the
abdus salam
international centre for theoretical physics

SMR: 1098/8

WORKSHOP ON THE STRUCTURE OF BIOLOGICAL MACROMOLECULES

(16 - 27 March 1998)

"Structure Determination of Biomacromolecules by NMR"

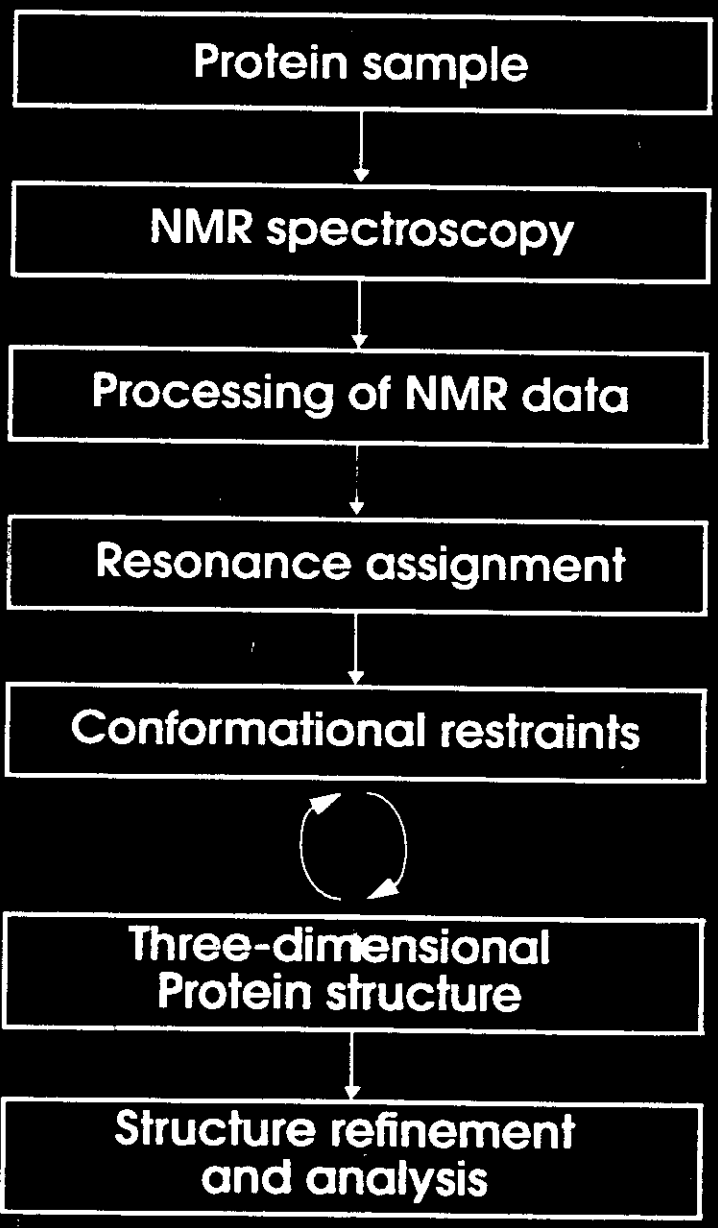
presented by:

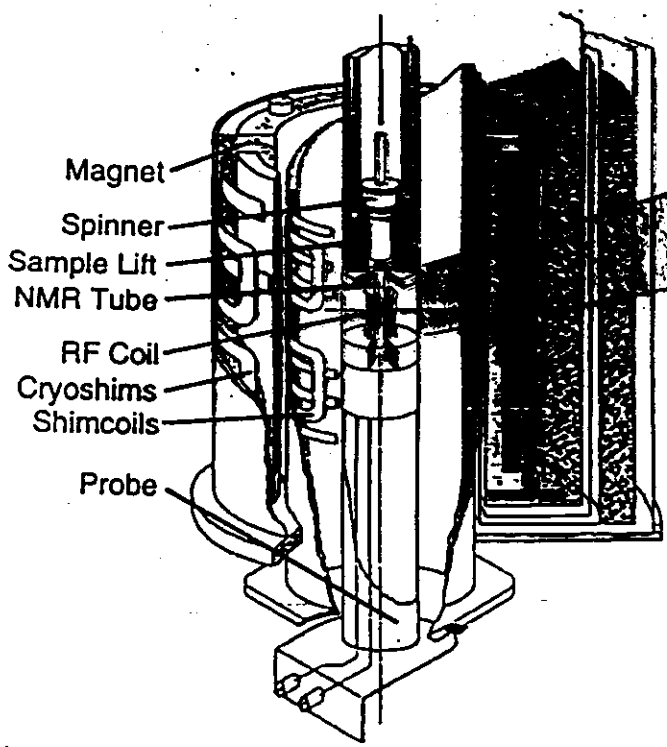
Peter GUENTERT

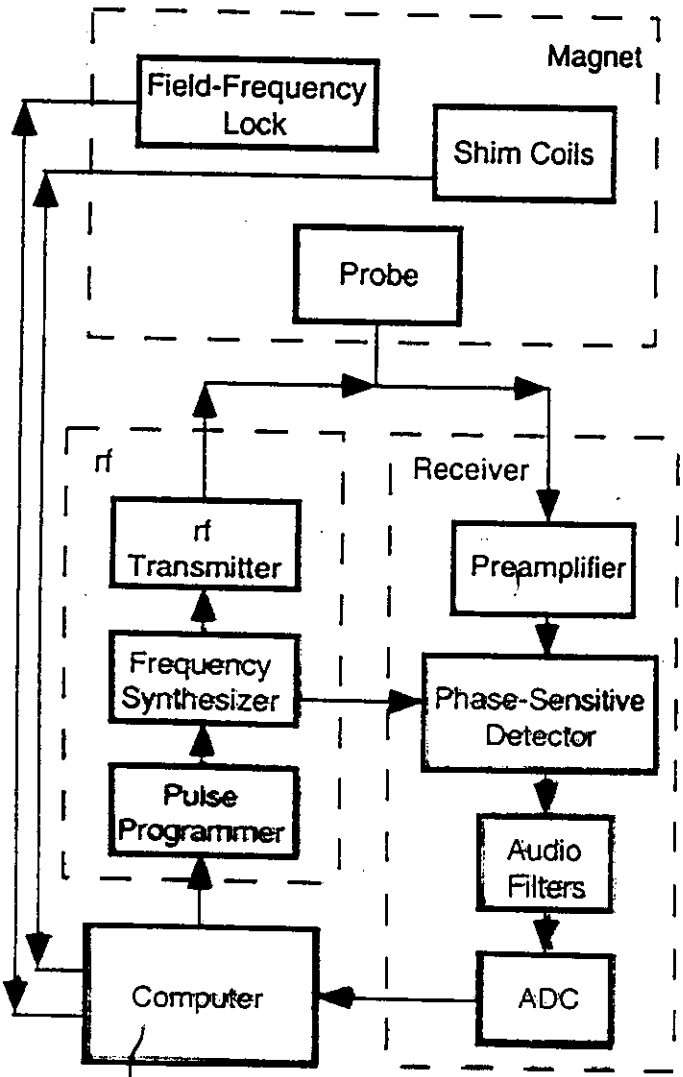
Eidgenossische Technische Hochschule Honggerberg
Institut für Molekularbiologie und Biophysik
CH-8093 Zürich
Switzerland

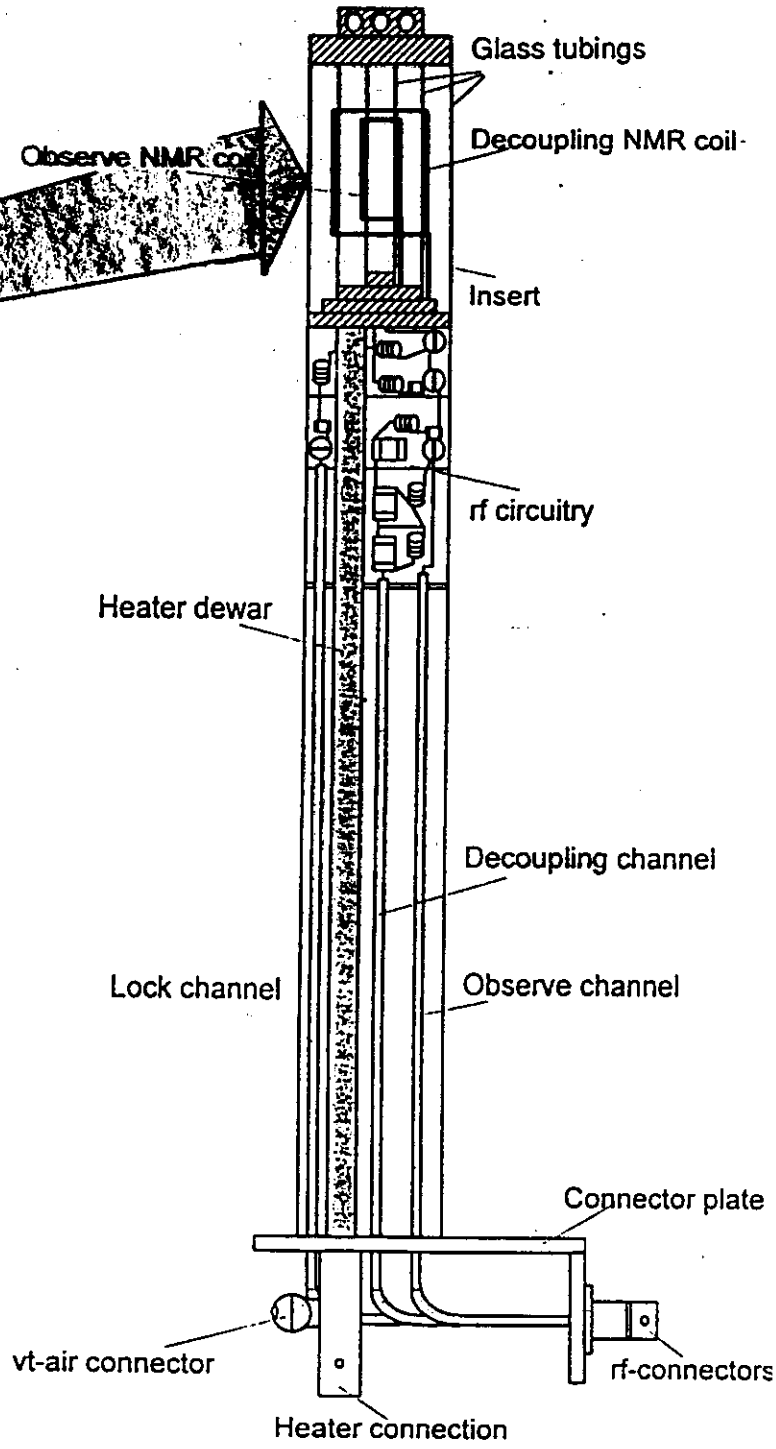
STRUCTURE DETERMINATION OF BIOMACROMOLECULES BY NMR

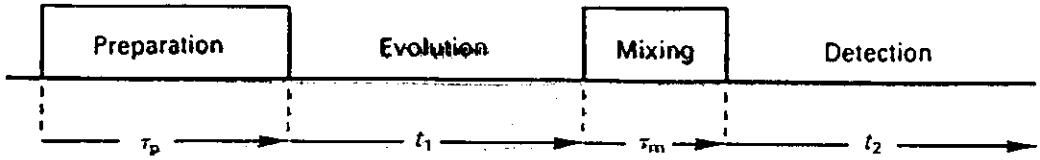
Protein structure determination by NMR

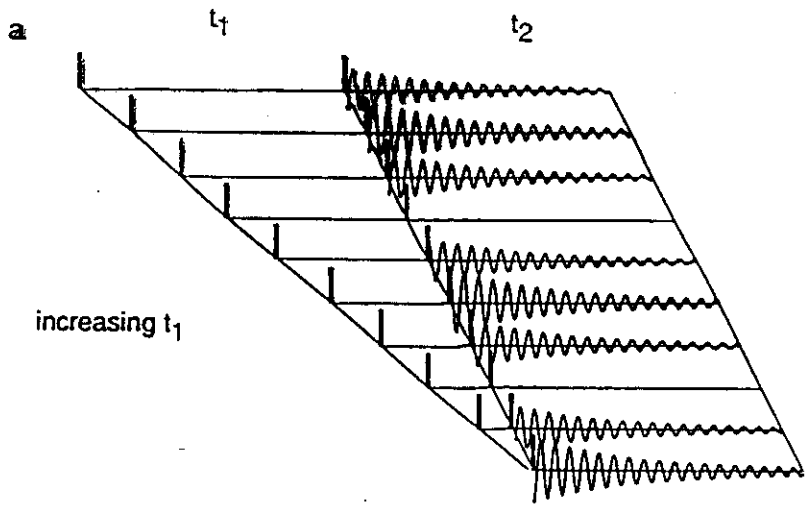



Magnet
Spinner
Sample Lift
NMR Tube
RF Coil
Cryoshims
Shimcoils
Probe


Magnet
Field-Frequency Lock
Shim Coils
Probe
rf
Receiver
rf Transmitter
Preamplifier
Frequency Synthesizer
Phase-Sensitive Detector
Pulse Programmer
Audio Filters
Computer
ADC

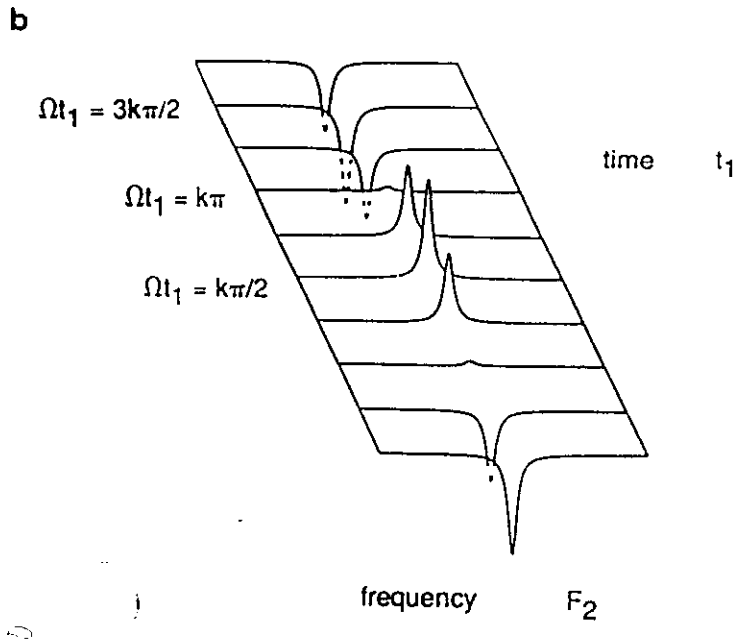

Glass tubings
Observe NMR coil
Decoupling NMR coil
Insert
rf circuitry
Heater dewar
Decoupling channel
Lock channel
Observe channel
Connector plate
vt-air connector
rf-connectors
Heater connection

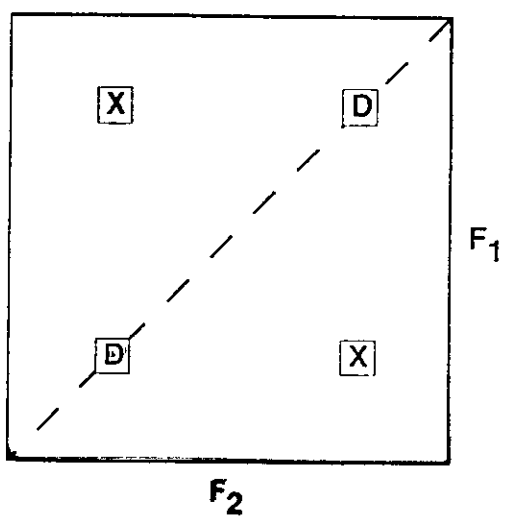




Fourier transformation

with respect to t_2





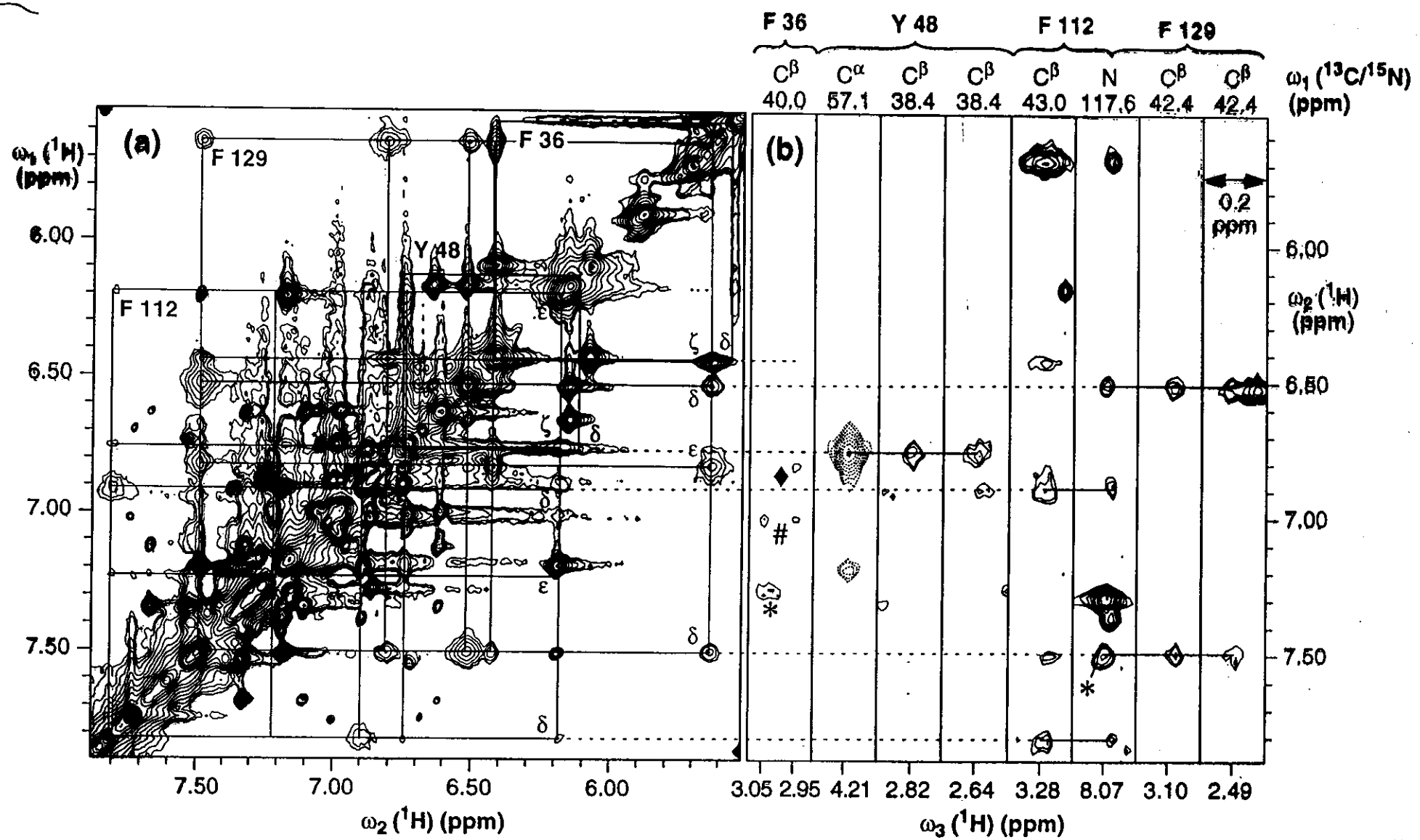
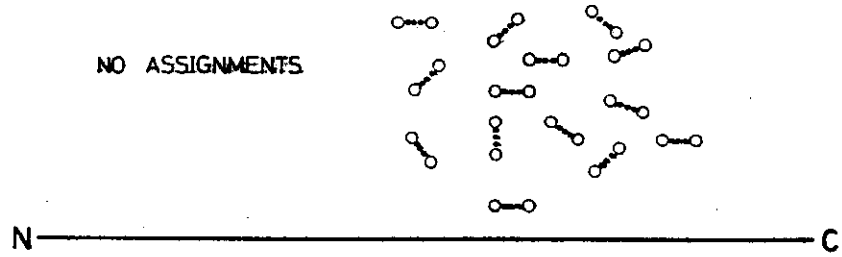
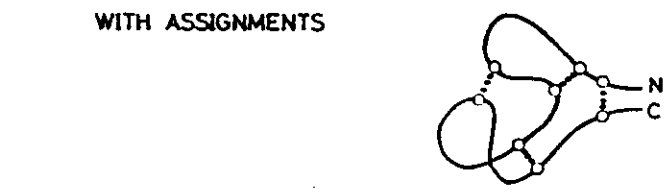
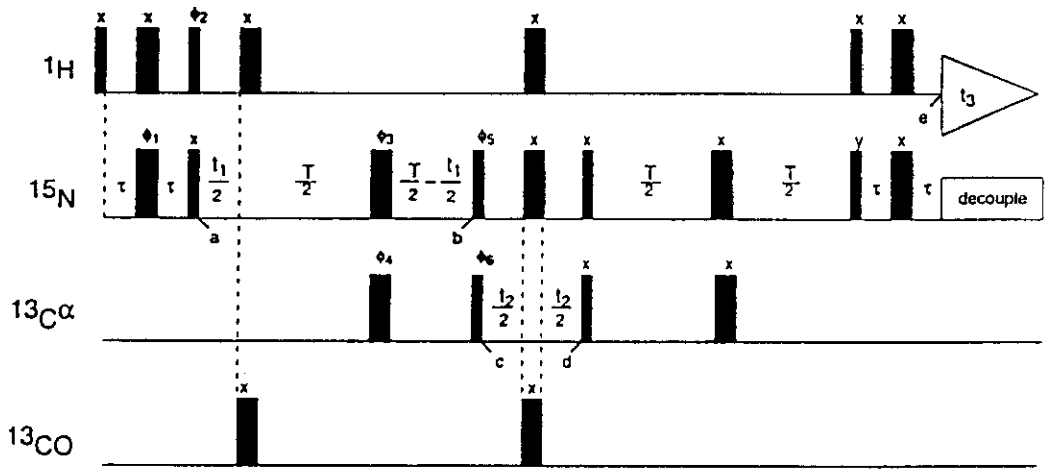

(a)
(b)
F 129
F 36
Y 48
F 112
F 36
Y 48
F 112
F 129
Cβ 40.0
Cα 57.1
Cβ 38.4
Cβ 38.4
Cβ 43.0
N 117.6
Cβ 42.4
Cβ 42.4
ω1 (13C/15N) (ppm)
0.2 ppm
ω1 (1H) (ppm)
ω2 (1H) (ppm)
6.00
6.50
7.00
7.50
7.50 7.00 6.50 6.00
3.05 2.95 4.21 2.82 2.64 3.28 8.07 3.10 2.49
ω2 (1H) (ppm)
ω3 (1H) (ppm)

Fig. 2

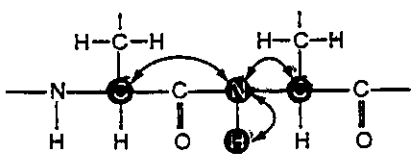






Pulse sequences for the HNCA experiment.

Magnetization transfer



Correlations observed

$^{1}H_{i}^{N}-^{15}N_{i}-^{13}C_{i}^{\alpha}$

$^{1}H_{i}^{N}-^{15}N_{i}-^{13}C_{i-1}^{\alpha}$

Interaction: $H\tau$
$\sigma(t) \rightarrow \sigma(t+\tau) = e^{-iH\tau}\sigma(t)e^{iH\tau}$

σ density matrix

H Hamilton operator (time-independent)

τ time during which the interaction H is active

Free precession: $H_f = \sum_k \omega_k I_{kz} + \sum_{k<l} 2\pi J_{kl} I_{kz} I_{lz}$

$\omega_k, \omega_l, \ldots$ chemical shifts

$J_{mn}, J_{pq}, \ldots$ scalar spin-spin couplings

I_{kx}, I_{ky}, I_{kz} spin operators

Radio-frequency pulse: $H_p\tau = \beta \sum_k [I_{kx}\cos\phi + I_{ky}\sin\phi]$

β rotation angle

ϕ phase

chemical shift: $\omega_k \tau I_{kz}$	scalar spin-spin coupling: $2\pi J_{kl}\tau I_{kz} I_{lz}$
$I_{kx} \rightarrow I_{kx}\cos\omega_k\tau + I_{ky}\sin\omega_k\tau$	$2I_{k\alpha}I_{l\beta} \rightarrow 2I_{k\alpha}I_{l\beta} \quad (\alpha, \beta \neq z)$
$I_{ky} \rightarrow I_{ky}\cos\omega_k\tau - I_{kx}\sin\omega_k\tau$	$2I_{kx}I_{lz} \rightarrow 2I_{kx}I_{lz}\cos\pi J_{kl}\tau + I_{ky}\sin\pi J_{kl}\tau$
	$2I_{ky}I_{lz} \rightarrow 2I_{ky}I_{lz}\cos\pi J_{kl}\tau - I_{kx}\sin\pi J_{kl}\tau$
	$I_{kx} \rightarrow I_{kx}\cos\pi J_{kl}\tau + 2I_{ky}I_{lz}\sin\pi J_{kl}\tau$
	$I_{ky} \rightarrow I_{ky}\cos\pi J_{kl}\tau - 2I_{kx}I_{lz}\sin\pi J_{kl}\tau$

pulse with rotation angle β and phase ϕ: $\beta[I_{kx}\cos\phi + I_{ky}\sin\phi]$

$I_{kz} \rightarrow I_{kz}\cos\beta + I_{kx}\sin\beta\sin\phi - I_{ky}\sin\beta\cos\phi$

$I_{kx} \rightarrow -I_{kz}\sin\beta\sin\phi + I_{kx}(\cos\beta\sin^2\phi + \cos^2\phi) + I_{ky}\sin^2(\beta/2)\sin 2\phi$

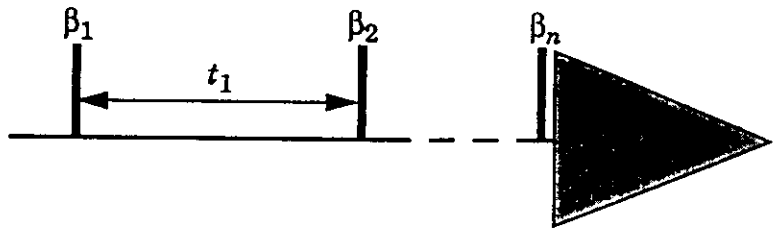
$I_{ky} \rightarrow I_{kz}\sin\beta\cos\phi + I_{kx}\sin^2(\beta/2)\sin 2\phi + I_{ky}(\cos\beta\cos^2\phi + \sin^2\phi)$

x-pulse: βI_{kx}	y-pulse: βI_{ky}	z-pulse: βI_{kz}
$I_{kz} \rightarrow I_{kz}\cos\beta - I_{ky}\sin\beta$	$I_{kz} \rightarrow I_{kz}\cos\beta + I_{kx}\sin\beta$	$I_{kx} \rightarrow I_{kx}\cos\beta + I_{ky}\sin\beta$
$I_{ky} \rightarrow I_{ky}\cos\beta + I_{kz}\sin\beta$	$I_{kx} \rightarrow I_{kx}\cos\beta - I_{kz}\sin\beta$	$I_{ky} \rightarrow I_{ky}\cos\beta - I_{kx}\sin\beta$

POMA: Product Operators in MATHEMATICA

(http://mol.biol.ethz.ch/wuthrich/software/poma)

spin[k, α]	**spin operator** $I_{k\alpha}$ ($\alpha = x, y, z$; $k = 1, 2, \ldots$)
nucleus[k] = X	set name of spin k to X
delay[t, {mn, pq, ...}]	free precession during time t with scalar couplings J_{mn}, J_{pq}, ...
delay[t, {mn, pq, ...}, {k, l, ...}]	free precession with chemical shifts only for spins {k, l, ...}
pulse[β, ϕ]	non-selective radio-frequency pulse with rotation angle β and phase ϕ
pulse[β, {ϕ_1, ϕ_2, ...}]	non-selective pulse with rotation angle β and phase cycle {ϕ_1, ϕ_2, ...}
pulse[β, {ϕ_1, ϕ_2, ...}, {k, l, ...}]	selective pulse for spins {k, l, ...}
receiver[{ψ_1, ψ_2, ...}]	receiver phase cycle {ψ_1, ψ_2, ...}
observable	all observable terms
observable[{k, l, ...}]	observable terms for spins {k, l, ...}
sort	sort according to Sørensen *et al.* (1983)



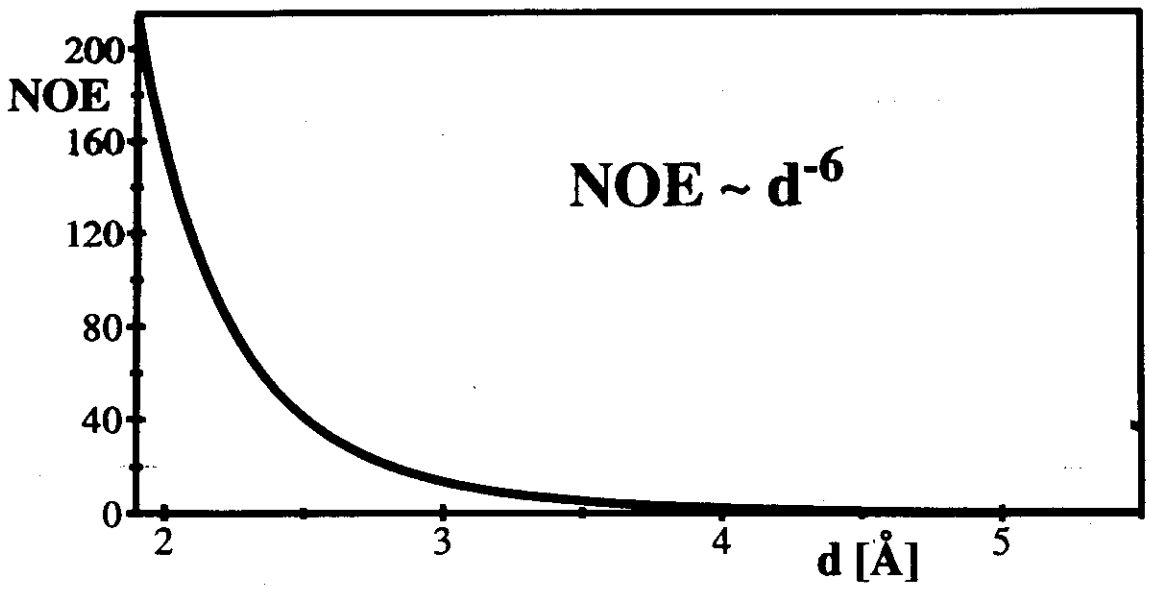
Pulse sequence: pulse β_1 — delay t_1 — β_2 — · · · — β_n — detection

In POMA: **pulse**[β_1, ...] // **delay**[t_1, ...] // **pulse**[β_2, ...] //... // **pulse**[β_n, ...] // **receiver**[...] // **observable**

Result: *analytical* expression for the density operator

STRUCTURAL DATA: NOEs

Calibration:



Use upper limits rather than exact distances
(NOE intensity may be reduced)

Calibration constant: **use NOEs of known distances**
(NOE intensities are not precise!)

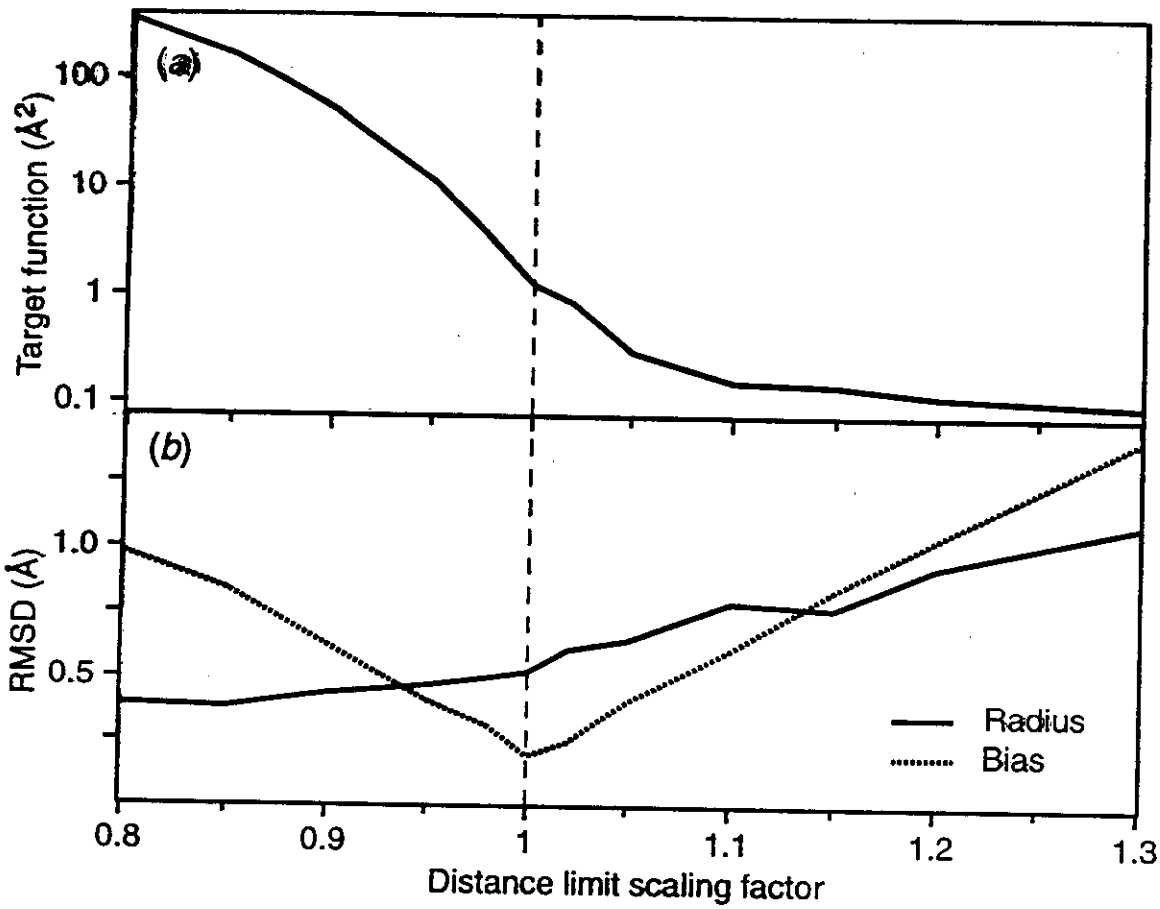
a) individual fixed distances

b) average backbone distance is 3.4 Å

c) distances in α-helices are well-defined

d) optimize calibration with preliminary structures

Influence of distance restraint calibration



TIME-AVERAGED NOEs

Example

consider the distance *d* in an ensemble of 1000 structures:

A: rigid, 1000 with d = 3.34Å

$$\langle NOE \rangle = k \cdot \left(\frac{1000}{3.34^6} \right) = k \cdot 0.72$$

B: mobile, 500 with d = 3Å and 500 with d = 5Å:

$$\langle NOE \rangle = k \cdot \left(\frac{500}{3^6} + \frac{500}{5^6} \right) = k \cdot (0.69 + 0.03) = k \cdot 0.72$$

⇒ a NOE consistent with the rigid structure A is also consistent with the mobile structure B

Memory function

- **if a constraint is satisfied, slowly forget it**
- **if a constraint is violated, slowly remember it**

$$E = w \cdot (\overline{d(t)} - u)^2 \qquad if\ \overline{d(t)} > u$$

$$\overline{d(t)^{-3}} = (1 - e^{-\Delta t/\tau}) \cdot d(t)^{-3} + e^{-\Delta t/\tau} \cdot \overline{d(t-\Delta t)^{-3}}$$

τ: "memory relaxation"

Problem

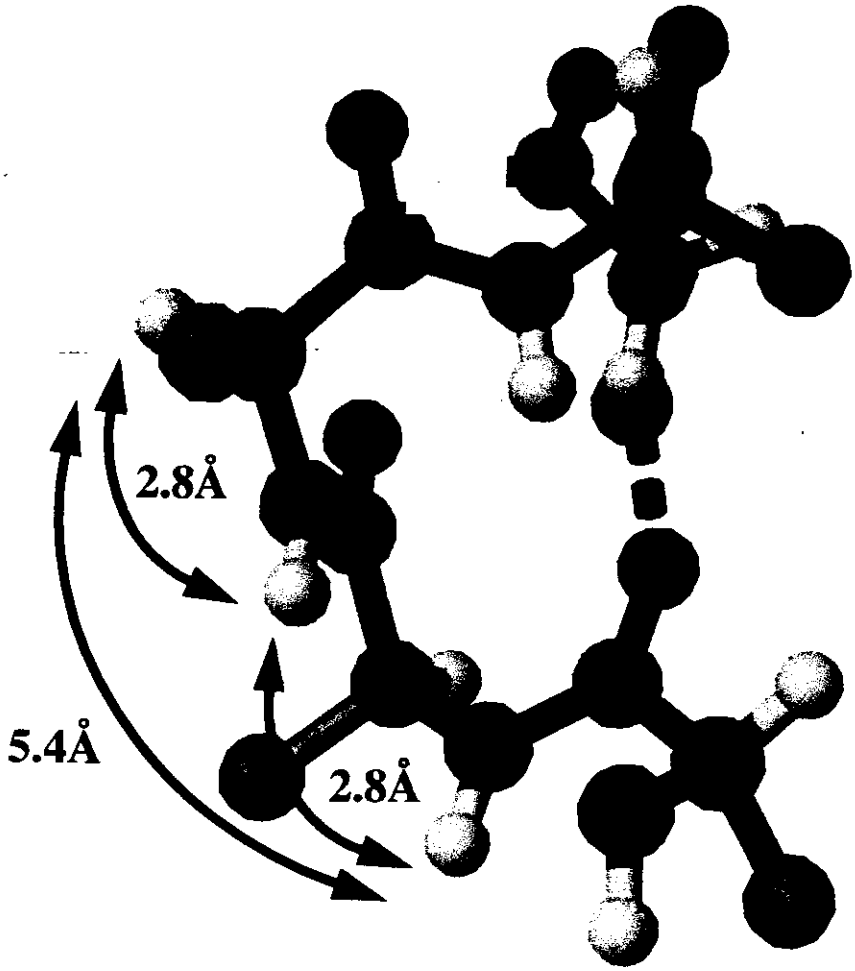
motions with given frequency (~1/τ) are induced

(van Gunsteren et al., *Methods Enzymol.* 239, 619-654, 1994)

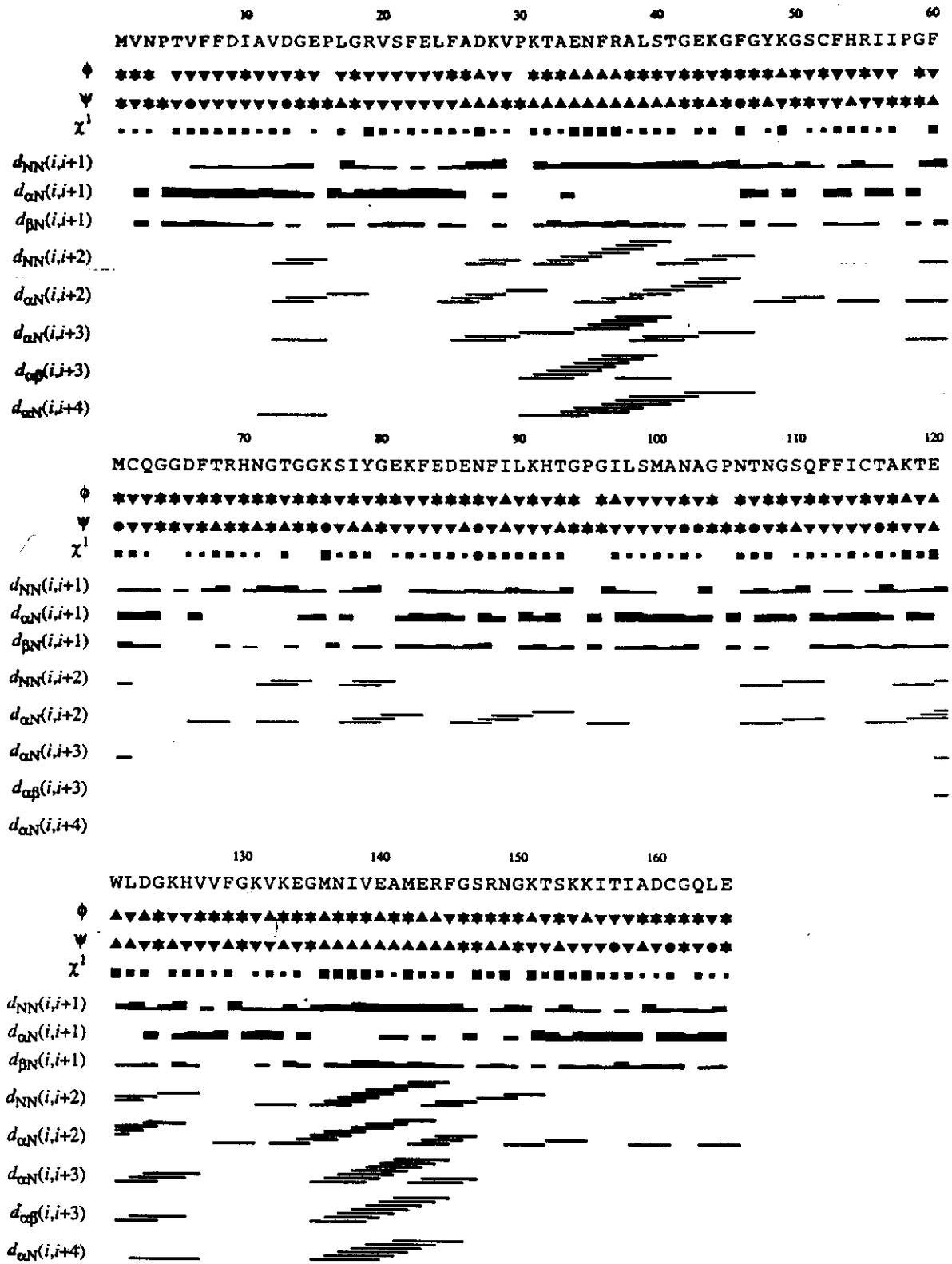
STRUCTURAL DATA: NOEs

Spin diffusion: $2\ \text{Å} \leq d \leq 5\ \text{Å}$

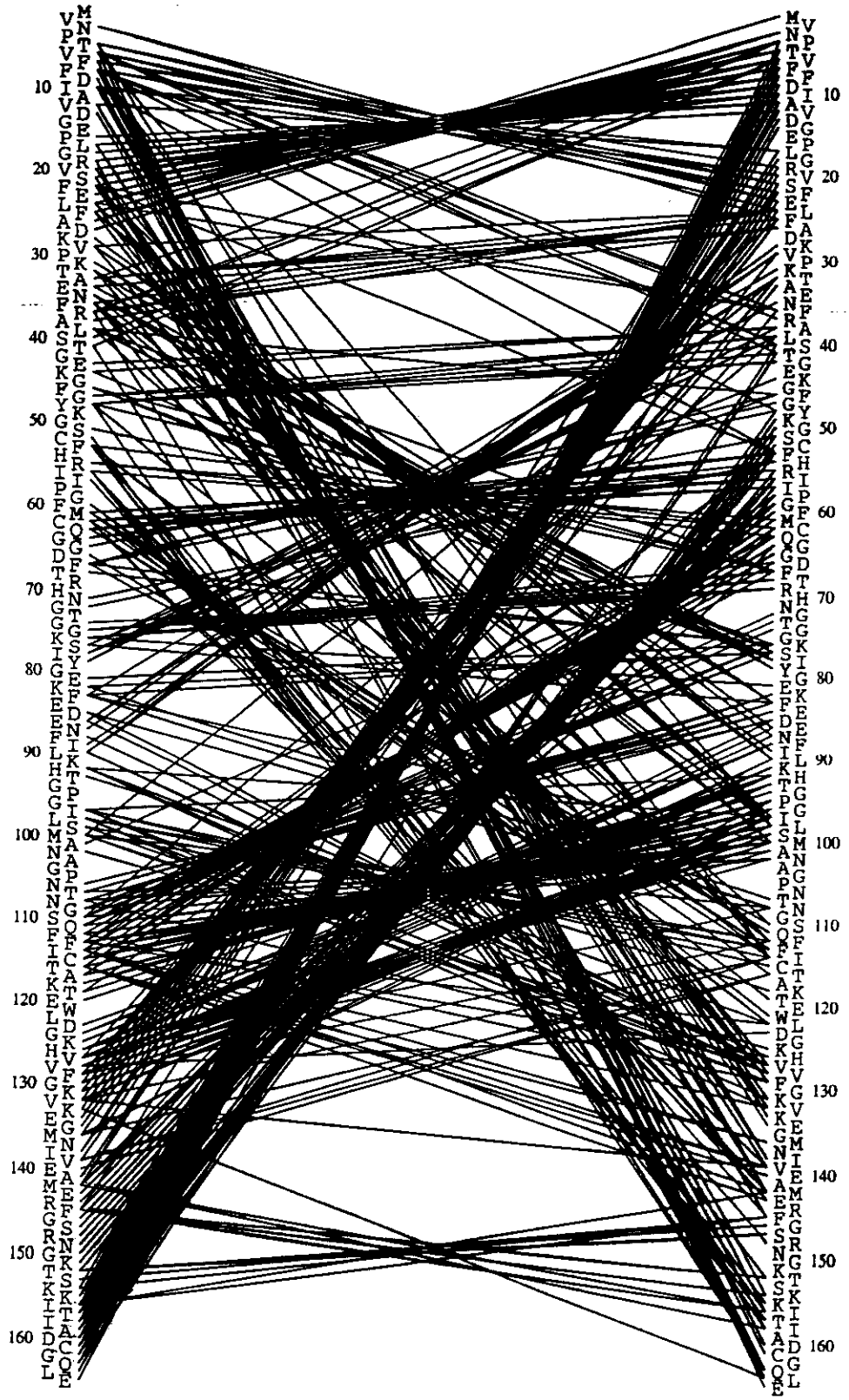
α-helix



Short- and medium-range NMR restraints for the protein cyclophilin A

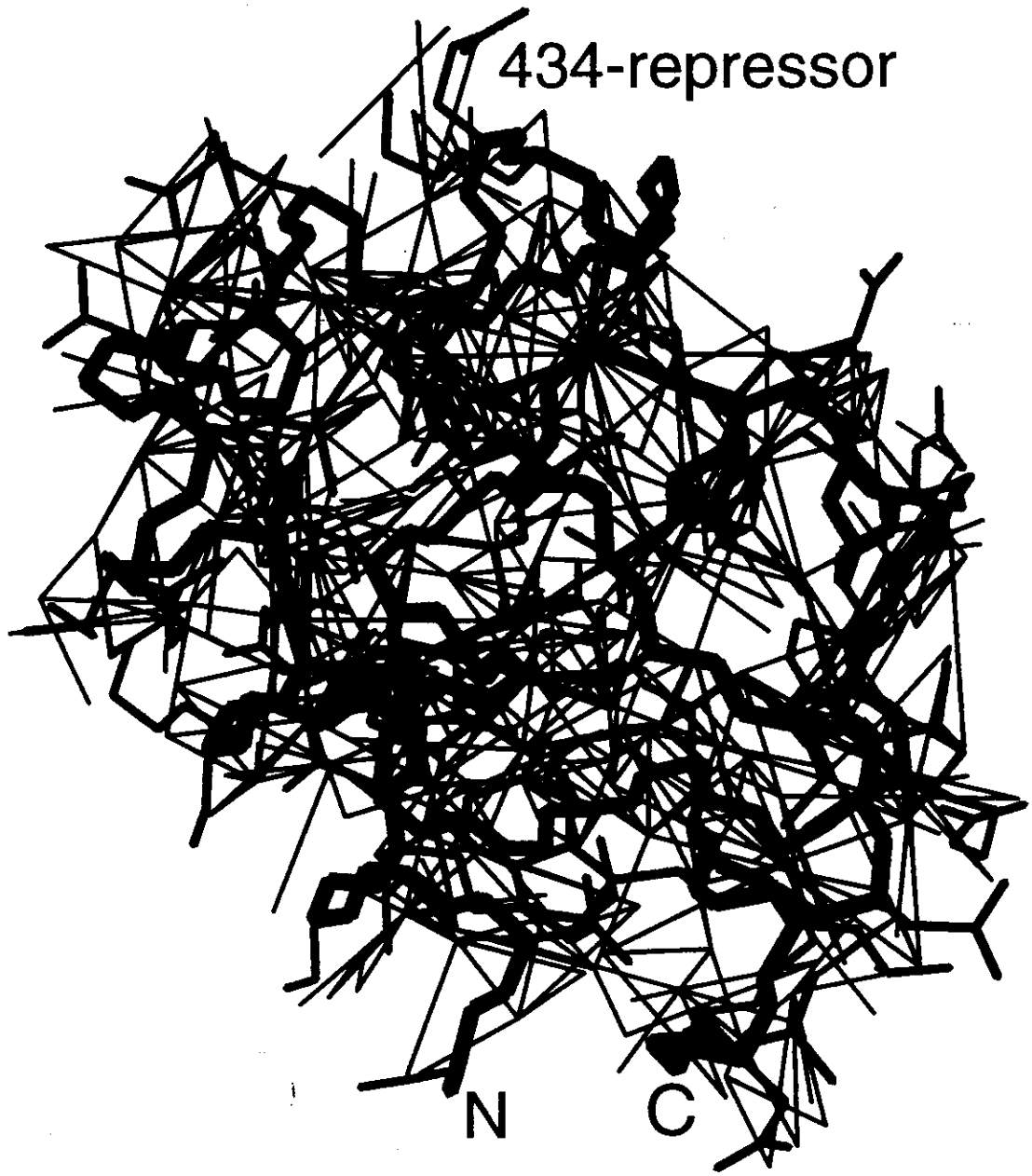


Medium- and long-range distance restraints derived from nuclear Overhauser effects (NOEs)



STRUCTURAL DATA: NOEs

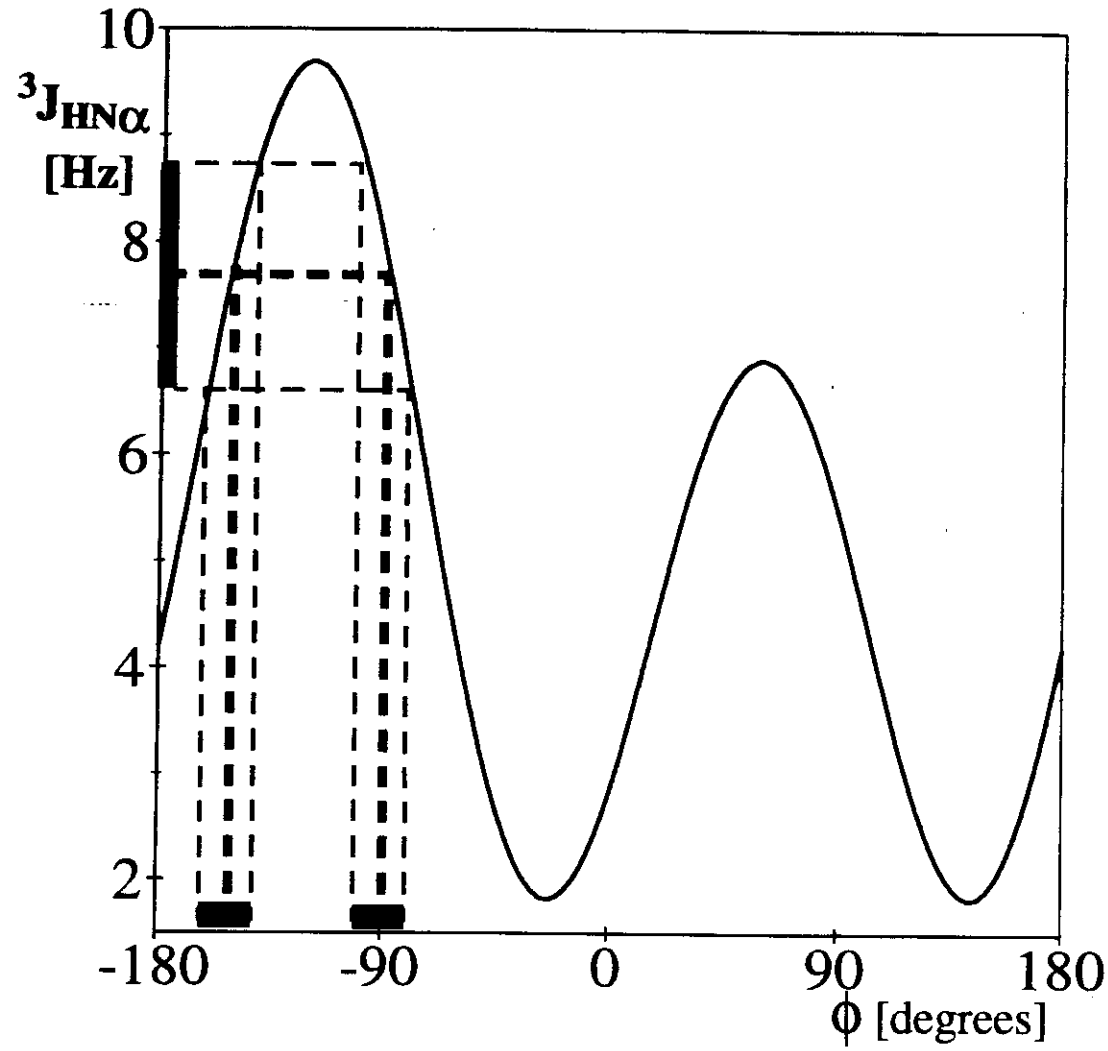
Density



STRUCTURAL DATA: J-COUPLING

Karplus-relation:

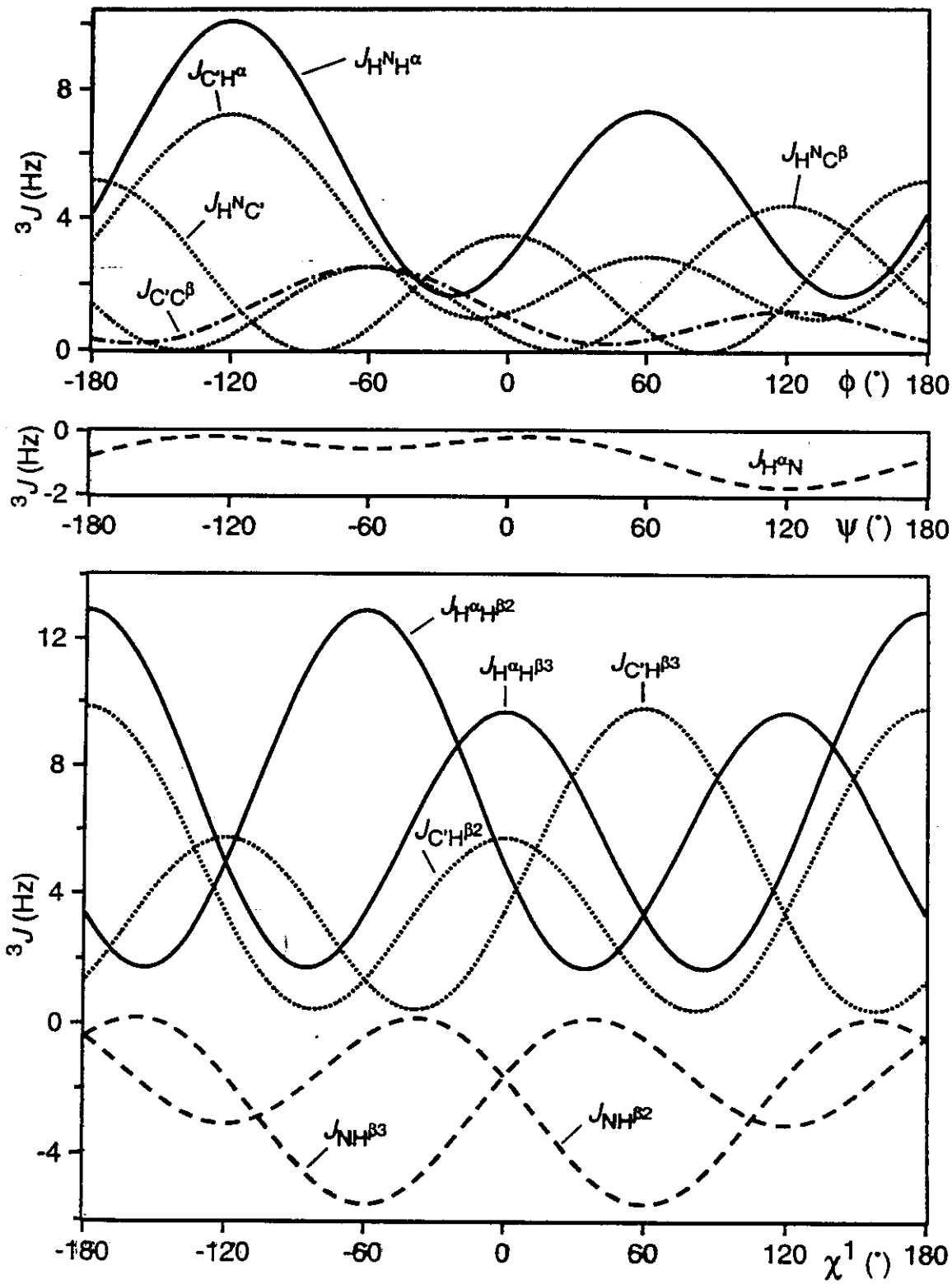
$$J(\theta) = A\cos^2\theta + B\cos\theta + C$$



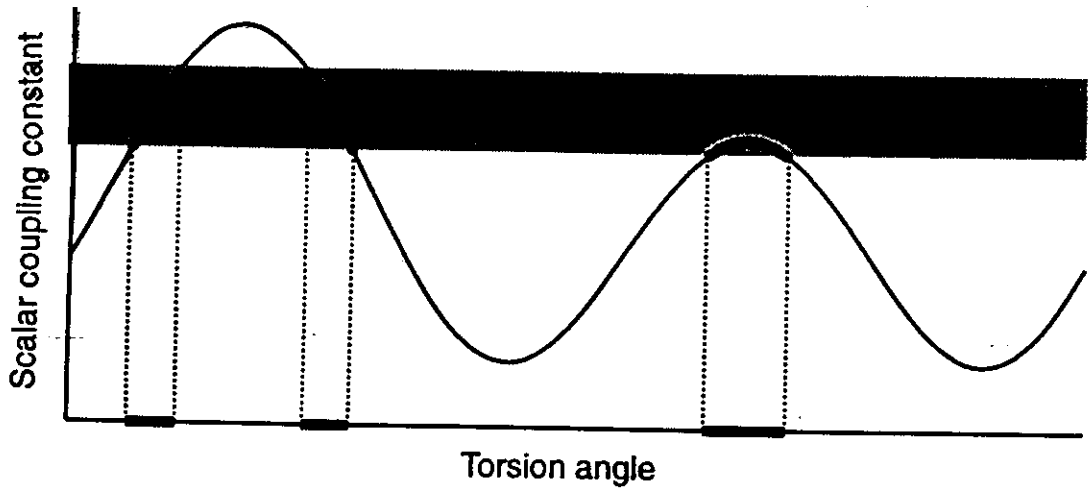
atoms coupled		*A*	*B*	*C*
HN	Hα	6.98	-1.38	1.72
HN	C'	4.32	0.84	0.00
HN	Cβ	3.39	-0.94	0.07
C'(i-1)	Hα	3.75	2.19	1.28
C'(i-1)	C'	1.33	-0.88	0.62
C'(i-1)	Cβ	1.59	-0.67	0.27

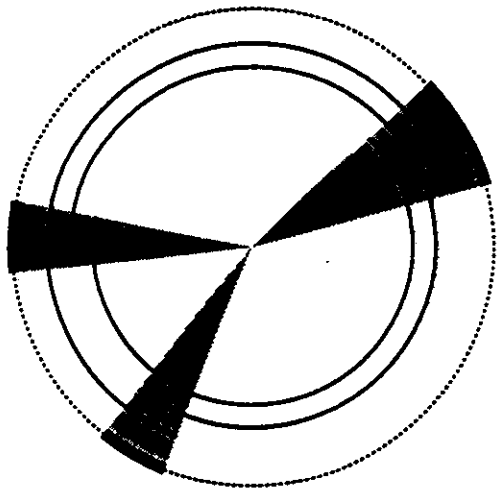
(Wang & Bax, 1996; Hu & Bax, 1997)

Karplus relations between ^3J scalar coupling constants and torsion angles

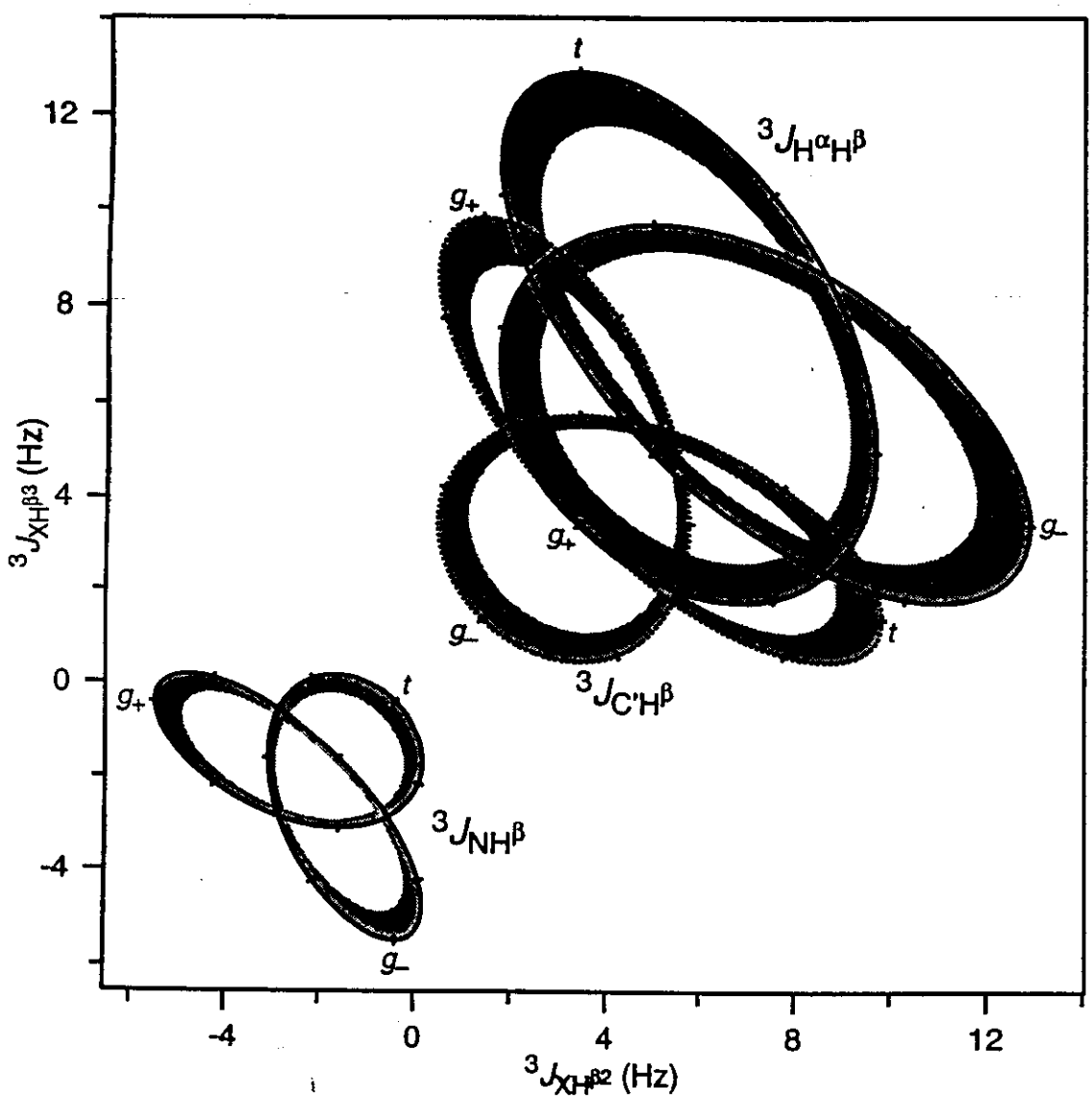


Torsion angle restraints from ^3J scalar coupling constants

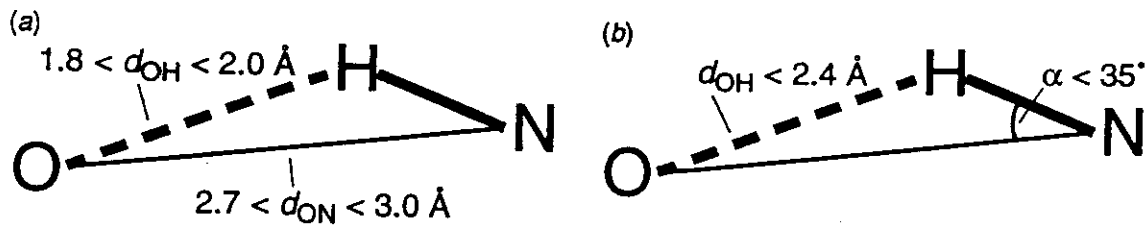




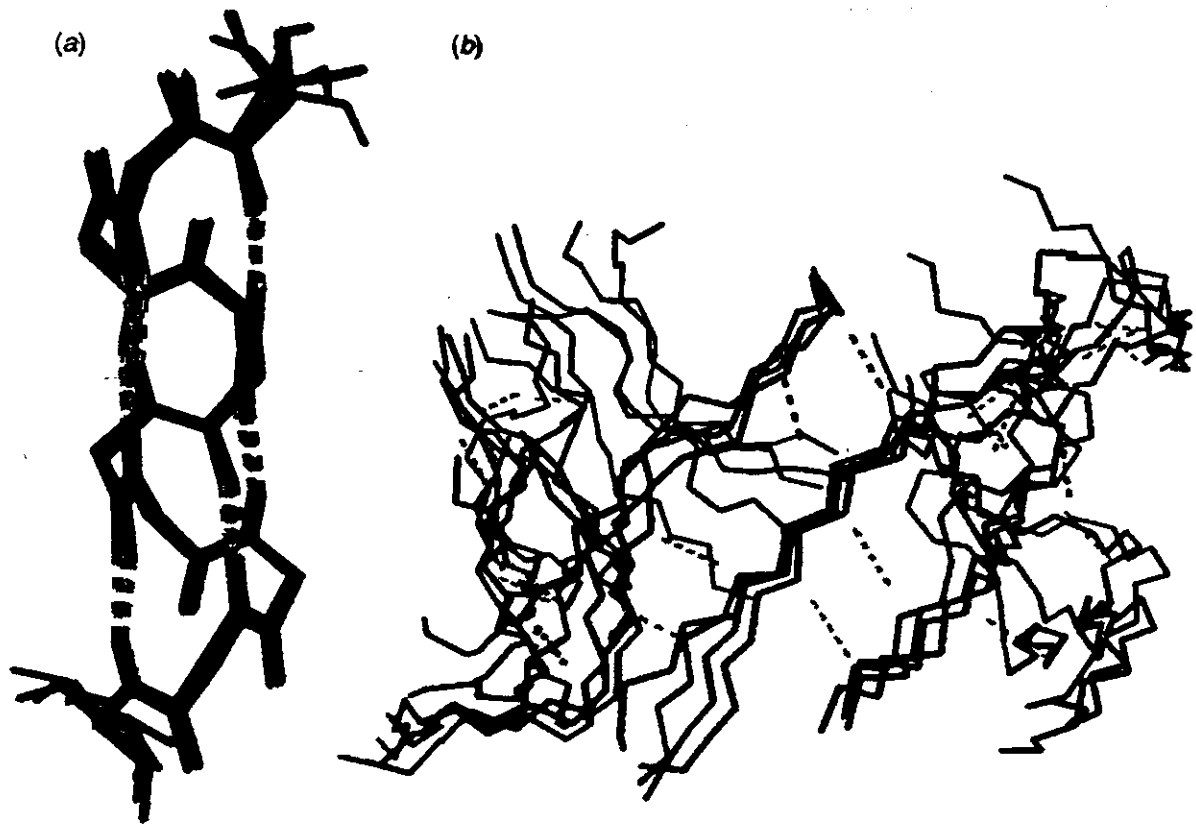
Correlations among scalar coupling constants for the torsion angle χ^1



Hydrogen bonds



Impact of hydrogen bond restraints on structure

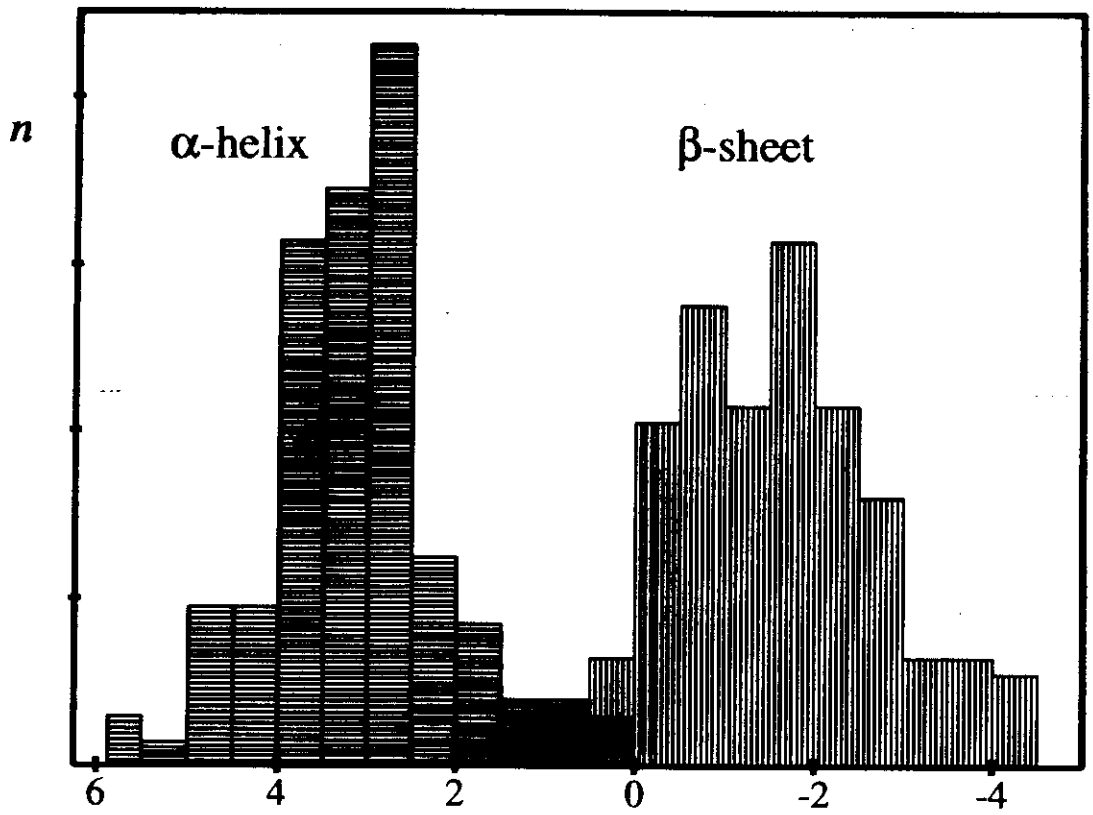


α-helix

β-barrel (cyclophilin A)

STRUCTURAL DATA: CHEMICAL SHIFTS

Chemical shifts of $^{13}C\alpha$

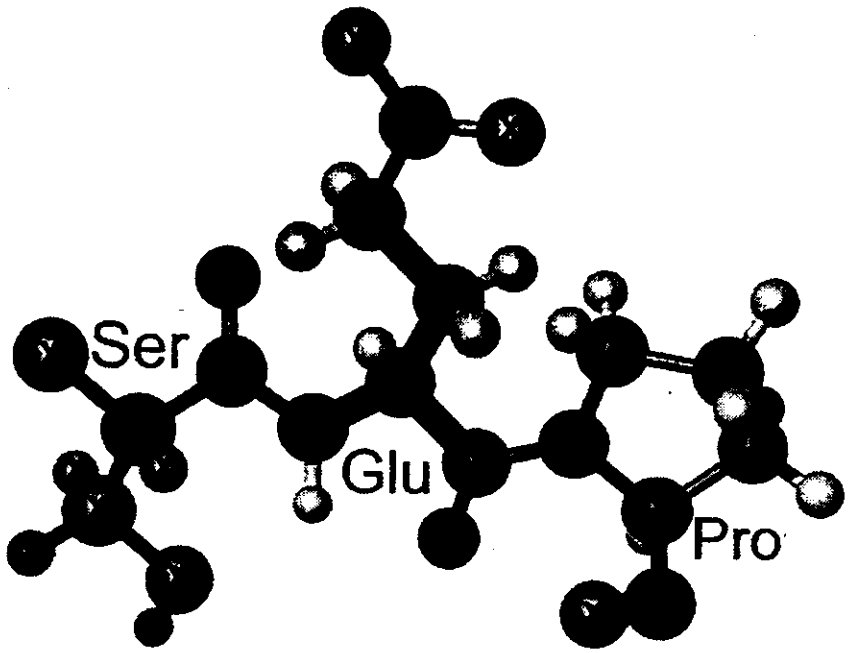


$\delta^{13}C\alpha > 2.0$ $\quad -120^{o} < \phi < -20^{o} \quad -100^{o} < \psi < 0^{o}$

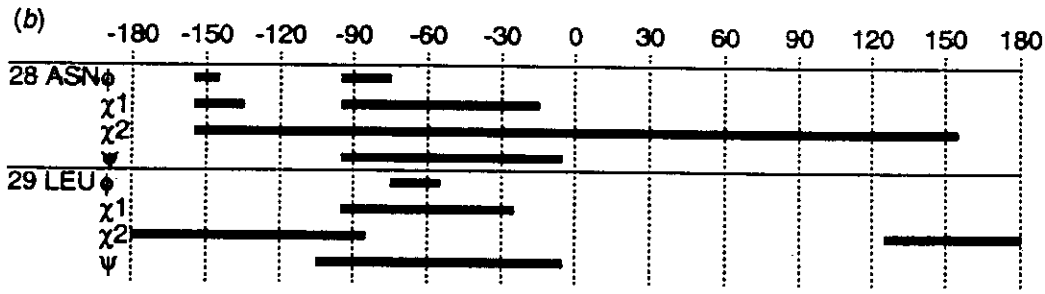
$\delta^{13}C\alpha < -1.5$ $\quad 160^{o} < \phi < 280^{o} \quad 40^{o} < \psi < 220^{o}$

Local conformation analysis

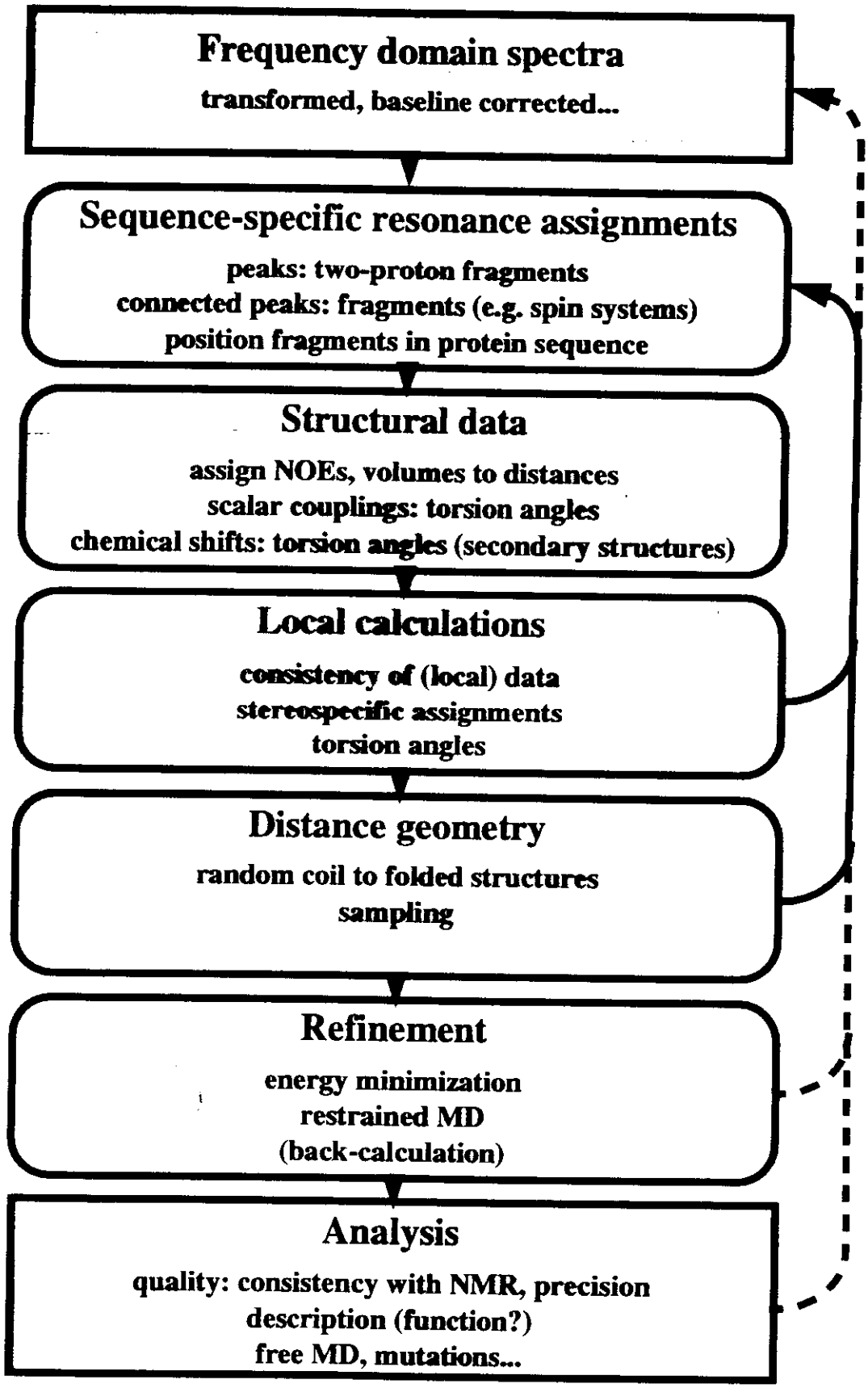
(a)



(b)

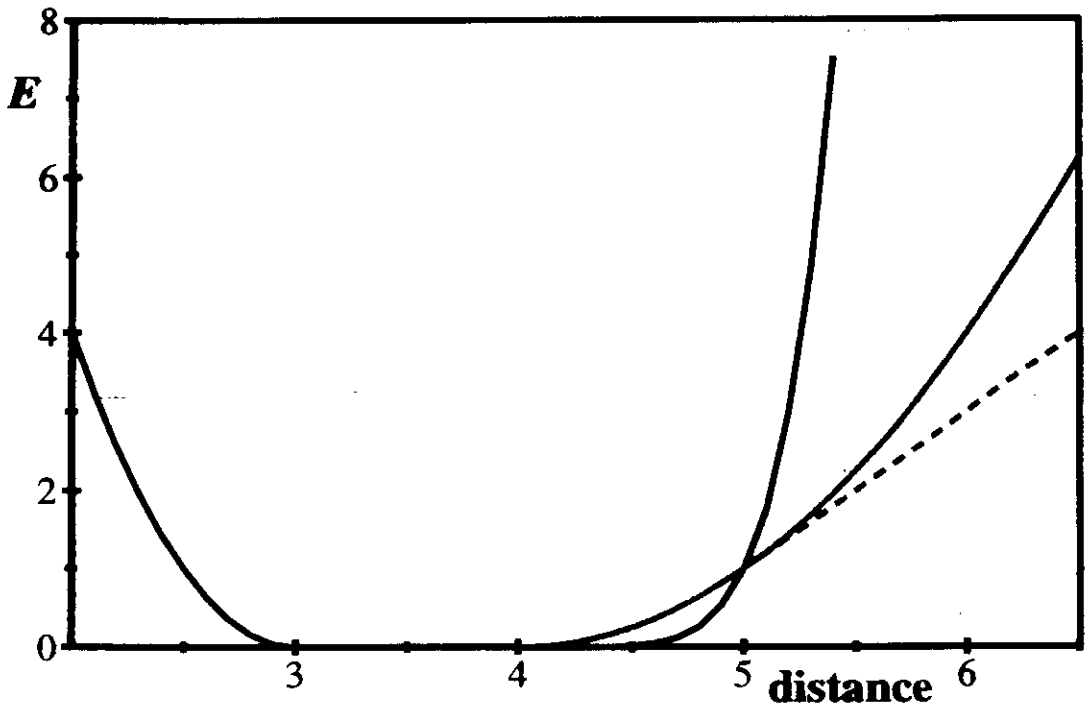


NMR Structure Determination

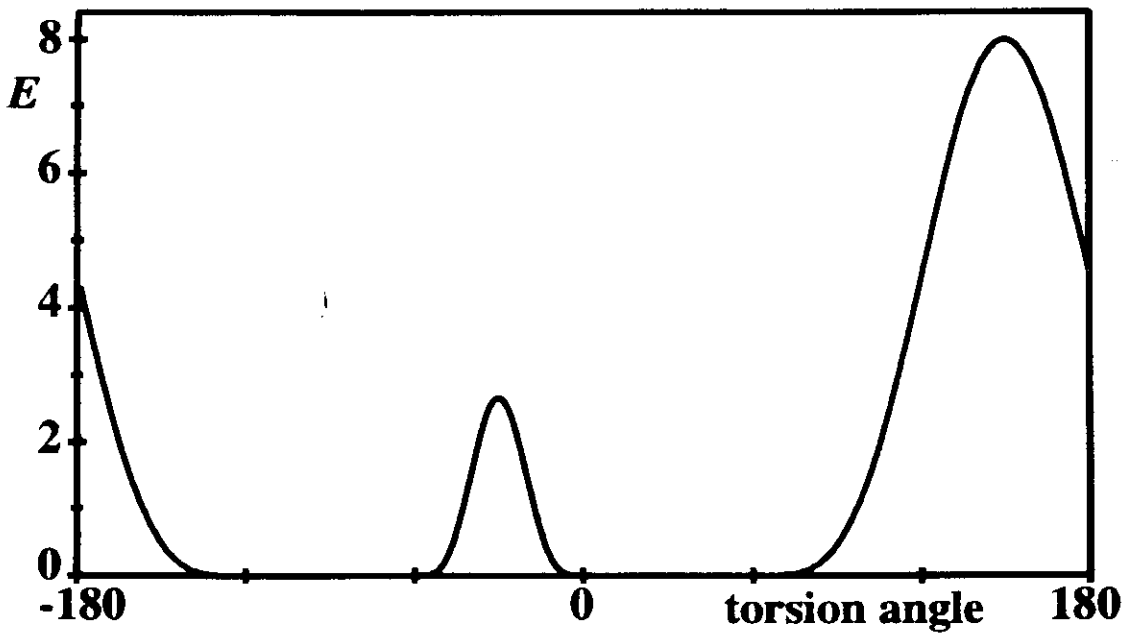


STRUCTURAL DATA: CONSTRAINTS

distance limits



angle constraints



STRUCTURE CALCULATION

(DISTANCE GEOMETRY)

Goal:

Determine 3D structures that are compatible with the NMR constraints and with general stereochemistry; no start structure is given.

Methods:

– metric matrix

– variable target function

– molecular dynamics / simulated annealing in Cartesian space

– molecular dynamics / simulated annealing in torsion angle space

– minimization (refinement)

STRUCTURE CALCULATION

Problem:

1) **find a structure that describes a globular protein consistent with the NMR**
2) **find a structure with minimal potential energy and vanishing violations of NMR constraints**
3) **minimize target function**

$$F = E_{pot} + N_{vio}$$

4) **find all structures with comparable low values of F**

The target function F
– is of high dimension (e.g. 3 × number of atoms)
– has many (local) minima

Random start structure: **find fold using mostly NMR data**

folded polypeptide: **refine using energy terms**

(or combine both steps)

TARGET FUNCTION

Goal: find minimum of target function F

$$F = E + N = \text{physical energy} + \text{NMR-constraints}$$

$$E = \sum K_i^b (b_i - b_i^o)^2 + \quad \text{bond lengths}$$

$$\sum K_i^\theta (\theta_i - \theta_i^o)^2 + \quad \text{bond angles}$$

$$\sum K_i^\omega (1 + \cos(\omega_i - \omega_i^o)) + \quad \text{torsion angles}$$

$$\sum (A_{ij}/r_{ij}^{12} - B_{ij}/r_{ij}^6) + \quad \text{Lennard-Jones}$$

$$\sum (q_i\, q_j) / (\varepsilon\, r_{ij}) \quad \text{Coulomb}$$

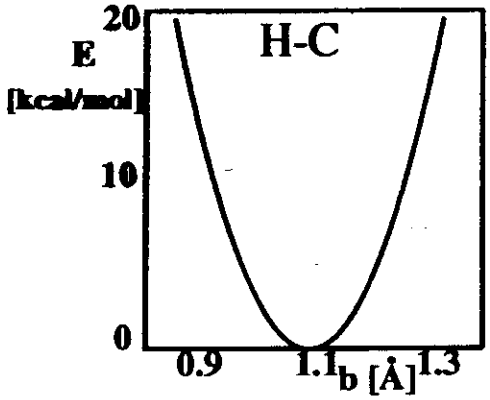
$$N = \sum U_i (d_i - d_i^o)^2 + \quad \text{distance constraints (if } d_l > d_l^o\text{)}$$

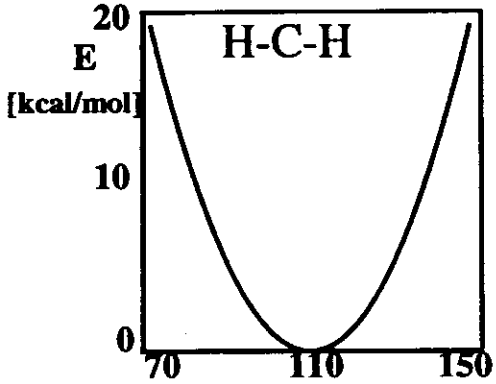
$$\sum W_i (1 + \cos(C_i\theta - \theta_i^o)) \quad \text{torsion angle constraints (if } \theta \text{ not allowed)}$$

b^o, θ^o, ω^o:	ideal values
i, j	summation indices
K, A, B, U, W:	force constants, weights

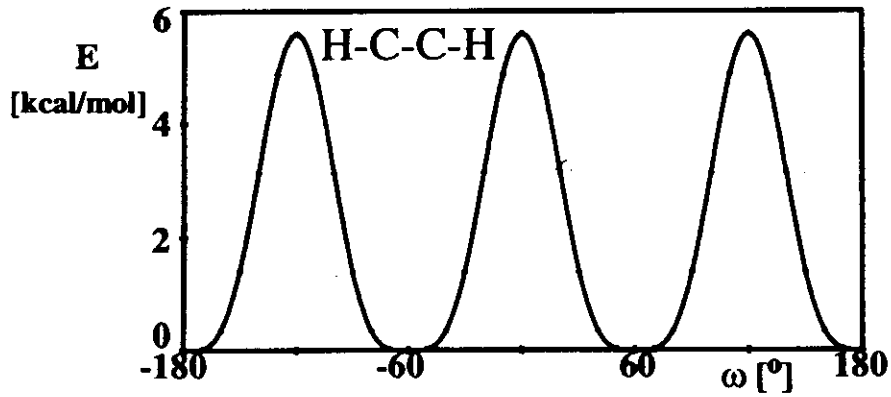
Energy Potentials

harmonic potential: bonds, bond angles

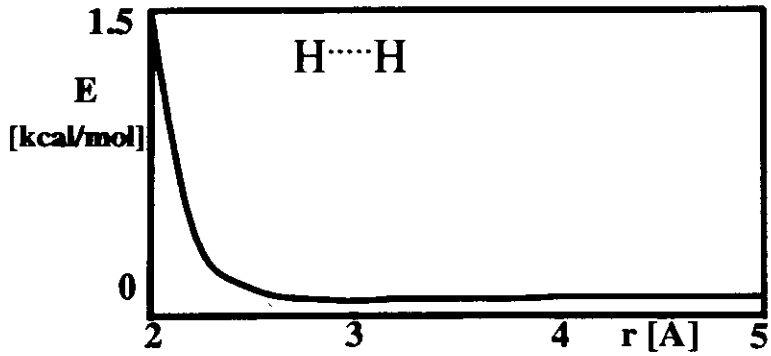




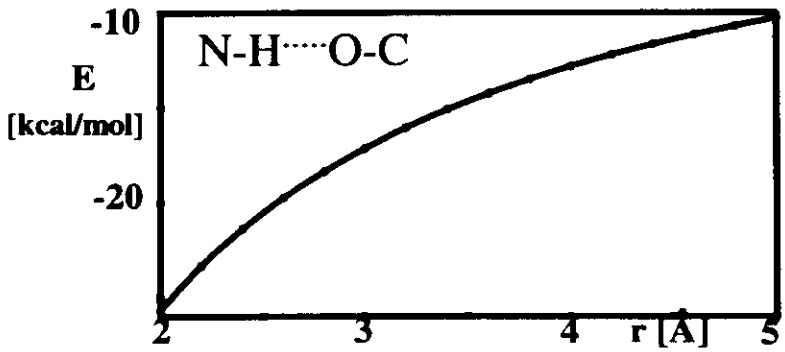
torsions angle



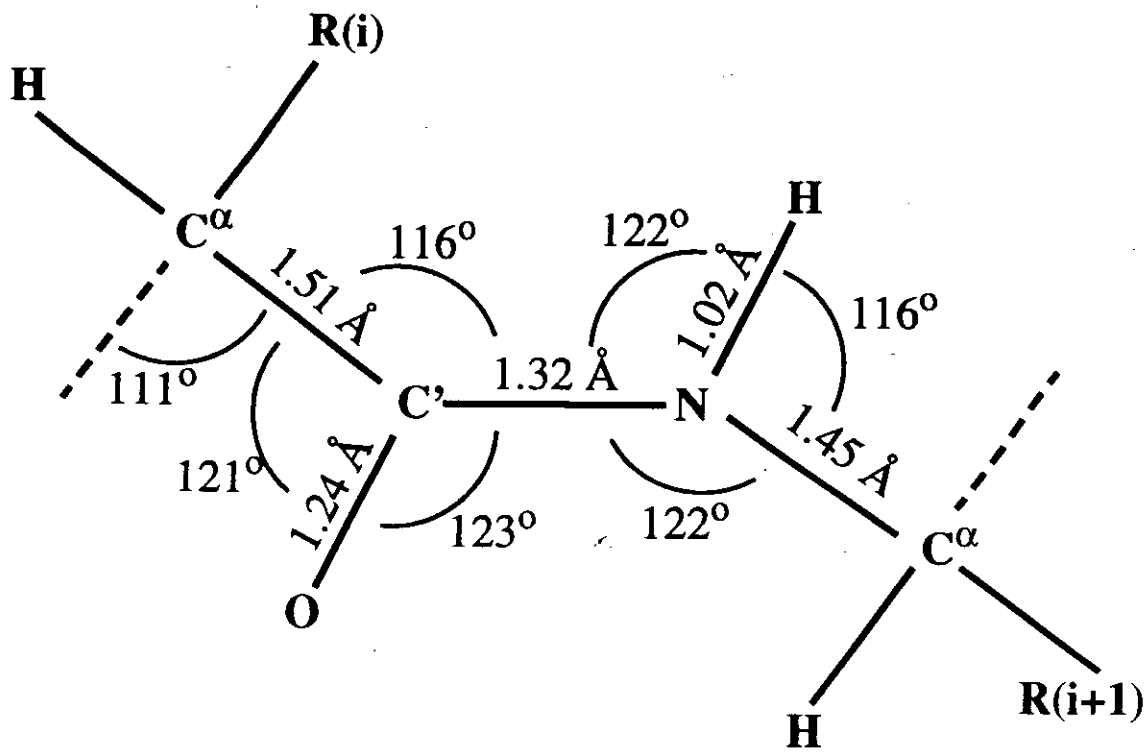
Lennard-Jones



Coulomb



DIMENSIONS IN A PROTEIN



Atom radii:

Atom	van der Waals (single bonds)	calculation, Ramachandran-plot
H	1.17	1.00
O	1.40	1.29
N	1.55	1.38
C	1.75	1.43
S	1.80	1.60

van der Waals distance = sum of two radii!

NMR Structure Calculation Methods

[1] Metric matrix distance geometry

distance matrix $D_{ij} = |\underline{x}_i - \underline{x}_j|$

↓

metric matrix $G_{ij} = \underline{x}_i \cdot \underline{x}_j$

↓

coordinates $\{\underline{x}_i\}$

[2] Simulated annealing by Cartesian space MD

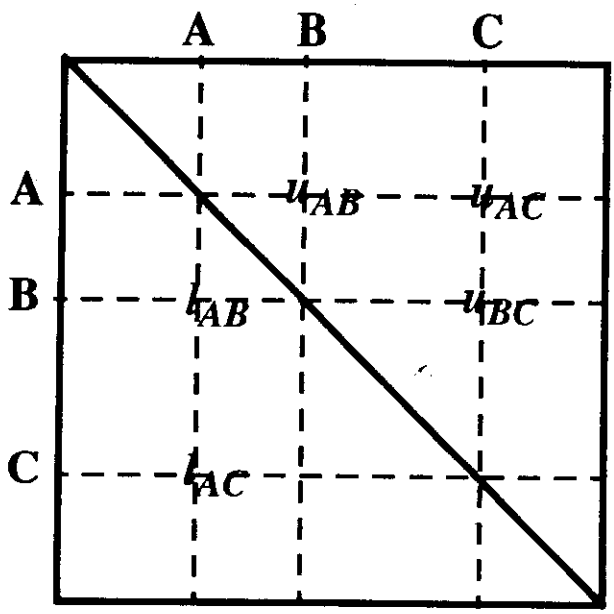
[3] Variable target function minimization in torsion angle space

- conjugate gradient minimization
- includes restraints in a stepwise fashion

METRIC MATRIX

Calculation in distance space, followed by embedding into Cartesian space

distance matrix (with atoms A, B, C)



upper limits:

$$d_{AB} \leq u_{AB}$$

$$d_{BC} \leq u_{BC}$$

$$d_{AC} \leq u_{AB} + u_{BC} \equiv u_{AC}$$

lower limits:

$$d_{AB} \geq l_{AB}$$

$$d_{AC} \geq l_{AB} - u_{BC} \equiv l_{AC}$$

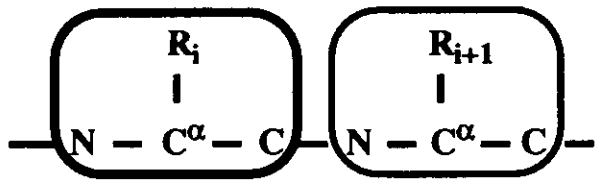
VARIABLE TARGET FUNCTION

(PROGRAM DIANA)

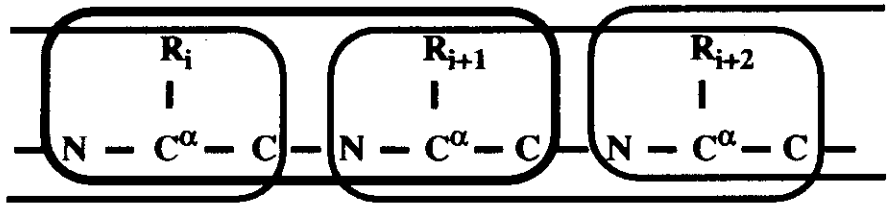
Torsion angle space

Level **Range of active constraints**

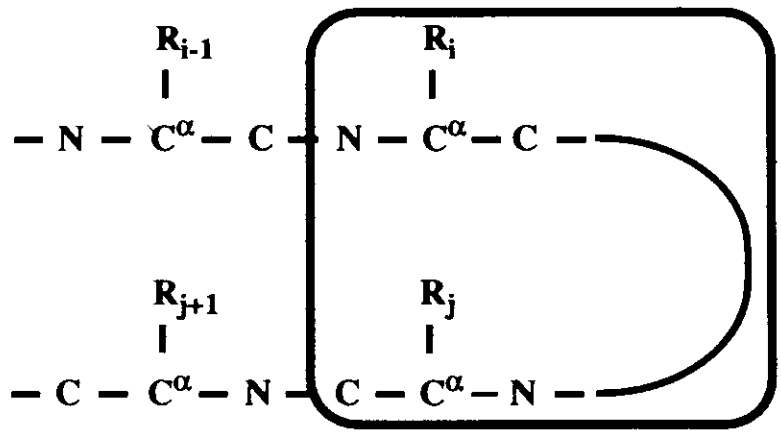
0

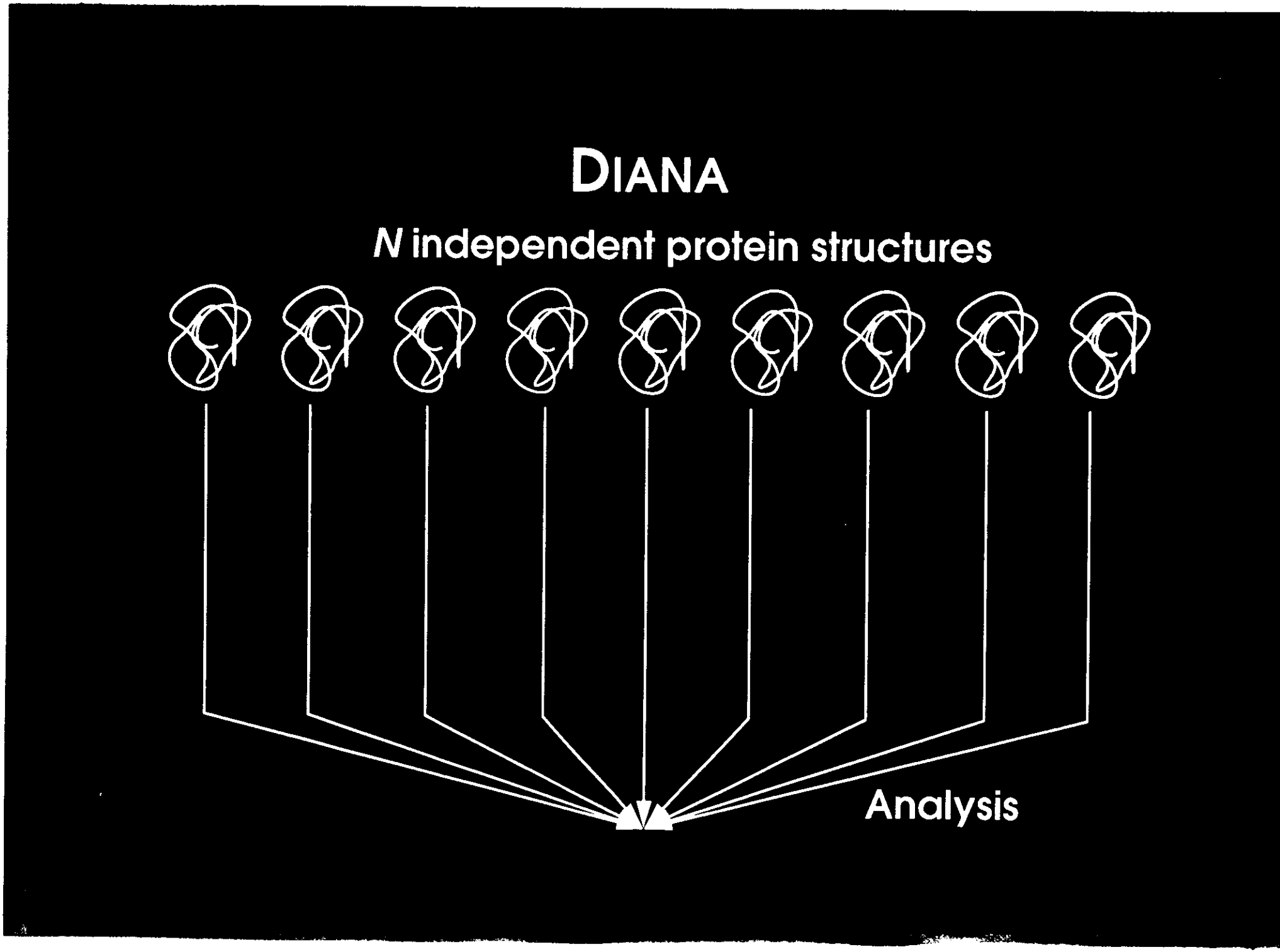


1



j - i




DIANA
N independent protein structures
Analysis

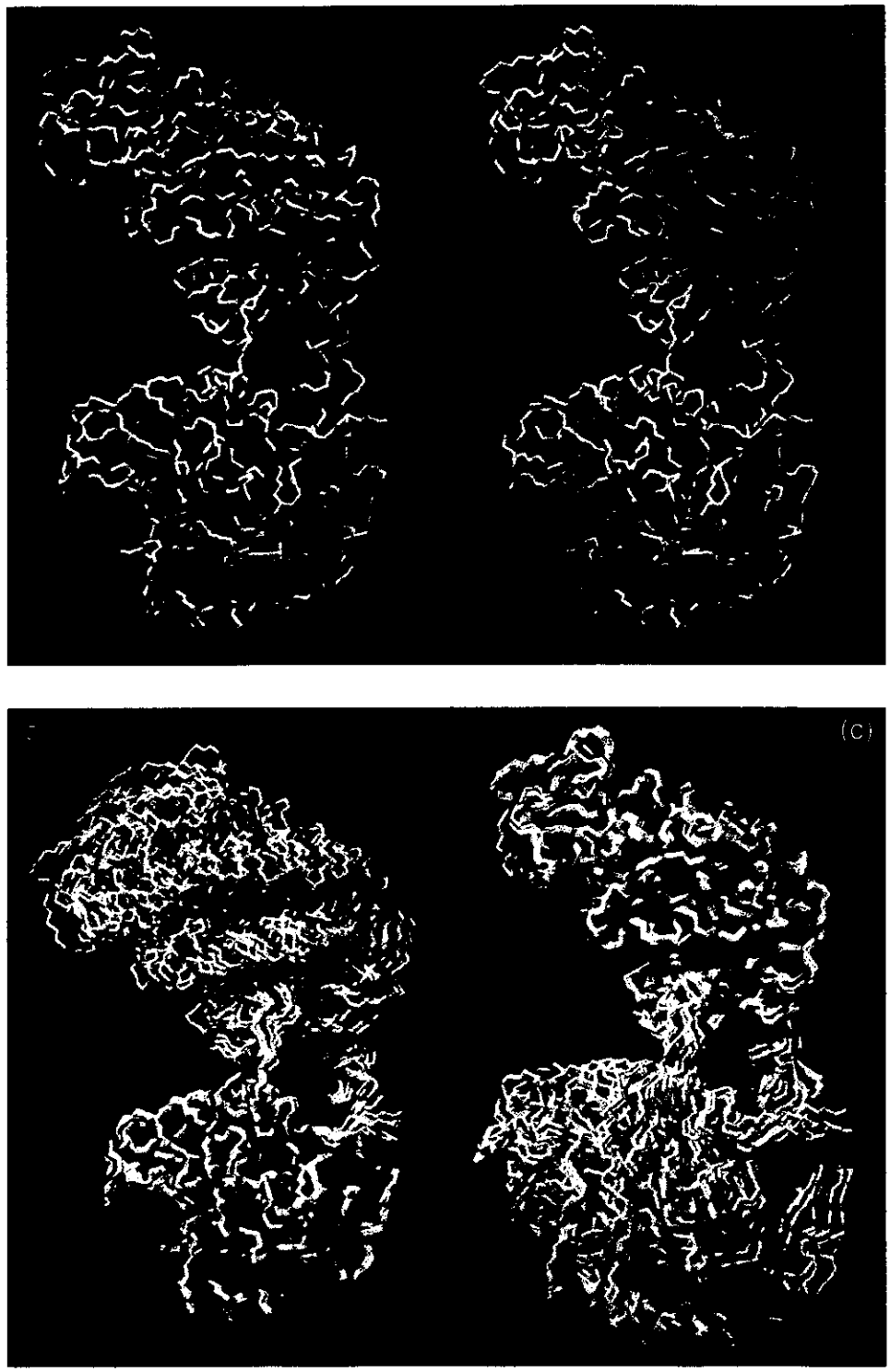


Fig. 7

DYANA

Torsion Angle Dynamics for NMR Structure Calculation and Automatic Assignment of NOESY Spectra

P. Güntert *C. Mumenthaler* *K. Wüthrich*

OPA

Molecular dynamics

Trajectory

Numerical integration of Newton's equations

Equations of motion

Cartesian coordinates: $x_1, \ldots, x_n$

$$m\ddot{x} = \frac{\partial L}{\partial x} \qquad \text{(Newton)}$$

Generalized coordinates: $q_1, \ldots, q_m$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0 \qquad \text{(Lagrange)}$$

with $L = E_{kin} - E_{pot}$

Integration of the equations of motion

e.g. "leap-frog" algorithm

$$q(t + \Delta t) = q(t) + \Delta t\, \dot{q}(t) + \tfrac{1}{2}\Delta t^2\, \ddot{q}(t) + O(\Delta t^3)$$

$$\dot{q}(t + \Delta t) = \dot{q}(t) + \tfrac{1}{2}\Delta t\, \ddot{q}(t) + \tfrac{1}{2}\Delta t\, \ddot{q}(t + \Delta t) + O(\Delta t^3)$$

q coordinates (Cartesian or torsional)

$\dot{q} = \dfrac{dq}{dt}$ velocities

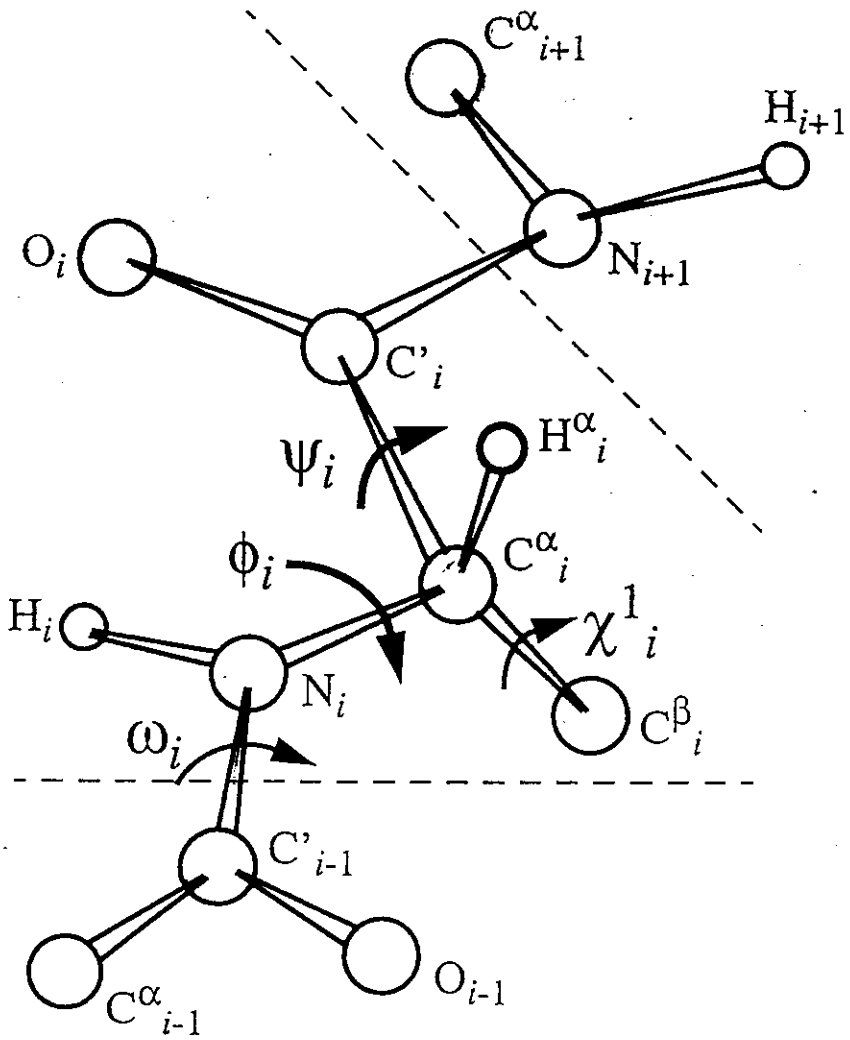
$\ddot{q} = \dfrac{d^2q}{dt^2}$ accelerations

Δt time step

Torsion Angle Dynamics (TAD)

- Torsion angle dynamics = molecular dynamics (MD) in torsion angle space.
- Classical mechanical equations of motion are solved for a system with N torsion angles as the only degrees of freedom.
- About 10 times less degrees of freedom than in conventional Cartesian space MD.
- Fixed bond lengths and bond angles:
  → no high frequency motions
  → longer integration time-steps, higher annealing temperatures

DIHEDRALWINKEL IN PROTEINEN 1



Winkel	Bindung	Atome
ω_i	$C'_{i-1} - N_i$	$C^{\alpha}_{i-1} - C'_{i-1} - N_i - C^{\alpha}_i$
ϕ_i	$N_i - C^{\alpha}_i$	$C'_{i-1} - N_i - C^{\alpha}_i - C'_i$
ψ_i	$C^{\alpha}_i - C'_i$	$N_i - C^{\alpha}_i - C'_i - N_{i+1}$
χ^1_i	$C^{\alpha}_i - C^{\beta}_i$	$N_i - C^{\alpha}_i - C^{\beta}_i - C^{\gamma}_i$

Molecular Dynamics

Cartesian space		Torsion angle space
$3N$ atom coordinates $x_1, \ldots, x_{3N}$	*Degrees of freedom*	n torsion angles $\theta_1, \ldots, \theta_n$
Newton's equations $m_i \ddot{x}_i = f_i$	*Equations of motion*	Lagrange equations $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0$ $L = T - V$
$E_{kin} = \frac{1}{2}\sum_{i=1}^{3N} m_i \dot{x}_i^2$	*Kinetic energy*	$T = \frac{1}{2}\sum_{i,j=1}^{n} M_{ij}(\theta)\dot{\theta}_i\dot{\theta}_j$
diagonal, constant (elements m_i)	*Mass matrix M*	non-diagonal, non-constant, $n \times n$
$\ddot{x}_i = \frac{1}{m_i}\frac{\partial V}{\partial x_i}$	*Accelerations*	$M(\theta)\ddot{\theta} = C(\theta, \dot{\theta})$ (n linear equations)
proportional to N	*Computational complexity*	solving linear system of equations: $\sim n^3$ **exploiting tree structure of molecule: $\sim n$** (Jain et al., *J. Comp. Phys.* 106, *p* 258, 1993)

JOURNAL OF COMPUTATIONAL PHYSICS 106, 258–268 (1993)

A Fast Recursive Algorithm for Molecular Dynamics Simulation

A. JAIN

Jet Propulsion Laboratory/California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109

N. VAIDEHI

A. A. Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California 91125

AND

G. RODRIGUEZ

Jet Propulsion Laboratory/California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109

Received December 16, 1991

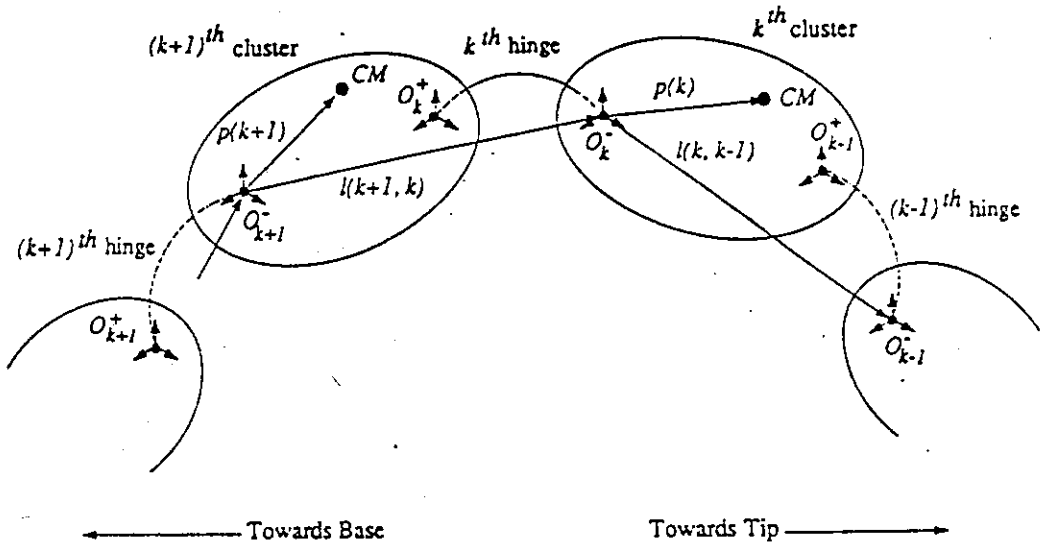
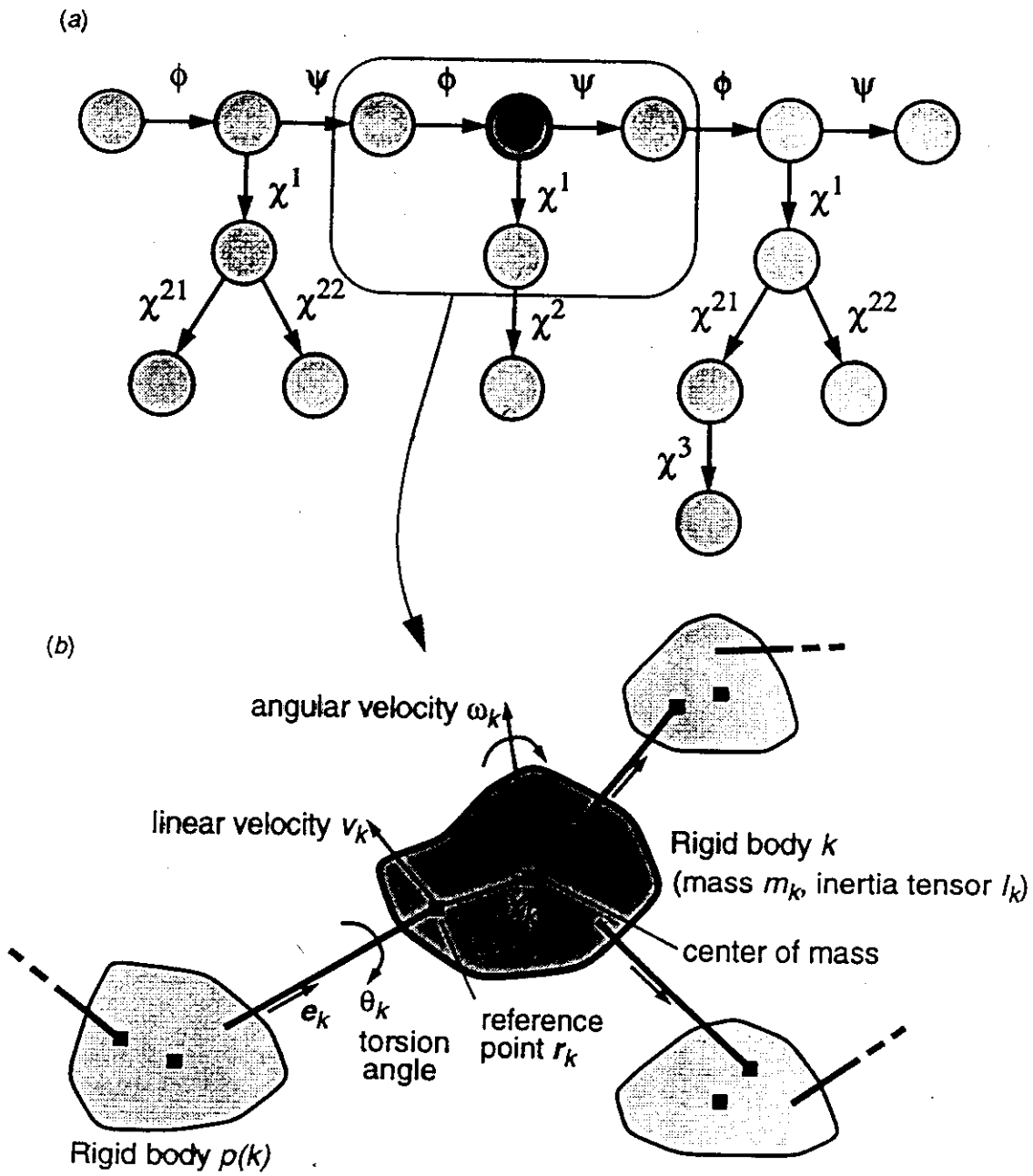


FIG. 2. A molecular model consisting of atomic clusters coupled together in a serial chain.

Torsion angle dynamics:
Tree structure of the molecule



Kinetic energy

For all rigid bodies, $k = 1, \ldots, n$, the angular velocity vector, $\vec{\omega}_k$, and the linear velocity of the reference point, $\vec{v}_k = \vec{v}_{\ldots}$, are calculated recursively (Jain *et al.*, 1993):

$$\vec{\omega}_{\ldots} = \ldots$$

$$\vec{v}_k = \ldots + \vec{\omega}_{\ldots}$$

$\vec{Y}_k$ vector from reference point to center of mass

m_k total mass of rigid body *k*

I_k inertia tensor of rigid body *k*

Kinetic energy:

$$E_{\text{kin}} = \frac{1}{2} \sum_{1}^{n} \left[m_k \vec{v}_k^2 + \vec{\omega}_k \cdot \boldsymbol{I}_k \vec{\omega}_k + \vec{v}_k \cdot (\vec{\omega}_k \times m_k \vec{Y}_k) \right]$$

Potential energy = target function

For torsion angle dynamics in DYANA the role of the potential energy, E_{pot}, is taken by the DIANA target function, V:

$$V = \sum_{\alpha,\beta} w_{\alpha\beta} \sum_{(\alpha,\beta)} f_c(d_{\alpha\beta}, b_{\alpha\beta})$$

$$+ w \sum_k \left(1 - \frac{1}{2}\left(\frac{\Delta_k}{\Gamma_k}\right)^2\right)\Delta_k^2$$

$w_{\alpha\beta}$ weights

$b_{\alpha\beta}$ upper/lower bound on distance $d_{\alpha\beta}$

Γ_k half-width of forbidden torsion angle interval

Δ_k size of the torsion angle constraint violation

Different choices for the function $f_c(d, b)$ that measures the contribution of a violated distance constraint to the target function.

$$f_c(d, b) = \left(\frac{d^2 - b^2}{2b}\right)^2$$

corresponds to the definition used in DIANA,

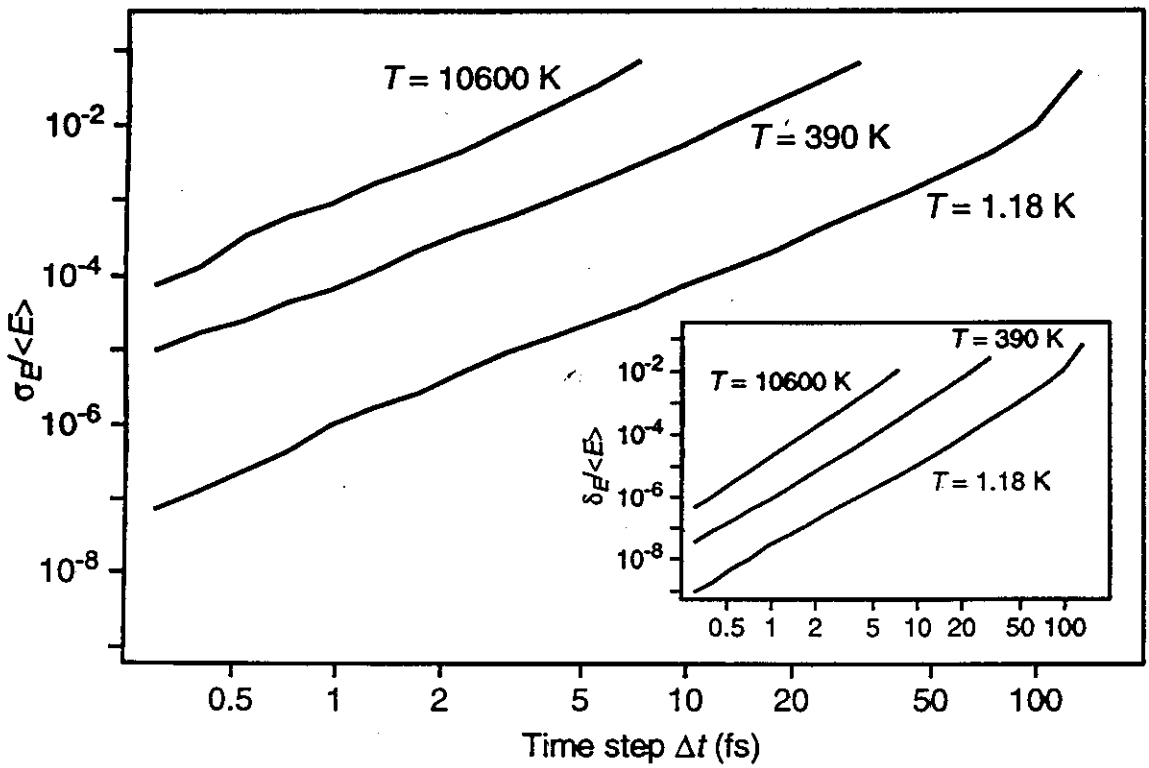
$$f_c(d, b) = (d - b)^2$$

is a simple square potential, and

$$f_c(d, b) = 2\beta^2 b^2 \left[\sqrt{1 + \left(\frac{d - b}{\beta b}\right)^2} - 1\right]$$

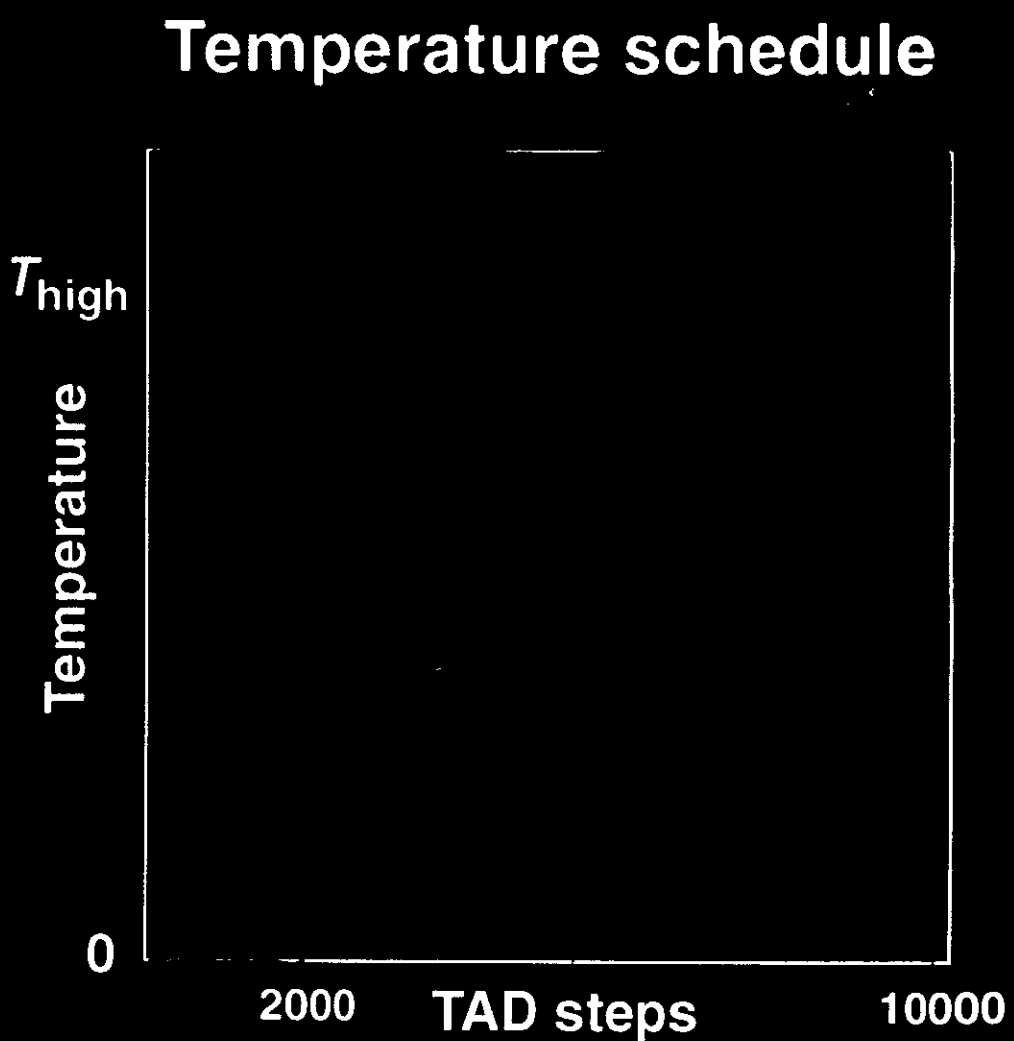
is a function with a linear asymptote for large constraint violations. β is a dimensionless parameter that weights large violations relative to small ones.

Torsion angle dynamics: Accuracy of energy conservation



Simulated Annealing using TAD

- Start from random structure
- Use all constraints simultaneously (no variable target function)
- Van der Waals check only for heavy atoms (steric repulsion of hydrogens is activated only at the end of the calculation)
- User-defined parameters:
  – start temperature, T_{high}
  – number of TAD steps



Temperature control

Weak coupling to a heat bath is used to control the temperature:

$$\theta \leftarrow \theta \sqrt{1 + \frac{T^{\mathrm{ref}} - T}{\tau T}}$$

θ torsional velocities
T instantaneous temperature, $T = \frac{2E_{\mathrm{kin}}}{nk_B}$
τ coupling constant

(Berendsen et al., *J. Chem. Phys.* 81, 3684–3690, 1984)

Time-step adaption

The same idea is used to adapt the integration time-step, Δt, such as to maintain a user-defined accuracy of energy conservation:

$$\Delta t \leftarrow \Delta t \sqrt{1 + \frac{\varepsilon^{\mathrm{ref}} - \varepsilon}{\tau\varepsilon}}$$

ε is the relative change of the total energy, $E = E_{\mathrm{kin}} + E_{\mathrm{pot}}$:

$$\varepsilon = \left| \frac{E(t - \Delta t) - E(t)}{E(t)} \right|$$

Standard protocol: $\varepsilon^{\mathrm{ref}} = 0.005 \ldots 0.0001$, $\tau = 20$

Torsion angle dynamics: Simulated annealing

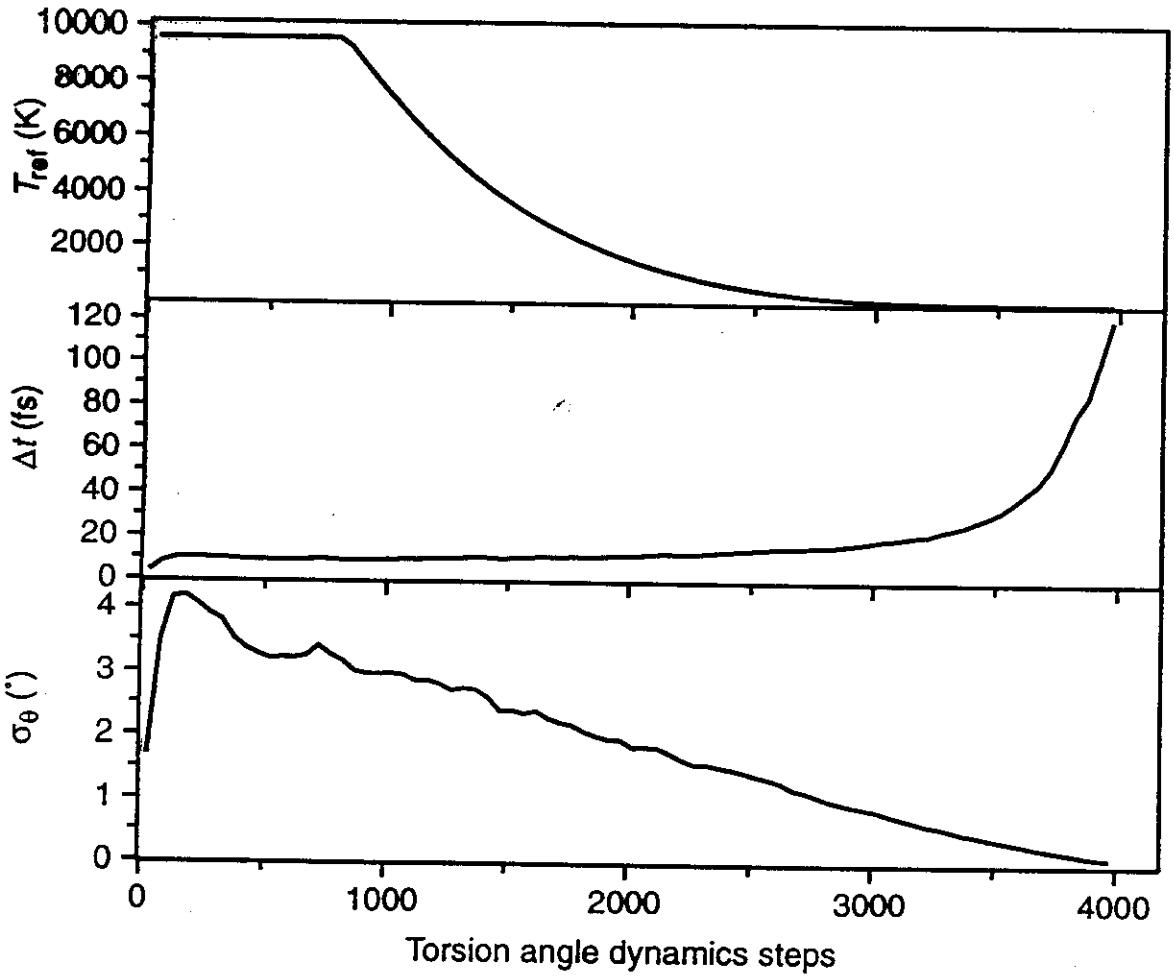
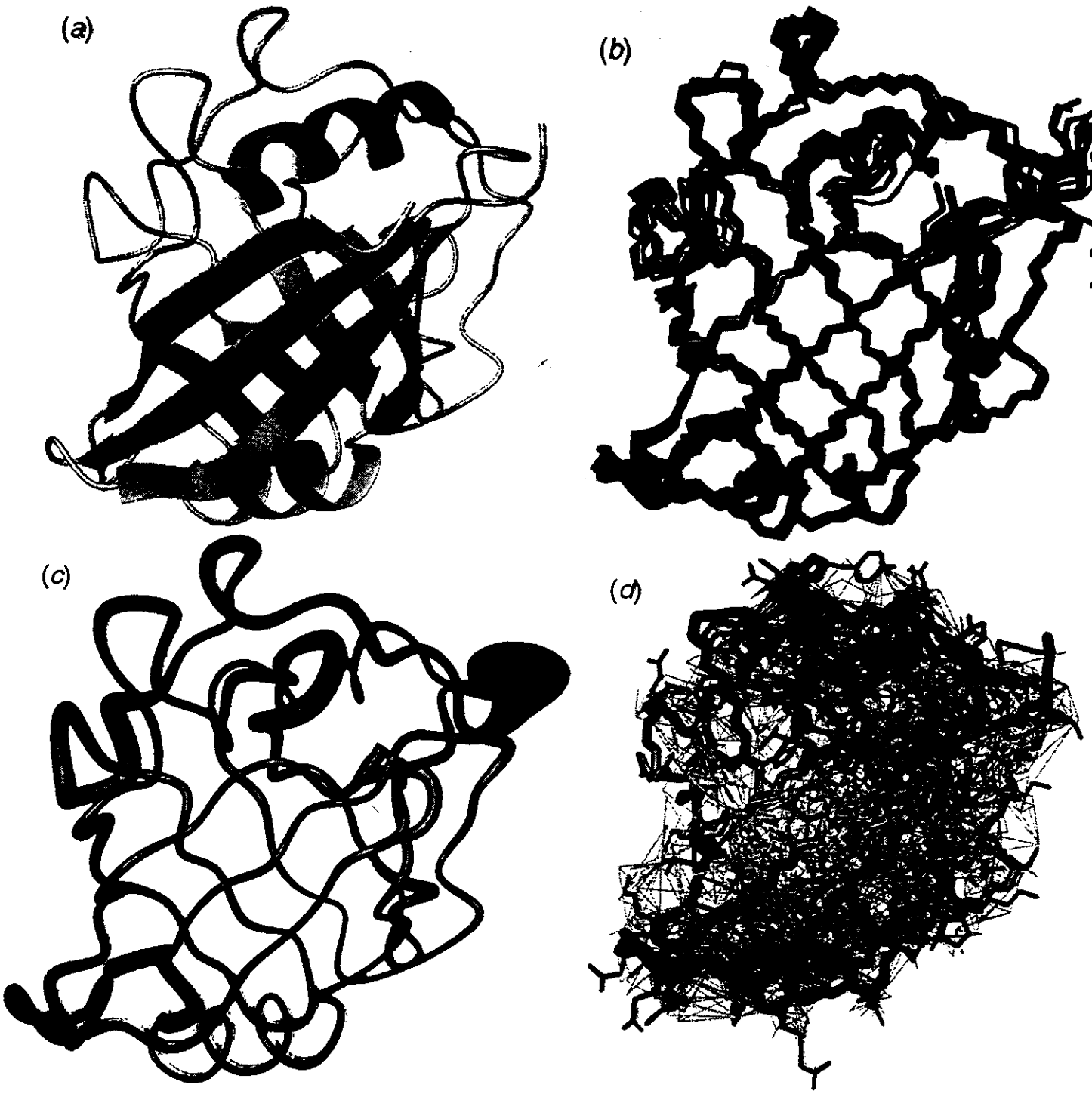


Table 6. *Structure calculations using the torsion angle dynamics algorithm of the program DYANA*[a]

Quantity[b]	BPTI	Antp	ADR	PrP	Cyp	PGK	APK
Size and experimental input data:							
Residues	58	68	81	113	165	394	32
Torsion angles	241	351	379	521	722	1778	318
Upper distance bounds	651	894	795	1508	4093	14161	929
Torsion angle restraints	115	171	168	256	371	1117	284
Restraints/torsion angle	3.2	3.0	2.5	3.4	6.2	8.6	3.8
Structure calculation:							
Accepted conformers (%)	88	75	88	47	58	63	45
Computation time (s)[c]	23	36	37	101	159	521	72
Target function ($Å^2$)	0.33	0.94	0.60	2.19	2.19	3.99	0.32
Backbone RMSD (Å)[d]	0.37	0.38	0.92	1.11	0.68	1.62	3.08
Average or sum[e] of restraint violations:							
Upper bounds (Å)	0.004	0.005	0.003	0.006	0.003	0.001	0.001
Steric lower bounds (Å)	1.2	2.5	1.9	5.2	5.1	11.0	0.8
Torsion angle restraints (°)	0.04	0.11	0.04	0.10	0.03	0.01	0.03
Maximal restraint violations:							
Upper bounds (Å)	0.19	0.28	0.25	0.41	0.44	0.47	0.22
Steric lower bounds (Å)	0.09	0.17	0.20	0.25	0.23	0.39	0.13
Torsion angle restraints (°)	1.6	4.5	2.7	5.3	3.7	6.9	1.4

Cyclophilin A



Overview of restraint violations

Overview:

Number of structures : 20
Cutoff for upper limits : 0.20 A
lower limits : 0.20 A
van der Waals : 0.20 A
angle constraints : 5.00 deg

struct	target function	upper limits #	upper limits sum	upper limits max	lower limits #	lower limits sum	lower limits max	van der Waals #	van der Waals sum	van der Waals max	torsion angles #	torsion angles sum	torsion angles max
1	0.94	1	6.6	0.35	0	0.0	0.00	0	2.8	0.15	0	2.2	1.7
2	0.98	2	7.3	0.25	0	0.0	0.00	0	2.6	0.16	0	3.8	2.4
3	1.03	4	8.6	0.27	0	0.0	0.00	0	3.2	0.12	0	5.1	1.5
4	1.08	3	8.1	0.34	0	0.0	0.00	0	3.4	0.18	0	4.6	1.5
5	1.13	2	9.4	0.24	0	0.0	0.00	0	3.5	0.12	0	5.5	2.2
6	1.17	1	6.2	0.22	0	0.0	0.00	2	3.3	0.34	0	6.9	4.2
7	1.21	2	7.1	0.26	0	0.0	0.00	1	3.2	0.28	1	9.1	7.3
8	1.22	3	9.1	0.35	0	0.0	0.00	0	3.2	0.12	0	5.6	2.7
9	1.25	3	9.4	0.35	0	0.0	0.00	0	3.2	0.17	0	2.9	1.4
10	1.27	5	9.1	0.36	0	0.0	0.00	0	3.1	0.15	0	7.0	2.0
11	1.30	2	9.9	0.22	0	0.0	0.00	0	4.5	0.15	0	5.8	3.3
12	1.30	2	8.5	0.39	0	0.0	0.00	2	3.2	0.25	0	4.5	2.9
13	1.36	5	9.8	0.35	0	0.0	0.00	0	3.5	0.14	0	2.2	0.9
14	1.40	6	9.0	0.42	0	0.0	0.00	0	3.1	0.18	0	4.7	2.3
15	1.43	4	10.5	0.25	0	0.0	0.00	1	4.2	0.24	0	4.9	1.5
16	1.43	2	9.5	0.38	0	0.0	0.00	0	4.4	0.16	0	9.9	2.2
17	1.44	2	9.3	0.42	0	0.0	0.00	1	3.8	0.24	0	7.1	3.6
18	1.53	4	10.1	0.39	0	0.0	0.00	0	4.0	0.12	0	9.8	2.1
19	1.56	3	8.4	0.29	0	0.0	0.00	2	4.2	0.38	1	10.5	6.1
20	1.66	7	10.6	0.42	0	0.0	0.00	0	3.8	0.16	0	4.0	1.0
Average	1.28	3	8.8	0.32	0	0.0	0.00	0	3.5	0.19	0	5.8	2.6
+/-	0.19	2	1.2	0.07	0	0.0	0.00	1	0.5	0.07	0	2.4	1.6
Minimum	0.94	1	6.2	0.22	0	0.0	0.00	0	2.6	0.12	0	2.2	0.9
Maximum	1.66	7	10.6	0.42	0	0.0	0.00	2	4.5	0.38	1	10.5	7.3

Constraint violation overview:

							max	1 5 10 15 20
Upper	QG	MET	1 - QB	ALA	26		0.23	*
Upper	HN	ASN	3 - HB2	ASN	3		0.22	*
Upper	HN	ASN	3 - HB3	ASN	3		0.42	++ + *
Upper	HA	ASN	3 - HD2	PRO	4		0.28	*
Upper	HN	THR	5 - HN	GLU-	165		0.22	*
Upper	HN	LEU	24 - HB3	LEU	24		0.22	*
Upper	HN	GLU-	43 - HB2	GLU-	43		0.20	*
Upper	HA	THR	73 - QG2	THR	73		0.42	+ +++ +++ +*+ +
Upper	HA	ILE	78 - HN	GLY	80		0.25	* +
Upper	QA	GLY	80 - HN	LYS+	82		0.22	*
Upper	QB	GLU-	81 - HN	LYS+	82		0.21	*
Upper	HN	LYS+	82 - HB2	LYS+	82		0.21	*
Upper	HN	LYS+	82 - HB3	LYS+	82		0.27	*
Upper	HN	LYS+	82 - QB	LYS+	82		0.25	+ *
Upper	HN	LEU	90 - HN	LYS+	91		0.26	+ +* ++
Upper	QD2	LEU	98 - HN	PHE	129		0.25	+ ++ *
Upper	HN	SER	99 - HN	PHE	129		0.27	*
Upper	HA1	GLY	104 - HD2	PRO	105		0.23	* +
Upper	HN	ASN	106 - HN	THR	107		0.23	+ *
Upper	HN	ASN	106 - HN	ASN	108		0.29	+ + + + +*
Upper	HA	CYS	115 - HN	ALA	117		0.27	*+++ + + + +++
Upper	HB2	LYS+	133 - HN	GLU-	134		0.24	*
Upper	QG2	ILE	158 - HG13	ILE	158		0.27	+ *
Angle	PSI	GLU-	81				7.35	*
Angle	CHI1	ASN	106				6.07	*

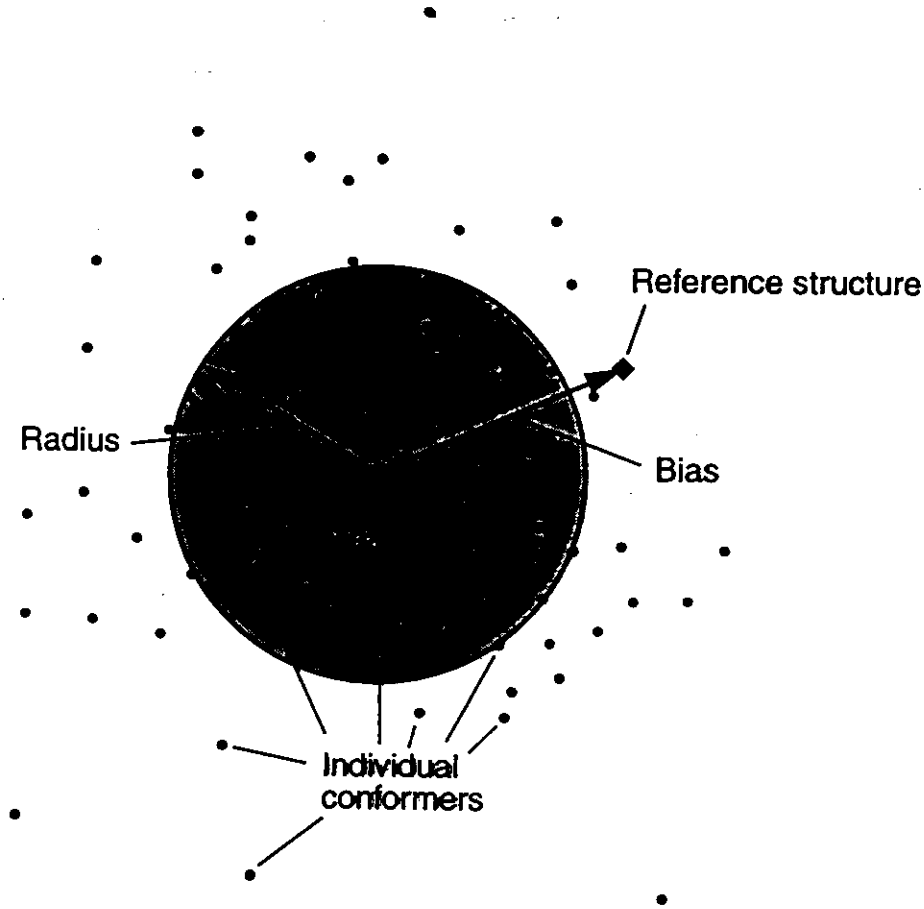
Table 7. *Experimental data and structural statistics for cyclophilin A*

Quantity[a]	Value	Comments
Resonance assignments:		
Sequence-specific[b]	88%	^{1}H 88%, ^{13}C 91%, ^{15}N, 78%
Stereospecific and individual NH_2	73	46 CH_2, 16 $C(CH_3)_2$, 11 NH_2
Experimental restraints:		
Upper distance bounds	4093	
Torsion angle restraints	371	135 ϕ, 135 ψ, 101 χ^1
Structure calculation:		
Target function	1.28 Å^2	range 0.94–1.66 Å^2
RMSD radius	0.53 Å	for N, C^{α}, C′ of all residues
Maximal restraint violations:		
Upper bounds	0.32 Å	average violation 0.0021 Å
Steric lower bounds	0.19 Å	sum of violations 3.5 Å
Torsion angle restraints	2.6°	average violation 0.016°

[a] Averaged over all accepted conformers, when applicable.

[b] Includes all protons and methyl groups with the exception of hydroxyl protons, and all ^{13}C and ^{15}N atoms with a directly bound proton.

RMSD values for a bundle of conformers



Ramachandran plot

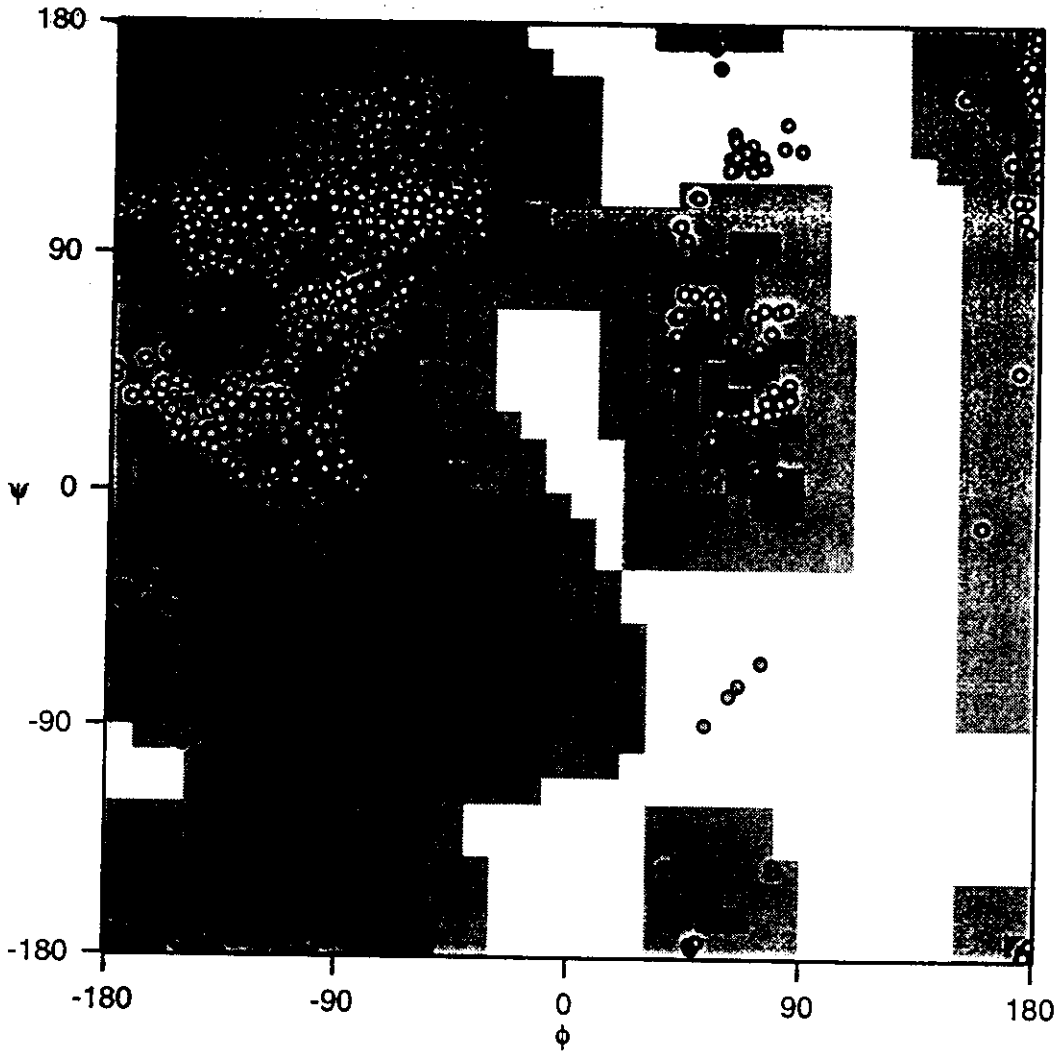


Table 8. *DYANA structure calculations for the protein Cyclophilin A using the complete experimental NMR data set and different subsets thereof*[a]

Data set[b]	Distance restraints	Success rate (%)[c]	Target function (Å^2)	Backbone RMSD (Å)[d]	
				radius	bias
All experimental restraints	4093	82	1.28	0.53	0.17
No stereospecific assignments	4394[e]	80	1.29	0.61	0.29
No angle restraints	4093	66	1.55	0.53	0.24
75% of all NOEs[f]	3054	63	1.02	0.64	0.43
50% of all NOEs	2055	65	0.72	0.79	0.70
25% of all NOEs	1038	55	0.76	1.07	1.20
10% of all NOEs	404	52	0.76	1.91	2.46
5% of all NOEs	213	64	0.83	4.41	4.43
Only backbone and Hβ NOEs	1656	26	3.86	1.35	1.25
Only HN–HN NOEs	254	38	2.89	9.46	10.09

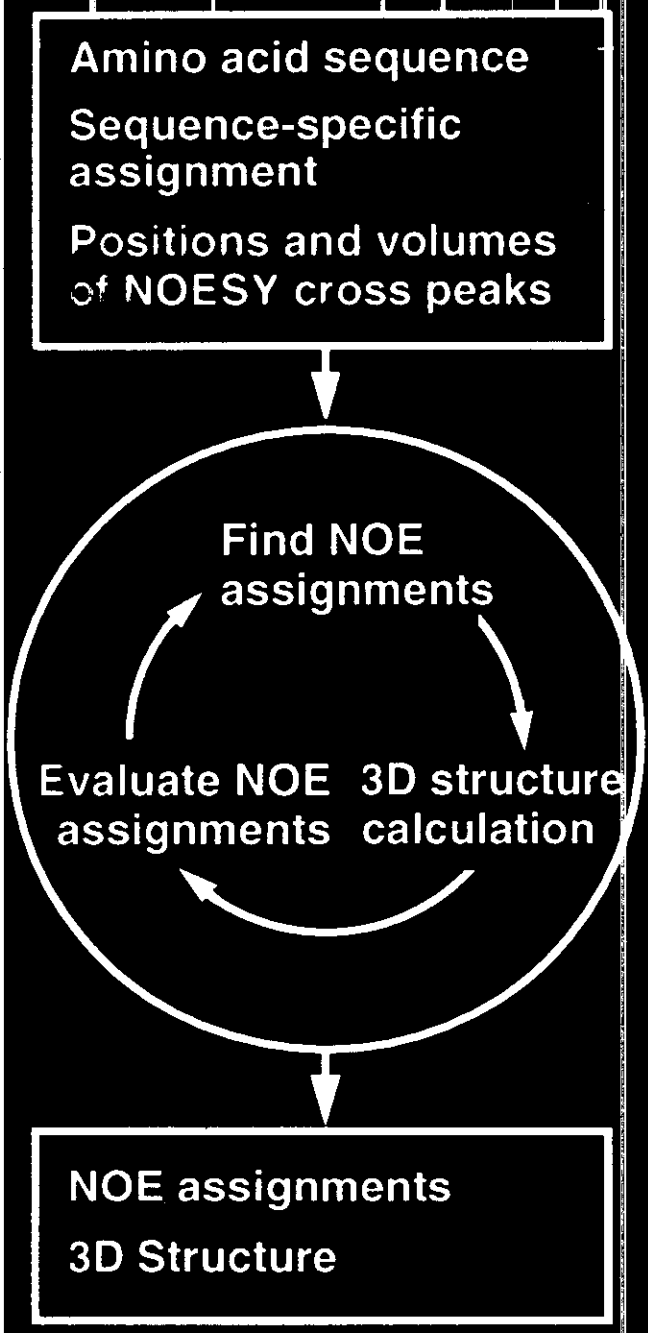
[a] For each data set 50 conformers were calculated using the standard simulated annealing protocol of the program DYANA with 8000 torsion angle dynamics steps and a target function with linear asymptote for large violations. The 20 conformers with the lowest final target function values were analyzed.

[b] The different data sets were derived from the complete experimental NMR data set for Cyclophilin A (Ottiger *et al.*, 1997) that comprises 4093 meaningful upper distance limits obtained from NOE measurements and 371 restraints for the torsion angles ϕ, ψ and χ^1. The same torsion angle restraints were included in all data sets except the one without any torsion angle restraints.

Table 9. *Structure calculations for Cyclophilin A with data sets to which first 2% distance restraints with wrong assignments were added and from which subsequently consistently violated distance restraints were eliminated.*[a]

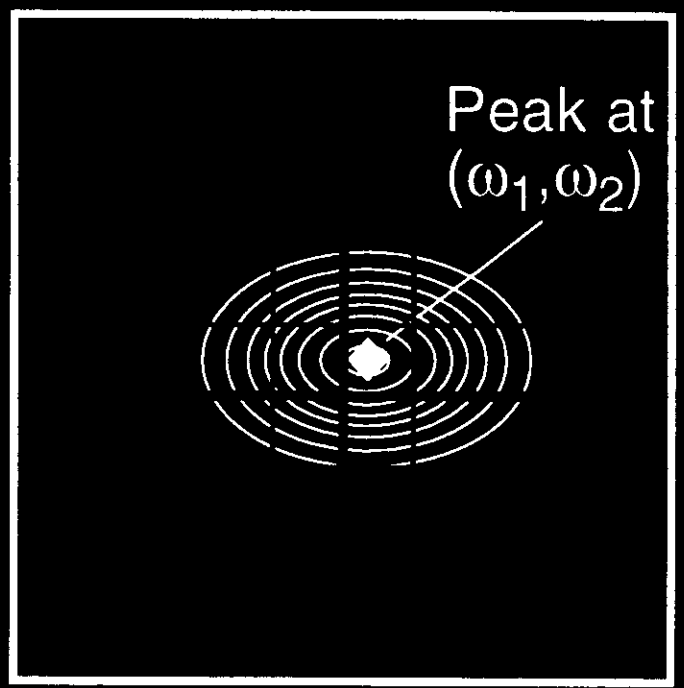
Data set	Restraints eliminated (%)[b]		Target function ($Å^2$)	Backbone RMSD (Å)[c]	
	wrong	correct		radius	bias
All experimental restraints	90	0.71	1.24	0.60	0.38
No stereospecific assignments	91	0.43	1.37	0.66	0.49
No angle restraints	89	0.66	1.59	0.59	0.44
75% of all NOEs	85	0.46	1.45	0.69	0.83
50% of all NOEs	82	0.51	1.33	0.79	1.05
25% of all NOEs	74	0.84	4.02	1.20	1.56
10% of all NOEs	50	0.49	4.02	2.05	3.37
5% of all NOEs	0	0.18	3.38	4.69	5.57
Only backbone and Hβ NOEs	70	0.96	9.54	1.51	1.62

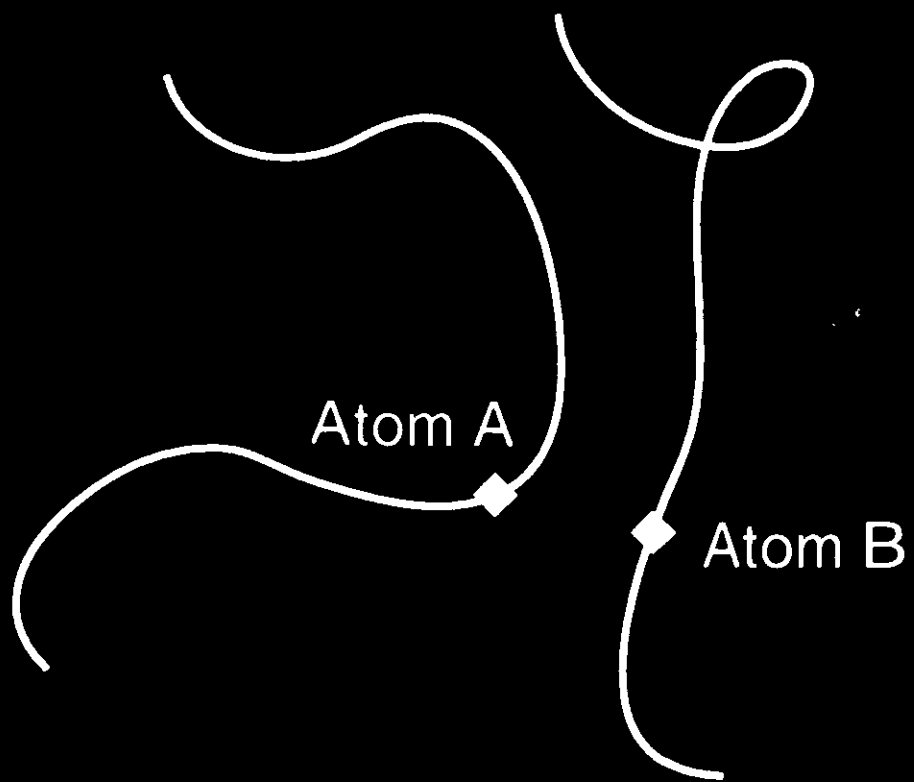
Automatic NOESY Assignment and Structure Calculation



Mumenthaler & Braun, *J. Mol. Biol.* 254, 465–480 (1995)
Mumenthaler, Günter, Braun & Wüthrich, *J. Biomol. NMR* (1997)

Conditions for valid assignments



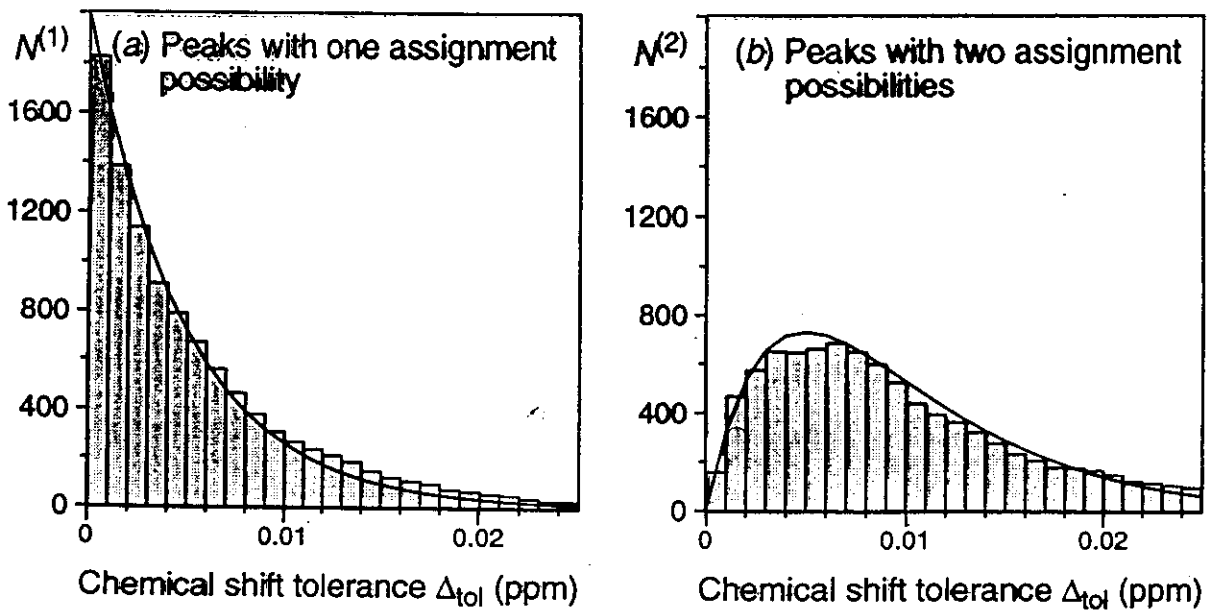


$|\omega_1 -$ $|<$

$|\omega_2 -$ $|<$

$d_{AB} <$

NOESY peaks with one and two possible assignments



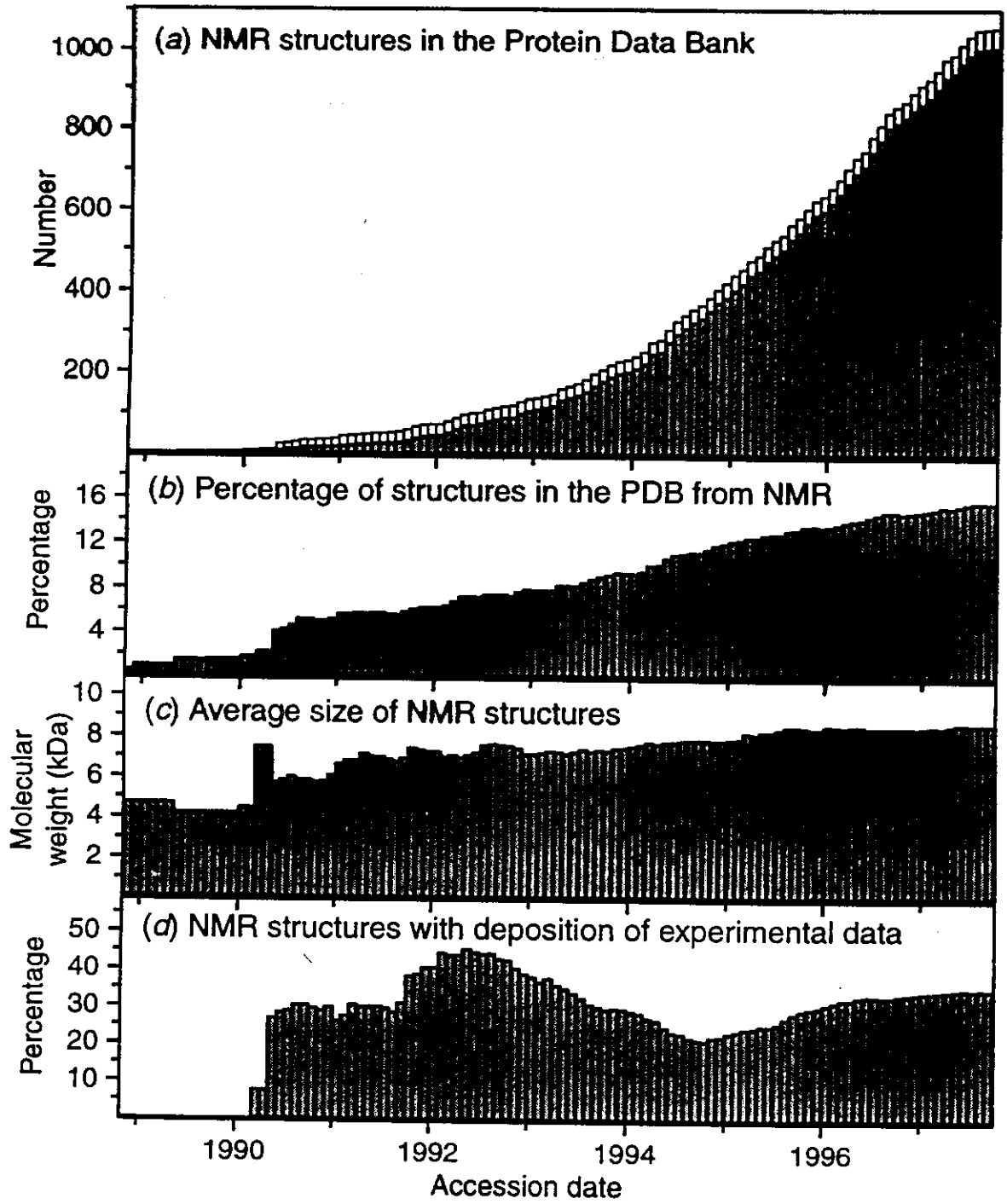
Automatic NOESY assignment

(*a*) Cycle 12

(*b*) Cycle 16

(*c*) Final structure

NMR structures solved until December 1997



Size distribution of NMR structures

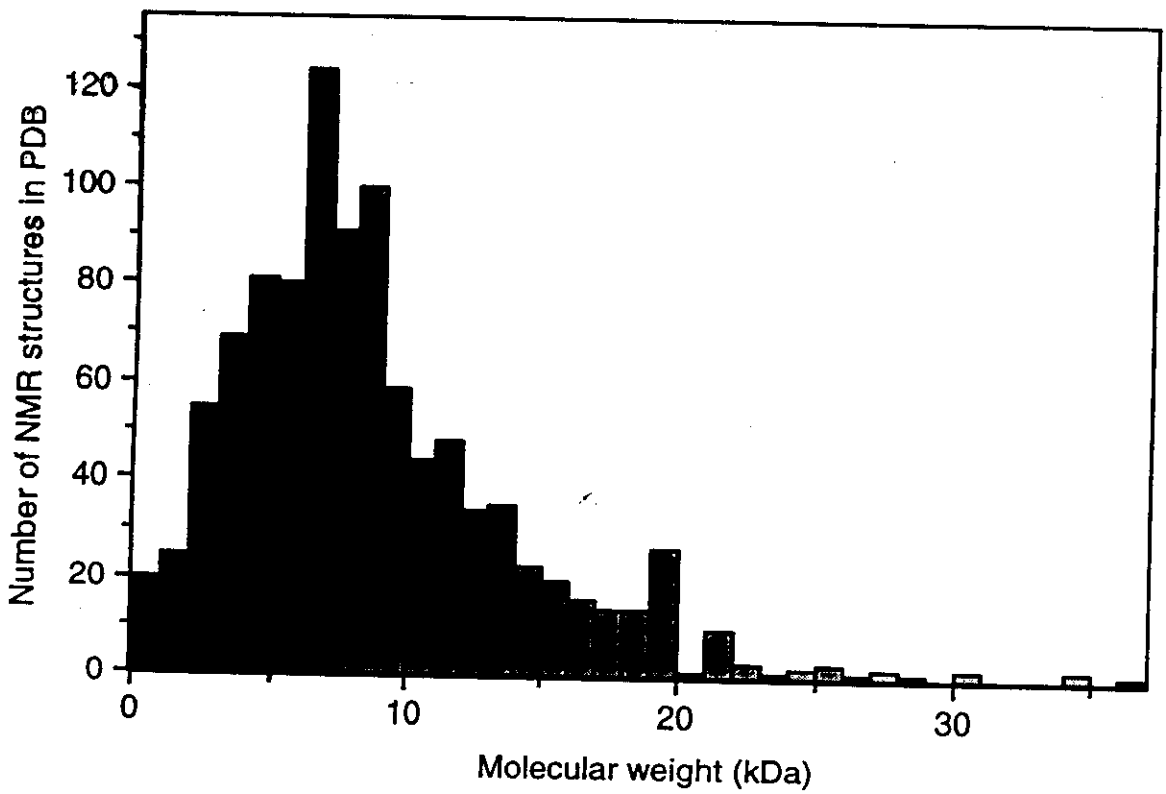


Table 2. *Structure calculation programs*

Program[a]	Structures[b]	Reference
Metric matrix distance geometry:		
DG-II	86	Havel (1991)
DISGEO	22	Havel & Wüthrich (1984)
DSPACE	27	Biosym, Inc.
EMBOSS	19	Nakai *et al.* (1993)
TINKER	3	Hodsdon *et al.* (1996)
Variable target function method:		
DIANA	124	Güntert *et al.* (1991a)
DISMAN	17	Braun & Go (1985)
Cartesian space molecular dynamics:		
AMBER	135	Pearlman *et al.* (1991)
CHARMM	88	Brooks *et al.* (1983)
DISCOVER	99	Molecular Simulations, Inc.
GROMOS	22	van Gunsteren *et al.* (1996)
SYBYL	6	Tripos, Inc.
XPLOR	570	Brünger (1992)
Torsion angle dynamics:		
DYANA	5	Güntert *et al.* (1997)

Table 1. *Journals that have published NMR structures available from the Protein Data Bank*[a]

Journal	Structures
Biochemistry	191
Journal of Molecular Biology	121
Nature Structural Biology	62
Structure	42
Science	33
Protein Science	23
European Journal of Biochemistry	21
Nature	20
Other journals	109

[a] The information was taken from the JRNL REF records of all unique coordinate files with NMR structures that were available from the Protein Data Bank in December 1997. About one third of these PDB coordinate files could not be considered because no precise reference is given (e.g. "to be published").