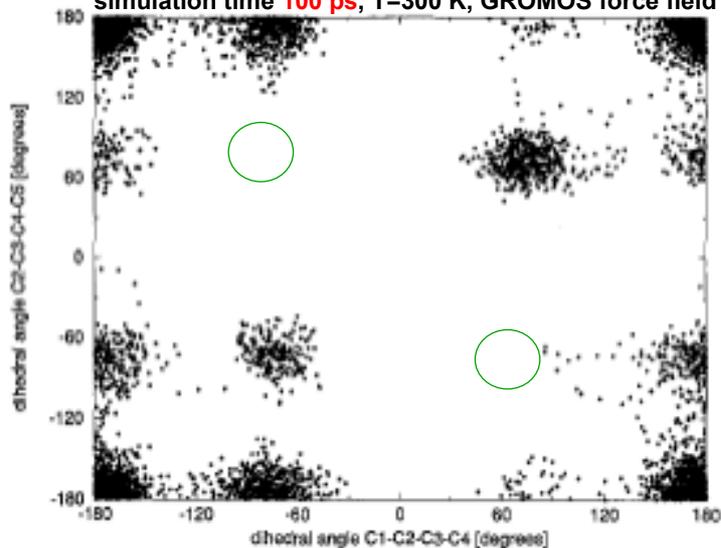


Thomas Huber et al., J. Comp. Aided Mol. Design 8 (1994) 695

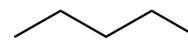
Test case: pentane

2 dihedral angles (3 minima each) → 9 low V_{phys} conformers

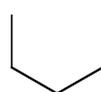
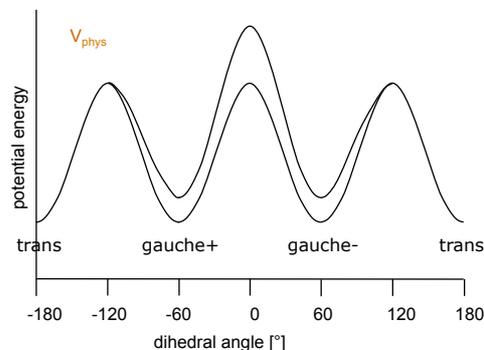
free SD-simulation (united atoms)
simulation time 100 ps, T=300 K, GROMOS force field



Higher-energy conformers are not (yet) sampled in 100 ps normal MD(SD) simulation



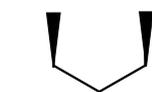
trans - trans *lowest*



cis - cis



gauche+ gauche+



gauche+ gauche-

highest

low

higher

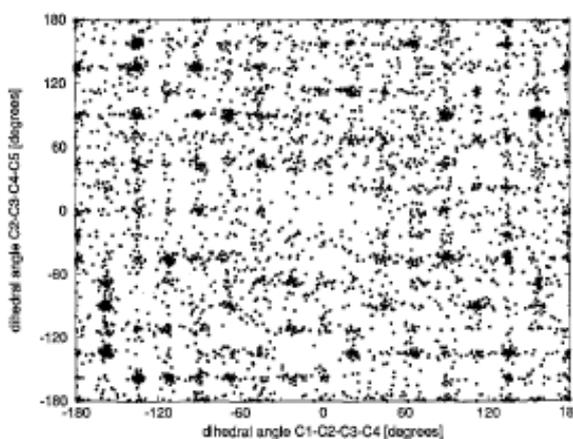
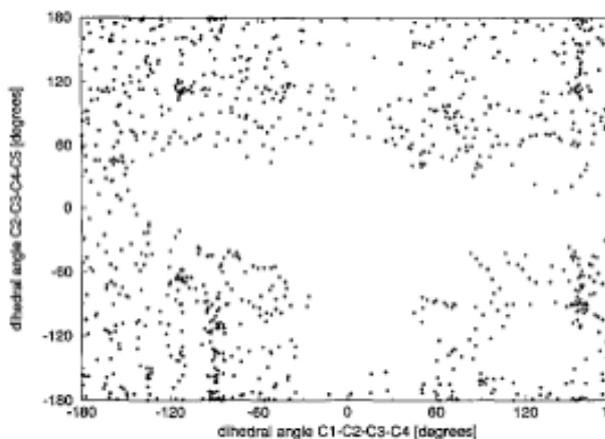
Thomas Huber et al., J. Comp. Aided Mol. Design 8 (1994) 695

Local elevation search: pentane

Local-elevation simulation of pentane (united atoms) T=300 K,
Gaussian local-elevation function with k=5kJ/mol per MD step

simulation time 20ps

simulation time 100ps



Higher-energy conformations are sampled in 20 ps local-elevation MD simulation

Almost all conformations are sampled in 100 ps LE-MD simulation

Application: ³J-value biasing

- Average ³J-coupling constants

$$\bar{J}(\mathbf{r}) = \frac{1}{\tau} \frac{1}{1 - e^{-t/\tau}} \int_0^t e^{-(t-t')/\tau} J(t') dt'$$

τ - memory decay time

- Restraining potential is the sum of N_{le} terms

$$V^{Jr}(\phi(\mathbf{r})) = \sum_{i=1}^{N_{le}} V_i^{le}(\phi)$$

- which are Gaussian Local Elevation functions, with an extra factor

$$V_i^{le}(\phi) = k^{Jr} w_{\phi_i}(t) e^{-(\phi - \phi_i^0)^2 / 2(\Delta\phi^0)^2}$$

a possible potential V_i at equally separated values of ϕ_i , distance $\Delta\phi^0$
 $N(\phi = \phi_i^0)$ in original local elevation formulation replaced by $w_{\phi_i}(t)$

Application: ³J-value biasing

- Local elevation potential added with weight $w_{\phi_i}(t)$

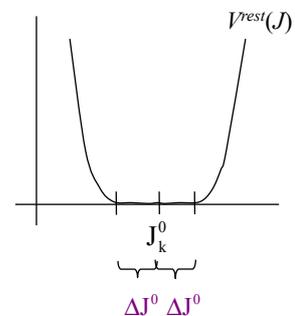
$$V_i^{le}(\phi) = k^{Jr} w_{\phi_i}(t) e^{-(\phi - \phi_i^0)^2 / 2(\Delta\phi^0)^2}$$

where

$$w_{\phi_i}(t) = \frac{1}{t} \int_0^t \delta_{\phi(t')\phi_i^0} V^{rest}(\bar{J}(t')) V^{rest}(J(t')) dt'$$

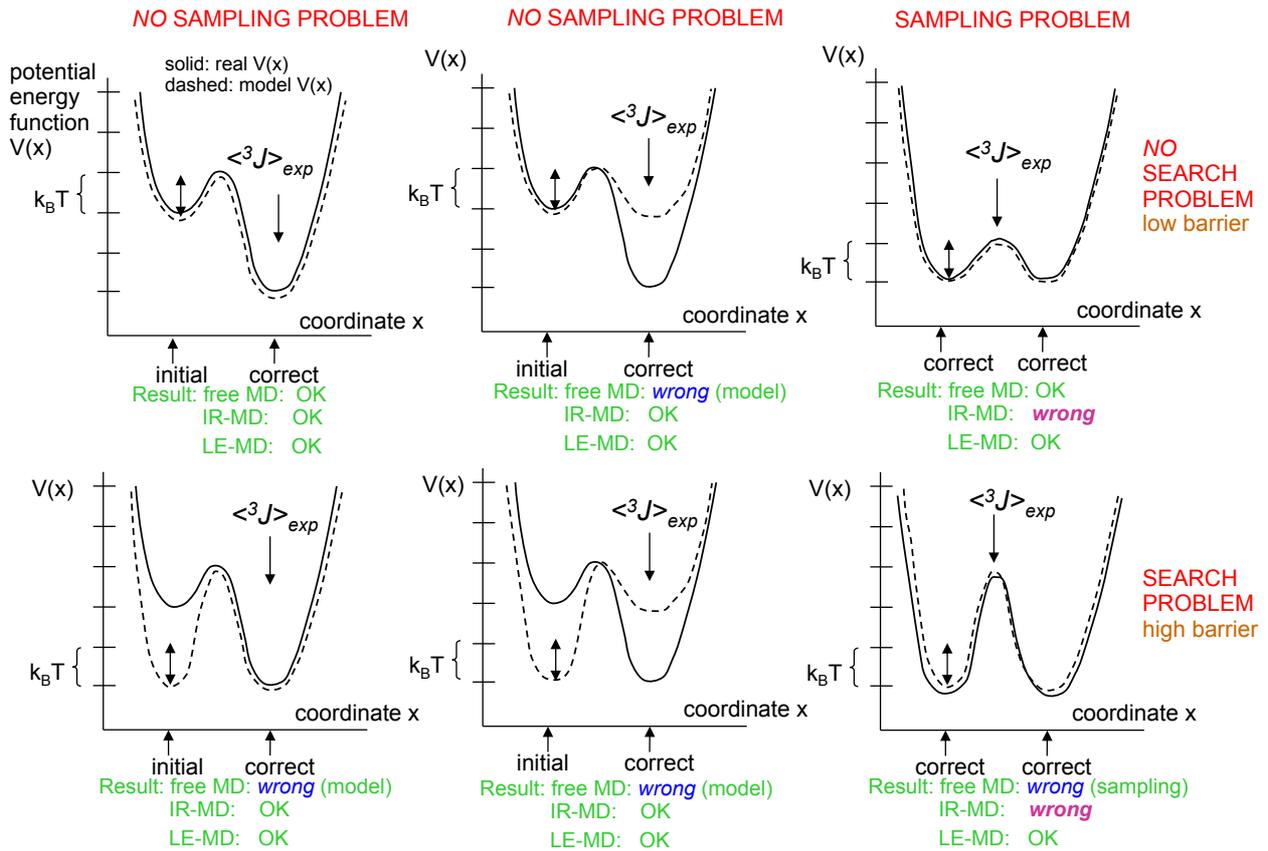
$$\delta_{\phi(t')\phi_i^0} = \begin{cases} 1 & \text{if } \phi(t') \text{ in } \phi_i^0 \text{ bin} \\ 0 & \text{otherwise} \end{cases}$$

$$V^{rest}(J(t')) = \begin{cases} (J(t') - J^0 - \Delta J^0)^2 & \text{for } J(t') > J^0 + \Delta J^0 \\ (J(t') - J^0 + \Delta J^0)^2 & \text{for } J(t') < J^0 - \Delta J^0 \\ 0 & \text{otherwise} \end{cases}$$



- We add to the V_i^{le} potential if:
 - We are in the bin i ($\delta_{\phi(t')\phi_i^0}$)
 - Both the average and the instantaneous J value are more than ΔJ^0 away from the target J^0 ($V^{rest}(\bar{J}(t'))$ and $V^{rest}(J(t'))$)

Why is local-elevation sampling better than instantaneous restraining ?



Example: GCN4

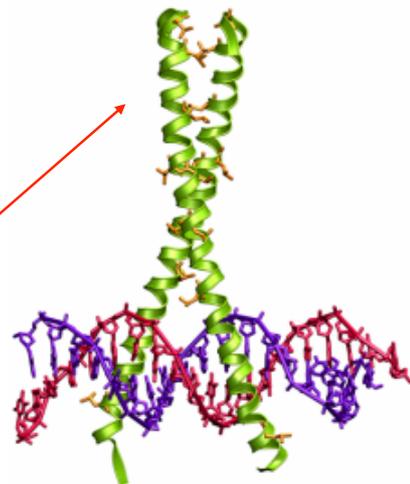
- parallel two-stranded **leucine zipper** of the yeast transcriptional activator
 - two-stranded structural motif present in many gene regulatory proteins

- 33 residues:

ARG - MET LYS GLN **LEU** GLU ASP LYS -
 - VAL GLU GLU **LEU** LEU SER LYS -
 - ASN TYR HIS **LEU** GLU ASN GLU -
 - VAL ALA ARG **LEU** LYS LYS **LEU** -
 - VAL GLY GLU ARG

N-terminal half
GCN4p2-17

C-terminal half
GCN4p16-31



- how does the zipper form ?**
C- terminal is a autonomous helical folding unit & trigger site for the **coiled-coil formation**

PDB ID 1YSA (the GCN4 leucine zipper binds to the DNA major groove, X-ray structure)

GCN4p16-31: NMR structure determination

Experimental data: **181 NOEs** and **15 $^3J_{\text{HN-H}\alpha}$** coupling constants

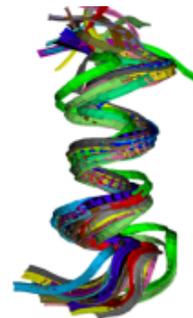
Single-structure refinement using the program XPLOR:

Simulated annealing from 200 different structures with randomised torsional angles using

- 172 NOE distance restraints
- 14 α -helical hydrogen bond restraints
- 8 ϕ -torsional angle restraints (residues 17 – 24)

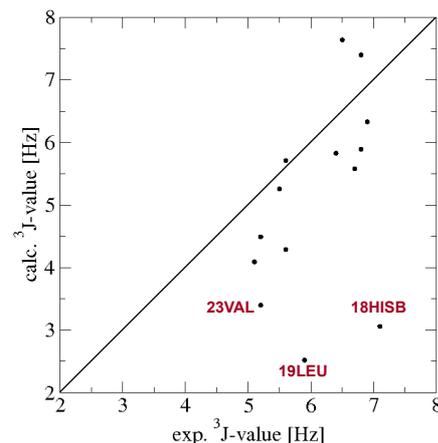
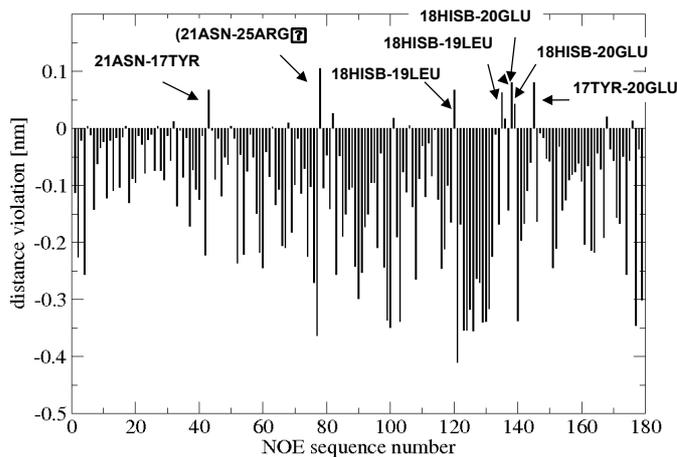
Result:

A set of 20 lowest energy structures, indicated as the set of **20 NMR model structures** →



M.O. Steinmetz et al., PNAS 104 (2007) 7062-7066

NOE and 3J -value analysis of NMR model structures



⇒ the set of NMR model structures violates the 3J coupling constants

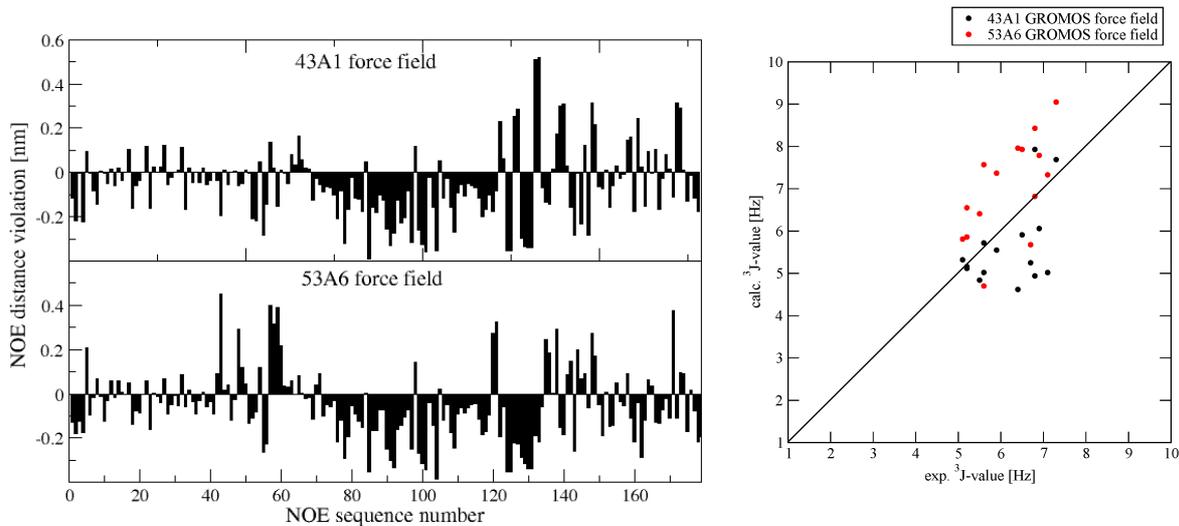
Are the NMR model structures representative for the real conformational ensemble ?

→ improve the protocol for structure determination by including time averaging

Are there inconsistencies between NOEs and 3J values ?

→ find a conformational Boltzmann ensemble that satisfies all the primary data

Unrestrained molecular dynamics simulations

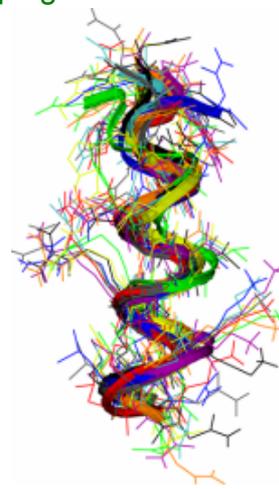
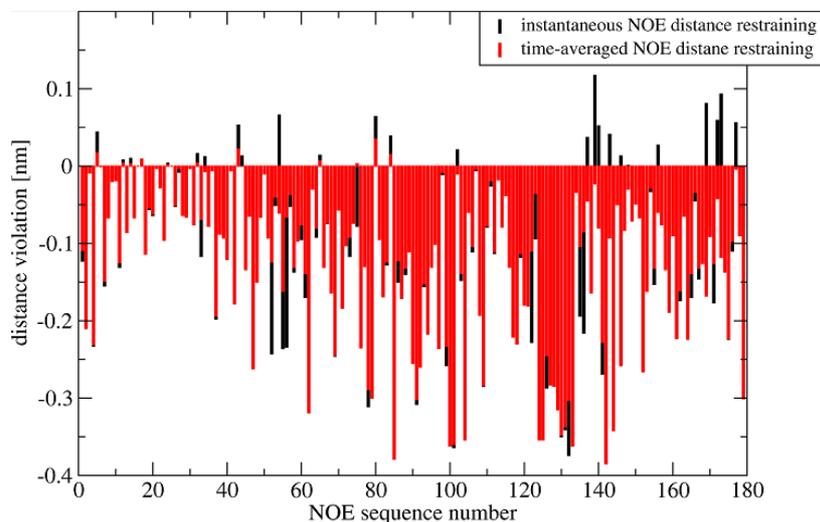


⇒ using unrestrained MD simulations and two different GROMOS force fields all the experimental data could not be satisfied on a 50 ns time scale:

→ force field problem ?
→ sampling problem ?

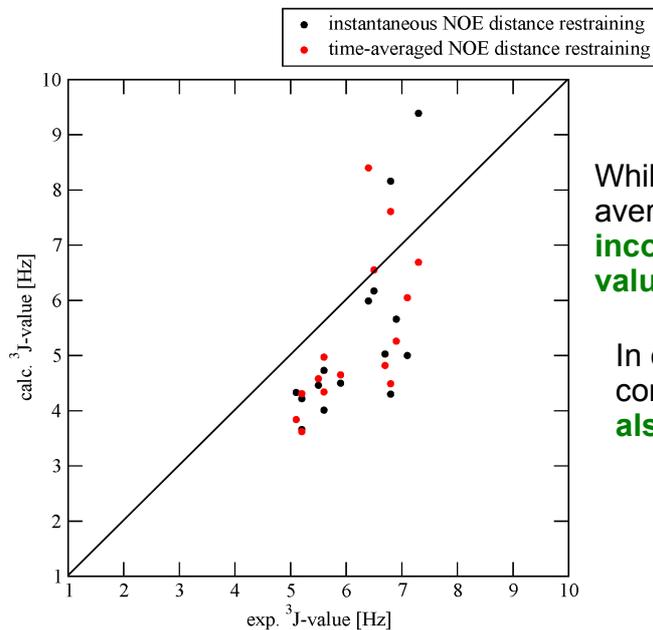
MD with distance restraints

1. **instantaneous distance restraints:** put an extra force on the atoms pulling them instantaneously to the experimental NOE distance → strain in the system
2. **time-averaged distance restraints:** put an extra force on the atoms such that the NOE distance is fulfilled *on average* → side chains flipping back and forth



NOEs are not compatible with *one* structure

MD with instantaneous distance restraints: average 3J -values

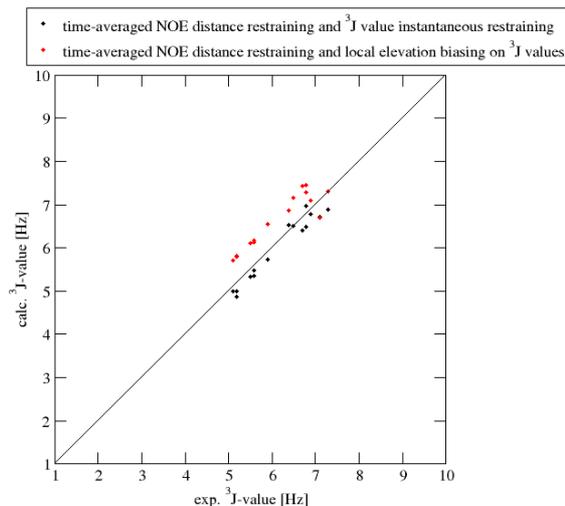
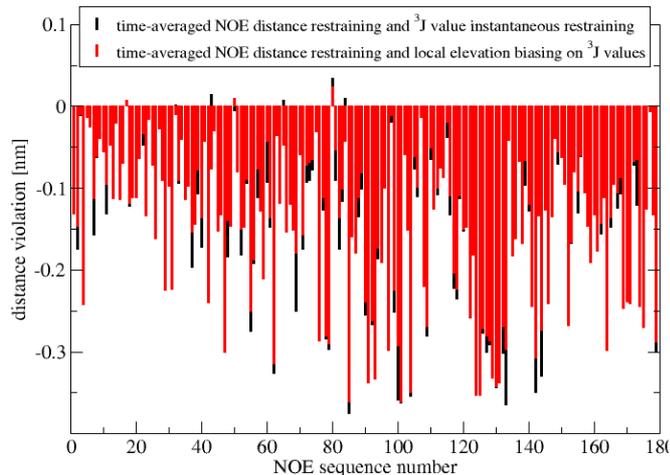


While restraining distances to the NOEs the averaged 3J -values turn out to be incompatible with the experimental 3J -values

In order to satisfy the 3J -coupling constants 3J -value restraints should also be used in the simulations



MD simulations with time-averaged NOE distance restraints and instantaneous restraining on 3J -values or local elevation biasing on 3J -values

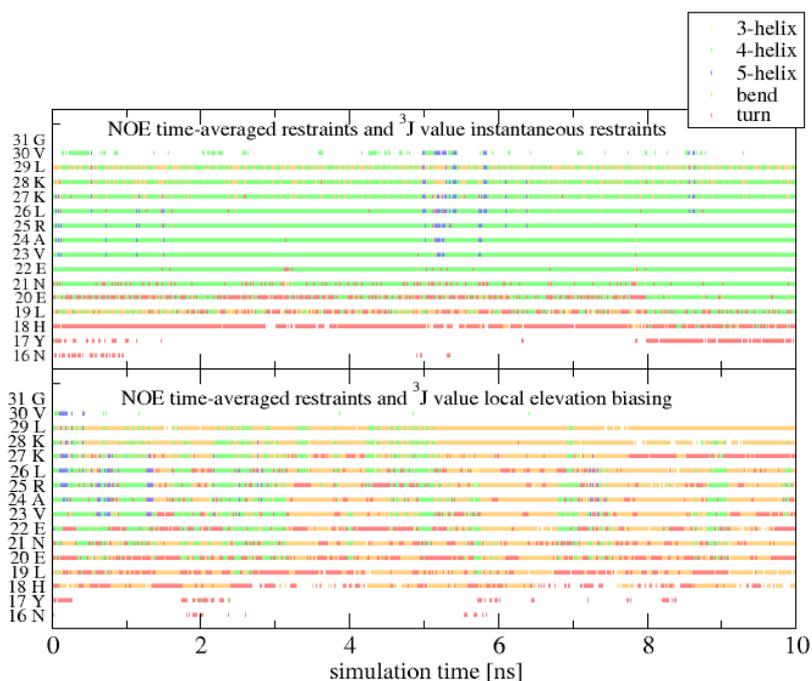


Using time-averaged distance restraints based on NOEs and local elevation biasing on 3J -coupling constants an ensemble of structures was obtained that satisfies the entire set of experimental data

The results are force field independent



Structural differences



Instantaneous restraining *artificially restricts* the motion
Local-elevation biasing MD allows more motion



FRET efficiencies

- Radiation-free energy transfer between fluorophores D and A
- FRET efficiency

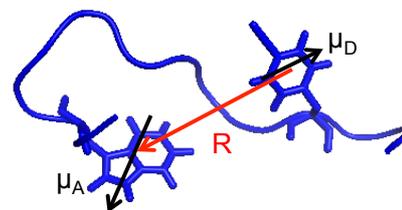
$$E(\mu_D, \mu_A, R) = \frac{1}{1 + (R/R_F)^6} = \frac{R_F^6}{R_F^6 + R^6}$$

- Förster radius

$$R_F = \left(\frac{\kappa^2 \Phi_D^0 J}{n^4} \right)^{1/6}$$

- and

$$\kappa = \cos \theta_{AD} - 3 \cos \theta_A \cos \theta_D$$



- So if we apply a restraint

$$V^{FRET}(\mathbf{r}) = \frac{1}{2} K [E(\mu_D, \mu_A, R) - E^0]^2$$

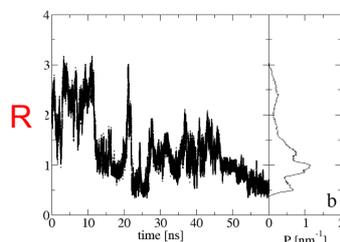
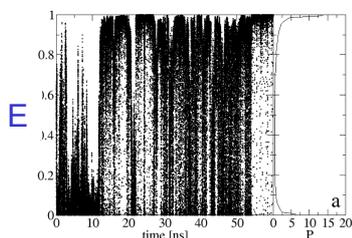
- Force is derivative of energy with respect to \mathbf{r}
 - Gives rise to terms due to R and due to κ^2



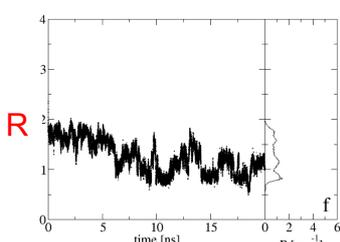
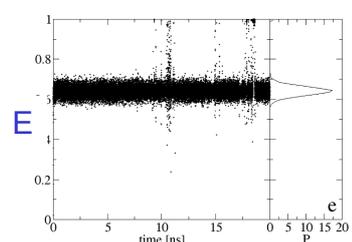
FRET efficiencies

- Small peptide in urea-water mixture
- Experimental FRET efficiency ~ 0.64

- Free simulation:

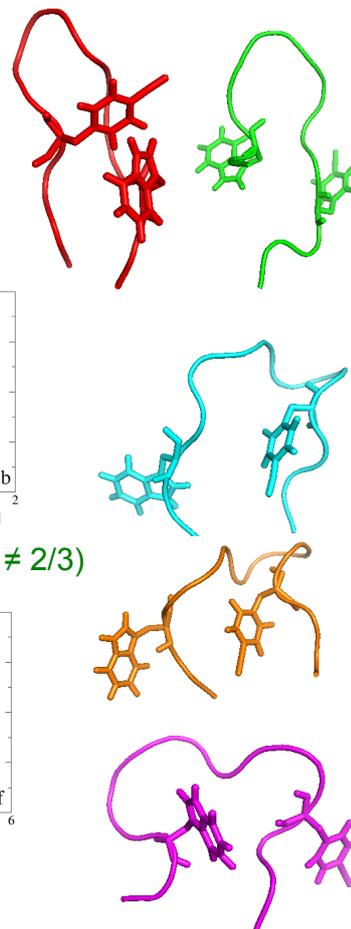


- Restrained simulation:



- Efficiency gets focused, still fluctuations in R

$(\kappa^2 \neq 2/3)$



The structure (ensemble) determination problem

Examples of (*observable*) quantities $Q(\mathbf{r})$:

- NOE intensities or distances $\langle I_{i,j} \rangle \sim \langle r_{ij}^{-p} \rangle$ with $p = 3$ or 6
- 3J -coupling constants $\langle ^3J_{i,j} \rangle = a \langle \cos^2 \theta_{i,j} \rangle + b \langle \cos \theta_{i,j} \rangle + c$
- Residual dipolar couplings $\langle D_{i,j} \rangle = a \langle \cos^2 \theta_{ij} \rangle + b$
- Chemical shifts $\langle \sigma \rangle$
- Structure factors (amplitudes) $\langle F_{hkl} \rangle$
- FRET Efficiencies $\langle E(R, \theta) \rangle$
- CD spectra $\langle I(\lambda) \rangle$

(implemented in GROMOS)



Summary

When using experimental data to bias the sampling:

1. Use only primary (measured) not secondary (derived) experimental data
2. Determine their accuracy and test for consistency
3. Do account for motional averaging
4. Do not restrain to a measured value of an observable if the function connecting structure to observable is multiple-valued, but use LE-biasing to enhance sampling and to compensate force-field deficiencies

Examples:

- (time-averaged) distance restraints
- Local elevation biased J-value restraining

Backup slides

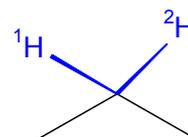
PSEUDOATOM AND MULTIPLICITY CORRECTIONS

Practical considerations

- Some force fields use united atoms for CH₁, CH₂, CH₃ groups

Apply the distance restraint to a **virtual atom**
calculate the position on the fly from the relevant heavy atoms
transfer the forces to the heavy atoms

- The signals of multiple protons may not be distinguishable
 - Intrinsically (chemically identical protons, e.g. -CH₃)
 - Given the experimental settings

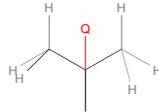
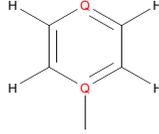


Apply the distance restraint to a **pseudo atom**
calculate the position on the fly from the relevant heavy atoms
transfer the forces to the heavy atoms
additional corrections required to the reference distance

Pseudo atom corrections

- Identical spins are considered to be represented by one pseudo-atom.
- A **multiplicity correction** for the fact that the real signal comes from more than one atom is added
- A **pseudoatom correction** for the fact that the 'closest' of the atoms will contribute most, while the pseudo atom is further away.

Corrections

	Wüthrich (1983)	Fletcher (1996)	GROMOS (1996)	
	corr (Å)	Z corr (Å)	corr (Å)	
3.		+1.0	$\times 2^{1/p}$ +0.7	+0.9
5.		+1.5	$\times 3^{1/p}$ +0.4	+1.0
6.		+2.9	$\times 6^{1/p}$ +1.5	+2.2
7.		+2.0	$\times 2^{1/p}$ +2.0	+2.1