Journal of Computer-Aided Molecular Design, 6 (1992) 97–112
ESCOM

J-CAMD 154

Conformational search by potential energy annealing: Algorithm and application to cyclosporin A

René C. van Schaik[a,*], Wilfred F. van Gunsteren[b] and Herman J.C. Berendsen[a]

[a]*Department of Biophysical Chemistry, Rijksuniversiteit Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands*
[b]*Laboratorium für Physikalische Chemie, ETH-Zentrum, CH-8092 Zürich, Switzerland*

Received 8 August 1991
Accepted 11 November 1991

Key words: Cyclosporin A; Conformational search; MD; NOE

SUMMARY

A major problem in modelling (biological) macromolecules is the search for low-energy conformations. The complexity of a conformational search problem increases exponentially with the number of degrees of freedom which means that a systematic search can only be performed for very small structures. Here we introduce a new method (PEACS) which has a far better performance than conventional search methods.

To show the advantages of PEACS we applied it to the refinement of Cyclosporin A and compared the results with normal molecular dynamics (MD) refinement. The structures obtained with PEACS were lower in energy and agreed with the NMR parameters much better than those obtained with MD. From the results it is further clear that PEACS samples a much larger part of the available conformational space than MD does.

INTRODUCTION

A major problem in modelling (biological) macromolecules is the search for low-energy conformations. Conformational search belongs to the class of NP-complete (nondeterministic polynomial time complete) problems, the complexity of which increases exponentially with the number of degrees of freedom [1]. This means that an exhaustive conformational search cannot be performed for all but the smallest molecules for which about 10 dihedral angle degrees of freedom are scanned [2,3]. All other search methods are heuristic in that they return the lowest local minimum found as an estimate of the global minimum. Therefore, general results such as global convergence cannot be expected for these methods [4]. Further, the performance of heuristic search

* To whom correspondence should be addressed.

0920-654X/$ 5.00

methods is rather problem-specific, which means in practice that search methods that work well for fluids are not guaranteed to work for covalently bonded systems and vice versa.

An inherent problem of conventional (heuristic) search methods such as molecular dynamics (MD) [5] and Monte Carlo techniques (MC) [5] is that they sample only a small part of configuration space. This is mainly due to the fact that these methods cannot easily overcome wide or high energy barriers between different conformations. Molecular dynamics surmounts barriers of about 1 kT per degree of freedom, while Monte Carlo searching may step through higher but narrow barriers. A disadvantage of Monte Carlo techniques is their inefficiency in Cartesian space for covalently bonded systems [6]. The performance of both MD and MC can be improved by searching at higher temperatures [7] which, however, favors the selection of higher entropy conformations. This problem can be successfully avoided with annealing techniques [8]; the conformational search starts at high temperatures (typically $\approx$ 1000 K) and the system is slowly cooled down during the run. Despite its disadvantages, MD (in combination with annealing techniques generally referred to as Simulated Annealing) is at present the method of choice for the refinement of crude structures based on experimental (NMR [9], X-ray [10]) data.

Other search methods proposed in the literature [11], change the potential energy function to diminish the number of local minima [12,13]. After having found a satisfactory local minimum they relax back to the original potential energy function. The methods described by Crippen [14], and Scheraga and Purisma [15] are a variation on this theme. However, they deal with the unchanged potential energy function in a multidimensional space where there is a unique minimum. After having found this minimum, a back transformation is performed to three dimensions. In general, these methods are hard to implement and there is no guarantee that low local minima found with reduced potential energy functions or in high dimensional space, directly correspond to low local minima of the original potential energy function in three dimensions.

The search algorithm proposed by Griewank [4] uses a different approach. The basic idea behind it is to define new dynamical laws which meet requirements for a good search method. A system subject to these new laws should explore configuration space more efficiently and evolve to low local minima if not the global minimum. Unfortunately, the algorithm involves a lot of computational effort which makes it impractical to use.

Donnelly and Rogers [16] transformed the ideas of Griewank into a new algorithm (SNIFR). This algorithm has proven to be very successful for Lennard–Jones fluids (R.A. Donnelly, private communication) and for an orientation problem related to the docking of two molecules. However, the performance of the method is very sensitive to changes in its parameters, and the absence of rules-of-thumb to set these parameters is a serious problem for its applicability. Besides this, SNIFR is inefficient when handling covalently bonded systems.

In this article we present two algorithms which are variations on the MD method. We have tried to improve the searching power of MD by simple modification of its equations of motion. Our first thought was to couple every individual particle to a temperature bath, instead of the normal overall temperature coupling. This should improve the search power because all particles preserve downhill as well as uphill the same velocity (cruise control), which should make it easier to overcome barriers. This method is further referred to as MD_{in}.

Our second algorithm is inspired by the PECT method of Cotterill and Madsen [17]. They proposed an algorithm to trace constant potential energy contours (PECT) in an MD simulation. We extended this idea to a searching scheme whereby the potential energy of the system is coupled to

an energy bath. The reference energy level of the bath is slowly decreased over the run, which should result in low potential energy structures at the end (**P**otential **E**nergy **A**nnealed **C**onformational **S**earch = PEACS).

It should be noted that although both methods are based on MD, they are non-Newtonian (center of mass, motion and energy are not conserved) and do not generate Boltzmann distributions. This is not a problem as long as one keeps in mind that both methods are intended for searching and not for simulation.

As an application of MD_{in} and PEACS we tried to refine structures of cyclosporin A, a cyclic peptide of eleven residues. The refinement of this molecule is known to be a hard case, because only two out of the nine different structures found with distance geometry calculations could be refined with MD to a structure of low energy which satisfies the experimental data [18]. In this article the following points will be addressed:

(1) A detailed comparison of the various refined structures is made to test the searching properties of the different refinement methods.
(2) The refined structures are compared with the previously found solution structure (MDS1) [19] in order to estimate the degree of convergence of the various refined structures.

METHOD

Our first algorithm MD_{in} which involves individual coupling of particles to a temperature bath, requires a minor change in the original MD algorithm. Normally, the temperature in the simulation is kept approximately constant by a weak coupling to a temperature bath [20]. This can be done by scaling of the velocities at each MD time step via

$$\mathbf{v}_i'(t) = \mathbf{v}_i(t)\left[1 + \frac{\Delta t}{\tau_T}\left(\frac{T_0}{T} - 1\right)\right]^{1/2} \quad i = 1,2...\mathrm{N} \tag{1}$$

where $\mathbf{v}(t)$ is the velocity resulting from an MD step and $\mathbf{v}'(t)$ is the scaled one. The time t is integrated in time steps Δt, τ_T is the temperature relaxation time, T_0 is the reference temperature and $T(t)$ is the actual temperature of the system.

$$T(t) = \frac{1}{\frac{1}{2}k_B N_{df}} \sum_{i=1}^{N} \tfrac{1}{2} m_i \mathbf{v}_i^2(t) \tag{2}$$

where the number of degrees of freedom of the system of N particles is denoted by N_{df}, and k_B is Boltzmann's constant. Every particle in the system has the same scaling factor.

For our searching purposes we changed $T(t)$ in Eq. (1) into $T_i(t)$, the individual particle temperature, which is calculated as

$$T_i(t) = \frac{1}{\frac{1}{2}k_B N_{df}/N}\, \tfrac{1}{2} m_i \mathbf{v}_i^2 \quad i = 1,2...N \tag{3}$$

The inclusion of the factor N_{df}/N in (3) is necessary since temperature is in principle not defined

for individual particles. It takes into account the number of degrees of freedom of the whole system and ascribes to each particle an average number of degrees of freedom.

In the PEACS algorithm, the potential energy $V(\{\boldsymbol{r}(t)\})$ of the system is coupled to an energy bath using the coupling equation:

$$\frac{dV(\{\mathbf{r}(t)\})}{dt}=\frac{1}{\tau_V}\left[V_0-V(\{\mathbf{r}(t)\})\right] \tag{4}$$

The potential energy V generally depends on the coordinates of all particles denoted by $\{\boldsymbol{r}(t)\}$, V_0 is the reference energy level of the bath and τ_V is the potential energy relaxation time. Equation (4) can also be written as:

$$\sum_{i=1}^{N}\nabla_i V(\{\mathbf{r}(t)\}).\mathbf{v}_i^{corr}(t)=\frac{1}{\tau_V}\left[\mathbf{V}_0-\mathbf{V}(\{\mathbf{r}(t)\})\right] \tag{5}$$

Equation (5) defines a so-called velocity correction $\mathbf{v}_i^{corr}(t)$ per particle. Since this correction is meant to keep the potential energy V nearly constant, a logical assumption is to take its direction along the gradient of V:

$$\mathbf{v}_i^{corr}(t)=\varepsilon_i\nabla_i V(\{\mathbf{r}(t)\}) \quad i=1,2...N \tag{6}$$

Inserting (6) into (5) yields one equation with N unknowns ε_i ($i=1,2,...,N$):

$$\sum_{i=1}^{N}\varepsilon_i\|\nabla_i V(\{\mathbf{r}(t)\})\|^2=\frac{1}{\tau_{\mathbf{v}}}\left[V_0-V(\{\mathbf{r}(t)\})\right] \tag{7}$$

In order to obtain N equations for the unknown ε_i ($i=1,2,...,N$), the ε_i are chosen such that all particles contribute equally to the left-hand side of Eq. (7), which then becomes:

$$N\varepsilon_i\|\nabla_i\mathbf{V}(\{\mathbf{r}(t)\})\|^2=\frac{1}{\tau_{\mathbf{v}}}\left[V_0-V(\{\mathbf{r}(t)\})\right] \quad i=1,2...N \tag{8}$$

The ε_i can be found from:

$$\varepsilon_i=\frac{[V_o-V(\{\mathbf{r}(t)\})]}{\tau_V N\|\nabla_i V(\{\mathbf{r}(t)\})\|^2} \quad i=1,2,...N \tag{9}$$

The velocity formula of the leap-frog integration scheme reads:

$$\mathbf{v}_i^{MD}(t+\tfrac{1}{2}\Delta t)=\mathbf{v}_i^{MD}(t-\tfrac{1}{2}\Delta t)-\frac{\Delta t}{m_i}\nabla_i V(\{\mathbf{r}(t)\}) \quad i=1,2,...N \tag{10}$$

The PEACS modification of it is obtained by addition of the 'correction' velocity $\mathbf{v}_i^{corr}$ as defined through (6):

$$\mathbf{v}_i^{PEACS}(t+\tfrac{1}{2}\Delta t)=\mathbf{v}_i^{PEACS}(t-\tfrac{1}{2}\Delta t)+\left(\varepsilon_i-\frac{\Delta t}{m_i}\right)\nabla_i V(\{\mathbf{r}(t)\}) \quad i=1,2,\ldots N \tag{11}$$

or using (9) we find:

$$\mathbf{v}_i^{PEACS}(\mathbf{t}+\tfrac{1}{2}\Delta t)=\mathbf{v}_i^{PEACS}(t-\tfrac{1}{2}\Delta t)-\left[1-\frac{m_i(V_0-V(\{\mathbf{r}(t)\}))}{N\tau_V\Delta t\|\nabla_i V(\{\mathbf{r}(t)\})\|^2}\right]\frac{\Delta t}{m_i}\nabla_i V(\{\mathbf{r}(t)\}) \quad i=1,2,\ldots N \tag{12}$$

Equation (12) together with the leap-frog position formula:

$$\mathbf{r}_i(t+\Delta t)=\mathbf{r}_i(t)+\mathbf{v}_i^{PEACS}(t+\tfrac{1}{2}\Delta t)\cdot\Delta t \quad i=1,2\ldots N \tag{13}$$

forms the leap-frog PEACS algorithm, which can perform conformational search at a certain potential energy level V_0.

However, to let the algorithm find low potential energy conformations, we have to bring down the reference level V_0 of the energy bath. This could simply be done by choosing a target level for V_0 close to the global energy minimum. The reference level V_0 at $t=0$ is set equal to the potential energy of the initial structure and subsequently V_0 is linearly decreased to reach the target level at the end of the run. A severe disadvantage of this method is that one needs to have some idea of the potential energy values of low local minima if not the global minimum. Since this information is, in general, not easily available this method is not practical.

We decided to use a method whereby the reference level $V_0(t)$ is made a function of time t and is brought down in accordance with the lowest energy values found in the trajectory, with a relaxation time τ_{V0}. In algorithmic form the procedure reads:

(1) $V_0(0)=V(\{\mathbf{r}(0)\})$
$V_{\min}=V(\{\mathbf{r}(0)\})$
(2) Calculate $F_i(t)=-\nabla_i V(\{\mathbf{r}(t)\})$ and $V(\{\mathbf{r}(t)\})$
(3) IF $V(\{\mathbf{r}(t)\})<V_{\min}$ THEN
$V_{\min}=V(\{\mathbf{r}(t)\})$
(4) $V_0(t)=V_0(t-\Delta t)\exp(-\Delta t/\tau_{V0})+V_{\min}(1-\exp(-\Delta t/\tau_{V0}))$
(5) Make a PEACS leap-frog step from t to $t+\Delta t$
(6) GOTO 2

With the memory function in step (4) we can select how fast the reference level is decreased. If τ_{V0} is chosen equal to the time step Δt then V_0 immediately follows V_{min}. For long memory times (i.e. $\tau_{V0}\gg\Delta t$) V_0 is scaled down very slowly, as required for an annealing scheme.

The PEACS algorithm introduces two extra parameters: τ_V the potential energy relaxation time and τ_{V0} the characteristic time of the V_0 memory function.

MOLECULAR MODEL AND COMPUTATIONAL PROCEDURE

The programs for carrying out conformational search and analyzing results were taken from the GROMOS87 library [21]. The potential energy function describing the interaction between protein atoms in vacuo (37D4 potential parameter set in Ref. [21]) is of the usual type applied to proteins. It is composed of terms representing bond angle bending, harmonic (out-of-plane, out-of-tetrahedral configuration) dihedral bending, sinusoidal torsion, van der Waals and electrostatic interactions [22,23]. Because of the small size of the system of interest, there was no need to apply a cutoff on the evaluation of the nonbonded forces. Bond lengths were constrained during the conformational search and energy minimization period by the SHAKE method [24,25] with relative tolerance 10^{-4}.

An extra term in the interatomic potential function was used, which forces the molecule to satisfy a set of 57 atom–atom distance constraints originating from nuclear Overhauser effects (NOEs) obtained from NMR spectroscopy. This energy term was harmonic with respect to the NOE distance constraint violations with a force constant of 4000 $kJ \cdot mol^{-1} \cdot nm^{-2}$ [19,26]. Note that one NOE (between 8-D-Ala-NH and 6-MeLeu-$C_\gamma H_3$) has been omitted from the original list of 58 in Ref. [26] since it has been misassigned.

The runs were performed at constant temperature whereby the system was weakly coupled to a thermal bath using the Berendsen thermostat [20], with coupling relaxation time $\tau_T = 0.1$ ps, and for the PEACS algorithm we used $\tau_V = 0.1$ ps and $\tau_{V0} = 2.5$ ps. In all simulations the time step was 0.002 ps [27].

Two types of $MD/MD_{in}/PEACS$ runs were performed:

1. Runs at constant temperature, which consist of 40 ps at $T = 300$ K whereby the initial velocities were taken from a Maxwellian distribution. The final structures were obtained from an energy minimization (250 steps of steepest descent with line minimizations) of the lowest potential energy structure from the trajectories.
2. Runs with temperature annealing. For these runs we used the following protocol:
   - Generate velocities from a Maxwellian distribution at 1200 K.
   - Perform 20 ps of conformational search at $T = 1200$ K.
   - Select the conformation with lowest potential energy from the generated trajectory and use this conformation for the next steps.
   - Generate velocities from a Maxwellian distribution at 1200 K.
   - Perform 20 ps of conformational search whereby the temperature is linearly decreased from 1200 K to 300 K over the first 15 ps.
   - Select the conformation with lowest potential energy from the generated trajectory and perform energy minimization using 250 steps of steepest descent with line minimizations.

In Ref. [18] the distance geometry algorithm DGEOM was repeatedly applied to CPA using the original set of 58 NOE distance constraints, which resulted in 27 structures. The mutual root-mean-square (RMS) positional differences were calculated for these structures and it appeared that it was possible to group the structures into nine distinct classes, denoted by xclass1 to xclass9. The first structure of each class is used as starting structure for the various refinement runs discussed below.

RESULTS AND DISCUSSION

(A) Results of the parametrization study for PEACS

The parametrization study for the PEACS parameters was carried out using structure xclass11 (the first structure of the first class of DG-structures [18]) as a test system. We tried different values for τ_{V0}, τ_V and τ_T while running at different temperatures with and without temperature annealing. We obtained the best results with the characteristic time of the V_0 memory function, τ_{V0} as large as possible. This is easy to explain, because the larger τ_{V0} the slower the level of the energy bath comes down, as required for good annealing schemes. So, in general τ_{V0} will be limited by the length of the run. For our runs of 40 ps we took $\tau_{V0} = 2.5$ ps.

The flow of kinetic energy to potential energy and backwards has a relaxation time of about 0.1 ps for biomolecules. This means that when $\tau_V \leqslant 0.1$ ps the potential energy will fluctuate around V_0. When $\tau_V \gg 0.1$ ps the PEACS algorithm reduces to normal MD, since the coupling to the energy bath becomes insignificant.

The results further indicated that there was no need to change the temperature relaxation time τ_T with respect to the value used in normal MD. We obtained the best results with $\tau_T = 0.1$ ps. However, smaller τ_T values are equally well possible.

We also tried a combination of MD_{in} and PEACS, i.e. PEACS with temperature coupling per particle. This idea was not very successful, numerical errors including SHAKE errors/high bond energies were the result of this attempt.

(B) Results of the refinement of cyclosporin A

In order to test the quality of our final structures, we compared the potential energies and sum of the NOE distance violations. Well-refined structures should have a low potential energy to make them statistically meaningful and a low sum of violations to let them fulfill the NMR dataset. We also compared the final structures to the previously found solution structure MDS1 [19]. However, it is not necessary for refined structures to deviate little from the MDS1 structure because CPA in solution might exist in different conformations. Structural comparisons are discussed on the base of RMS positional differences and pictures from graphics systems. Structural comparisons are also possible on the base of dihedral angle information, however for the comparisons below, this did not lead to extra information or different conclusions (results not shown).

The potential energy, total violation of the distance constraints and the root-mean-square (RMS) difference between final structure and MDS1 structure for refinement starting from the nine distinct structures are shown in Table 1. In Table 1A the results are shown for the runs at 300 K. On average, the results of MD_{in} are slightly better than those of MD. Both methods do however not lead to well-refined structures, the energies and sum of the distance-constraint violations remain too high. The results with PEACS are better, and give rise to one refined structure which is very close to the MDS1 structure. On average the PEACS structures have a lower potential energy and lower sum of violations.

The results for the runs at 1200 K shown in Table 1B reveal much more the benefits of PEACS. Again the results of MD_{in} are slightly better than those of MD. MD has found one refined structure (xclass91) close to the MDS1 structure and another structure (xclass41) with a reasonably low potential energy, which satisfies the NMR data. The MD_{in} results show two refined structures both close to the MDS1 structure (xclass21 and 91). The results of both methods show clearly the performance enhancement by making use of temperature annealing.

The results of PEACS with temperature annealing are even better. The final structures fulfill, in general, the NMR dataset much better than those of the MD and MD_{in} refinement and also the energies are much lower. Five of the structures become well-refined with this method (xclass21,

TABLE 1A
RESULTS OF THE REFINEMENT OF CYCLOSPORIN A, USING MD, MD_{in} AND PEACS AT 300 K

	MD at 300 K			MD_{in} at 300 K			PEACS at 300 K		
Structure[a]	E_{pot}[b] (kJ/mol)	Σ_{viol}[c] (Å)	Δr_{MDS1}[d] (Å)	E_{pot} (kJ/mol)	Σ_{viol} (Å)	Δr_{MDS1} (Å)	E_{pot} (kJ/mol)	Σ_{viol} (Å)	Δr_{MDS1} (Å)
xclass11	200	3.6	3.4	182	4.0	3.7	98	1.8	2.6
xclass21	77	3.2	1.7	80	3.1	1.9	75	1.3	2.6
xclass31	178	5.6	3.9	146	4.3	3.7	165	3.8	3.4
xclass41	170	2.3	2.3	124	2.3	2.5	94	2.4	2.7
xclass51	138	5.2	2.3	143	4.4	2.0	157	4.5	2.3
xclass61	153	4.7	3.9	146	5.3	3.7	195	6.5	3.7
xclass71	122	5.4	1.8	118	5.5	1.9	104	6.0	2.0
xclass81	124	1.8	2.3	146	3.6	2.2	118	3.2	2.4
xclass91	132	4.2	1.9	88	2.6	2.5	17	1.1	0.6

[a] The names of the structures are taken from Ref. [18].
[b] Potential energy of the system after energy minimization. Distance-restraint energy is not included.
[c] Sum of all distance violations in excess of the reference NOE distances.
[d] RMS positional differences of the C_α atoms with respect to the MDS1 structure.

TABLE 1B
RESULTS OF THE REFINEMENT OF CYCLOSPORIN A, USING MD, MD_{in} AND PEACS ANNEALED FROM 1200 K

	MD annealed from 1200 K			MD_{in} annealed from 1200 K			PEACS annealed from 1200 K		
Structure[a]	E_{pot}[b] (kJ/mol)	Σ_{viol}[c] (Å)	Δr_{MDS1}[d] (Å)	E_{pot} (kJ/mol)	Σ_{viol} (Å)	Δr_{MDS1} (Å)	E_{pot} (kJ/mol)	Σ_{viol} (Å)	Δr_{MDS1} (Å)
xclass11	140	3.2	3.1	120	2.5	3.5	84	1.4	2.0
xclass21	126	4.0	1.7	9	1.4	0.6	27	0.9	0.3
xclass31	104	3.9	2.9	112	3.4	2.8	23	1.1	0.4
xclass41	70	0.9	1.8	82	1.4	2.2	49	1.2	1.9
xclass51	153	3.6	2.8	158	3.4	2.6	48	1.0	0.6
xclass61	148	4.5	3.5	148	3.5	3.9	81	2.0	1.9
xclass71	96	3.2	1.9	80	2.5	2.7	125	1.1	1.3
xclass81	92	2.5	2.5	112	2.8	2.1	143	1.2	2.1
xclass91	2	1.4	0.3	17	0.7	0.3	30	2.5	0.5

[a] The names of the structures are taken from Ref. [18].
[b] Potential energy of the system after energy minimization. Distance-restraint energy is not included.
[c] Sum of all distance violations in excess of the reference NOE distances.
[d] RMS positional differences of the C_α atoms with respect to the MDS1 structure.

TABLE 1C
RESULTS OF MD AND PEACS ANNEALED FROM 1200 K, USING CONJUGATE GRADIENTS ENERGY MINIMIZATION WITHOUT USING SHAKE (SEE TEXT)

	MD annealed from 1200 K	MD_{in} annealed from 1200 K	PEACS annealed from 1200 K
Structure[a]	E_{pot}[b] (kJ/mol)	E_{pot} (kJ/mol)	E_{pot} (kJ/mol)
xclass11	57	39	0
xclass21	51	−70	−65
xclass31	35	41	−66
xclass41	−11	−5	−26
xclass51	73	78	−65
xclass61	73	76	−8
xclass71	−1	5	27
xclass81	31	92	150
xclass91	−82	−77	−50

[a] The names of the structures are taken from Ref. [18].

[b] Potential energy of the system after energy minimization (see text). Bond energy is included, distance-restraint energy is not included.

31, 41, 51 and 91), while two others (xclass11 and 61) have a reasonably low energy and a low sum of distance-restraints violations. We call structure xclass91 refined although it has a sum of the violations of 2.5 Å, this is because there are two bad NOE contacts in the side chains responsible for this larger violation, while the overall conformation of the peptide is correct.

To make it possible to compare the results of our method with methods which do not make use of bond-length constraints (e.g. SNIFR) we show in Table 1C the results of MD and PEACS at 1200 K, whereby the final structures were minimized using 1000 steps of conjugate gradients without bond-length constraints. The energies drop about 50 kJ/mol for the MD structures and about 80 kJ/mol for the PEACS structures compared to those in Table 1B. This energy difference is mainly caused by a reduction of the nonbonded energy as a result of the relaxation of the covalent bonds.

The final structures of the PEACS refinement at 1200 K fall into four groups (see Table 2 and Fig. 1). The first group is formed by the structures xclass21, 31, 51 and 91. All these structures are within 0.9 Å of each other and the MDS1 structure (see Fig. 2), and fulfill (except for xclass91) the NMR dataset very well. The second group of structures is formed by the structures xclass11, 41 and 61. The members of this group are higher in energy and fulfill the dataset less well than the former group. We were not able to obtain low-energy structures for xclass71 and 81 with any of the methods. The best results for these structures were obtained with MD and MD_{in} at 1200 K.

(C) Comparison of the searching power of the methods

We compared the searching power of the methods by looking at properties like root-mean-square (RMS) positional difference between initial and final structures, RMS positional fluctuations over the generated trajectories and the number of backbone dihedral transitions. These properties reflect in one way or another the extent of conformational space covered by the search, and therefore the searching power of the various methods.

TABLE 2
ROOT-MEAN-SQUARE POSITIONAL DIFFERENCES (Å) BETWEEN THE MDS1 STRUCTURE AND NINE REFINED STRUCTURES (OBTAINED WITH PEACS AT 1200 K USING TEMPERATURE ANNEALING)

	MDS1	11	21	31	41	51	61	71	81	91
MDS1	0.0	1.8	0.3	0.4	2.0	0.6	1.9	1.3	2.1	0.5
11	2.9	0.0	1.8	1.9	0.8	1.5	0.5	2.3	2.6	1.6
21	0.7	3.0	0.0	0.7	2.0	0.5	1.9	1.4	2.1	0.5
31	0.7	3.0	1.2	0.0	2.0	0.9	2.1	1.2	2.1	0.7
41	3.2	1.6	3.2	3.2	0.0	1.8	0.6	2.3	2.7	1.8
51	1.1	2.8	0.9	1.5	3.3	0.0	1.7	1.7	2.3	0.5
61	3.1	1.5	3.1	3.2	1.2	3.1	0.0	2.4	2.7	1.7
71	3.2	4.4	3.2	3.2	4.4	3.3	4.4	0.0	2.0	1.4
81	3.8	4.7	3.8	3.8	4.7	4.0	4.5	3.1	0.0	2.2
91	1.1	2.8	1.3	1.3	3.2	1.4	3.0	3.3	4.0	0.0

Upper righthand triangle: RMS positional differences for C_α atoms. Lower lefthand triangle: RMS positional differences for all atoms.

In Tables 3A and B we have listed the difference in potential energy between initial and final structures and the RMS positional fluctuations over the generated trajectory. A good search method should have a large RMS positional fluctuation because this indicates that a large portion of conformational space has been sampled. In the third column the RMS positional difference between initial and final structures is given. Also this number should be large, to indicate that the method can induce large structural changes. However, when the starting structure is already close to the wanted low-energy structure, one should expect only minor structural changes.

A remarkable result shown in Table 3A is that MD shows large RMS positional *differences* between C_α *atoms* of initial and final structures while MD_{in} shows large RMS positional *fluctuations* for *all atoms* over the generated trajectory. The conclusion could be that MD_{in} increases the side-chain motions and restricts the backbone motion. This is opposite to what one expects, because

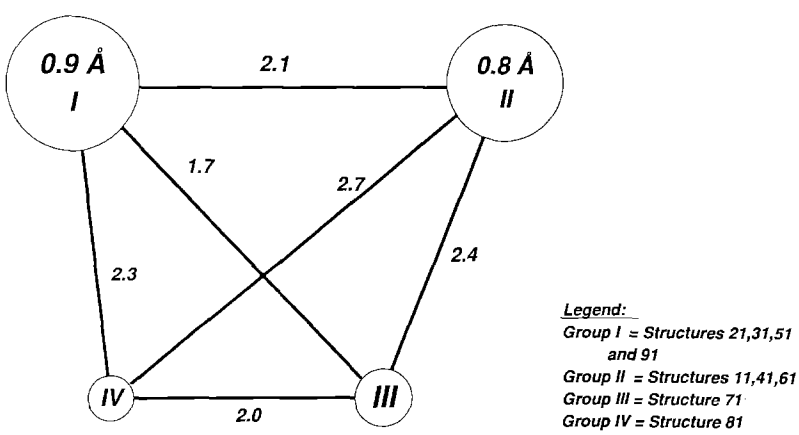


Fig. 1. Mutual RMS positional differences between all structures refined with PEACS annealed from 1200 K, this figure is a graphical interpretation of the structural differences given in Table 2.

TABLE 3A
COMPARISON OF THE SEARCHING POWER OF MD, MD_{in} AND PEACS AT 300 K

Structure[a]	MD at 300 K			MD_{in} at 300 K			PEACS at 300 K		
	ΔE_{pot}[b] (kJ/mol)	$(\Delta r^2)^{1/2}$[c] (Å)	$\Delta r_{initial}$[d] (Å)	ΔE_{pot} (kJ/mol)	$(\Delta r^2)^{1/2}$ (Å)	$\Delta r_{initial}$ (Å)	ΔE_{pot} (kJ/mol)	$(\Delta r^2)^{1/2}$ (Å)	$\Delta r_{initial}$ (Å)
xclass11	−984	1.0	1.6	−1002	3.7	0.8	−1086	1.8	2.2
xclass21	−1180	0.6	1.7	−1177	1.9	0.6	−1182	1.9	2.7
xclass31	−852	0.8	1.7	−884	3.7	0.8	−865	1.7	1.3
xclass41	−824	0.9	1.3	−807	2.5	1.1	−837	2.2	2.4
xclass51	−1102	0.9	2.7	−1097	2.0	0.8	−1083	1.4	2.7
xclass61	−1192	0.6	1.2	−1199	3.7	0.6	−1150	2.5	1.7
xclass71	−927	0.6	1.1	−931	1.9	0.6	−945	2.1	1.3
xclass81	−847	0.6	0.9	−869	2.2	0.7	−853	1.4	1.0
xclass91	−849	0.8	1.4	−893	2.5	0.9	−964	1.6	1.3
Averages	−973	0.8	1.5	−984	2.7	0.8	−996	1.8	1.8

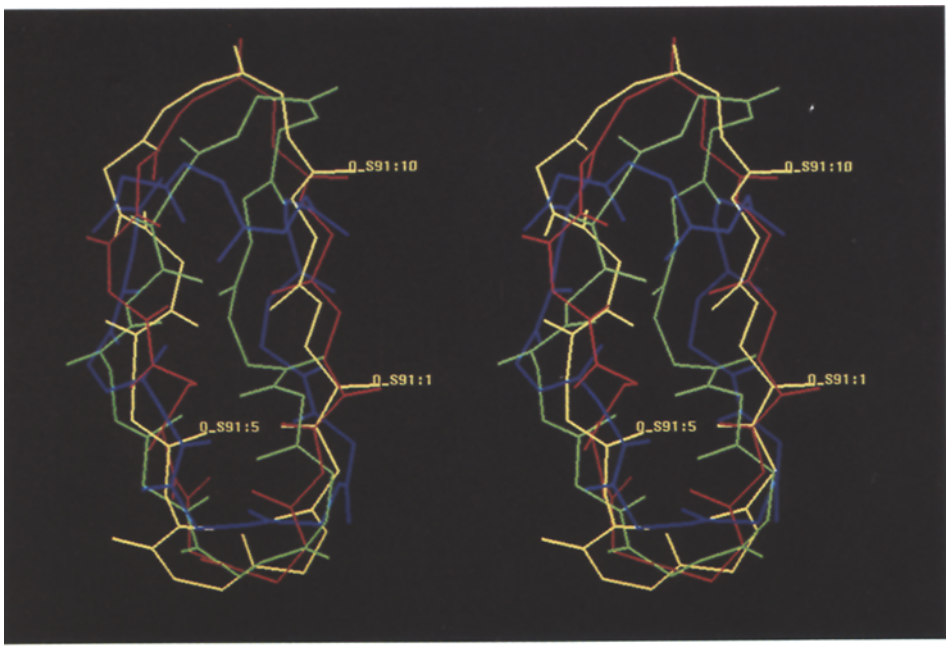
[a] The names of the structures are taken from Ref. [18].
[b] Potential energy difference between starting and final structures. Distance-restraint energy is not included.
[c] Average RMS positional fluctuation for all atoms.
[d] RMS positional difference of the C_α atoms between starting and final structure.

TABLE 3B
COMPARISON OF THE SEARCHING POWER OF MD, MD_{in} AND PEACS ANNEALED FROM 1200 K

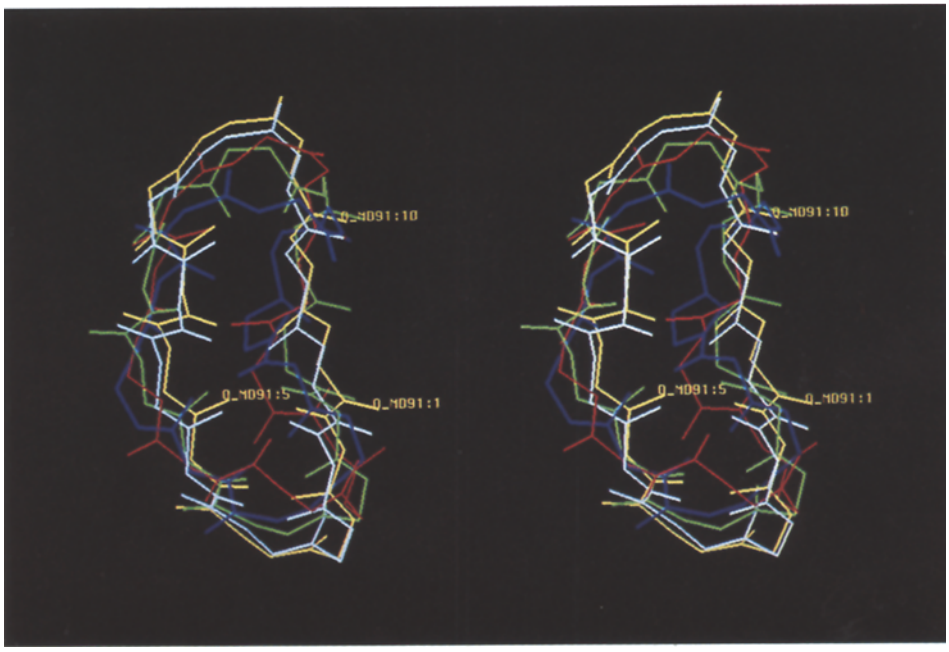
Structure[a]	MD annealed from 1200 K			MD_{in} annealed from 1200 K			PEACS annealed from 1200 K		
	ΔE_{pot}[b] (kJ/mol)	$(\Delta r^2)^{1/2}$[c] (Å)	$\Delta r_{initial}$[d] (Å)	ΔE_{pot} (kJ/mol)	$(\Delta r^2)^{1/2}$ (Å)	$\Delta r_{initial}$ (Å)	ΔE_{pot} (kJ/mol)	$(\Delta r^2)^{1/2}$ (Å)	$\Delta r_{initial}$ (Å)
xclass11	−1044	1.2	1.4	−1064	1.2	1.7	−1100	1.7	2.0
xclass21	−1131	1.3	1.6	−1248	1.4	2.2	−1230	1.8	2.1
xclass31	−926	1.6	2.3	−918	1.4	1.5	−1007	2.1	3.3
xclass41	−861	1.2	1.2	−849	1.2	1.1	−882	1.4	1.2
xclass51	−1087	1.3	3.5	−1082	1.5	3.2	−1192	1.8	1.7
xclass61	−1197	1.2	1.2	−1197	1.3	1.3	−1264	2.1	3.1
xclass71	−953	1.2	1.3	−969	1.4	2.3	−924	1.4	1.2
xclass81	−879	1.3	1.1	−859	1.2	1.3	−926	2.0	0.8
xclass91	−979	1.2	1.4	−964	1.2	1.6	−951	1.5	1.6
Averages	−1006	1.3	1.7	−1017	1.3	1.8	−1053	1.8	1.9

[a] The names of the structures are taken from Ref. [18].
[b] Potential energy difference between starting and final structure. Distance-restraint energy is not included.
[c] Average RMS positional fluctuation for all atoms.
[d] RMS positional difference of the C_α atoms between starting and final structure.

A


0_S91:10
0_S91:1
0_S91:5
0_S91:10
0_S91:1
0_S91:5

B


0_M091:10
0_M091:5
0_M091:1
0_M091:10
0_M091:5
0_M091:1

C

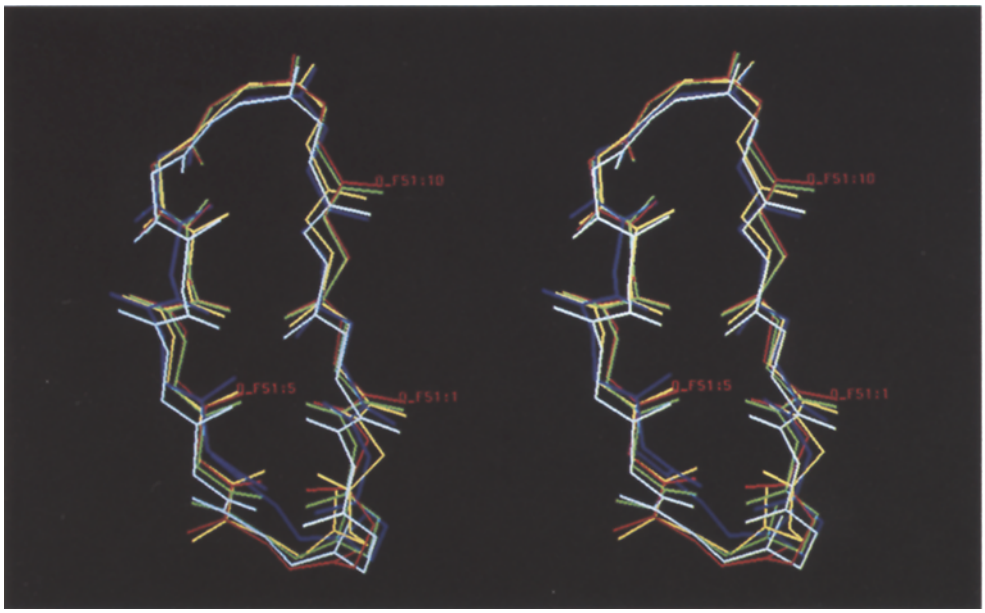


Fig. 2. A: distance geometry structures. B: MD refined structures using annealing from 1200 K + MDS1. C: PEACS refined structures using annealing from 1200 K + MDS1. Stereopairs of four CPA structures and the MDS1 structure; only backbone atoms are shown. The color code used is as follows: xclass21 green, xclass31 blue, xclass51 red, xclass91 yellow, and MDS1 white.

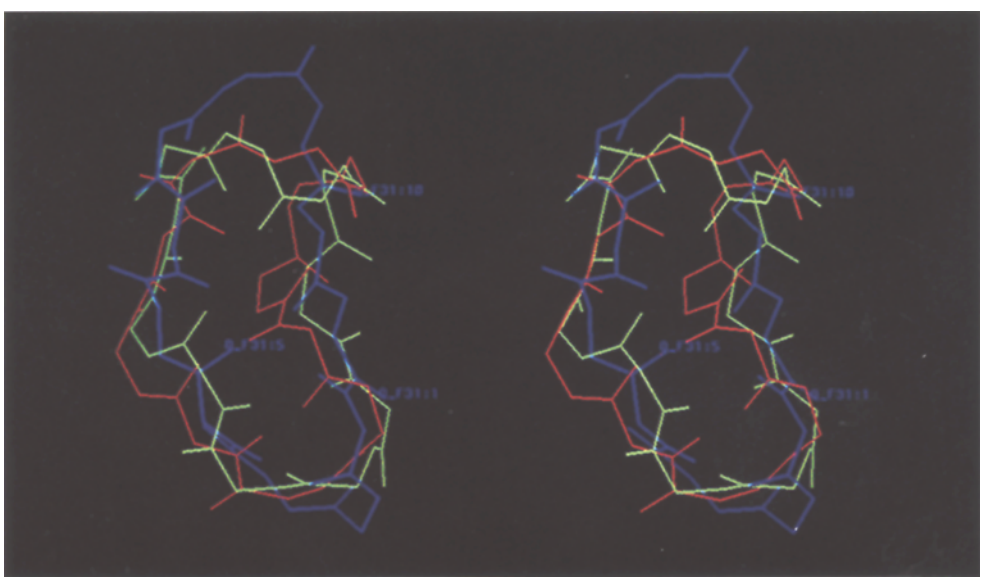


Fig. 3. Stereopairs of the starting and final structures of xclass31. The starting structure is shown in green, the structure obtained with MD in red and the structures obtained with PEACS in blue. The MD and PEACS structures are obtained using annealing from 1200 K.

in general side-chain atoms show larger amplitude motion than backbone atoms during MD simulations. However, different quantities – a *difference* between two structures, and a *fluctuation* over a trajectory – are compared for C_α atoms in the first case and all atoms in the last case. If we consider RMS *fluctuations* we find for C_α atoms 0.49 Å in case of MD and 0.54 Å in case of MD_{in}, to be compared with for all atoms 0.8 Å in case of MD and 2.7 Å in case of MD_{in}. So, MD and MD_{in} yield comparable amplitude of motion for the backbone atoms. For the side chains the MD_{in} algorithm searches a much larger part of space.

From Table 3A, it is further clear that PEACS combines the good properties of MD and MD_{in}. The PEACS method shows large RMS positional differences and large positional fluctuations. The average C_α positional fluctuation is 1.27 Å which is far larger than for the MD methods.

Table 3B shows that at 1200 K, the structural changes between initial and final structures are almost equal for all three methods (except that the structural changes from PEACS lead to more low-energy structures). The positional fluctuations are, for the PEACS method at this temperature, about a factor 1.4 for all atoms and a factor 1.5 for C_α atoms larger than those for MD and MD_{in}. Figure 3 shows xclass31 before and after PEACS refinement as an example of the power of the PEACS method for making structural changes.

Another interesting property to look at is the number of backbone dihedral transitions. A large number of backbone dihedral transitions indicates that a particular search method is capable of making significant structural changes. Figures 4A and B show bar diagrams with the number of backbone dihedral transitions for MD, MD_{in} and PEACS at 1200 K. Both diagrams show that the use of PEACS results in about 40% more backbone dihedral transitions. Unfortunately the increase in dihedral transitions mainly takes place in regions that were already relatively mobile (residues 1–7). Note that although structure 81 shows an overwhelming number of dihedral transitions (mainly for residues 1–7) when using PEACS this does not lead to a low-energy structure which satisfies the NMR constraints.

A problem of the PEACS method is that sometimes the system becomes instable because it is forced to a low energy for which a conformation is difficult to find. The forces acting on the atoms become enormous, which will end the run because of numerical errors. The easiest way to continue is to take the lowest energy structure found so far and perform the remaining steps starting with random velocities. This situation occurred four times when running at 300 K and never occurred when starting at 1200 K.

CONCLUSION

The performance of MD as a method for conformational search can be improved by coupling individual particles to a temperature bath (MD_{in}) or by coupling the system to a potential-energy bath (PEACS). The latter method has a far better performance than MD almost without any extra computational costs. It is able to find more low-energy structures and it samples much more of the available conformational space. At this moment the PEACS method is still under investigation; we are testing its properties for minimizing the energy in Lennard–Jones liquids and its performance when applied in structure refinement using X-ray diffraction data.

A point we have not discussed so far is that NOEs and J coupling parameters are based on ensemble averages. The interatomic distances calculated from the initial NOE build-up rates should therefore also be interpreted as average interatomic distances in the ensemble. This point of view

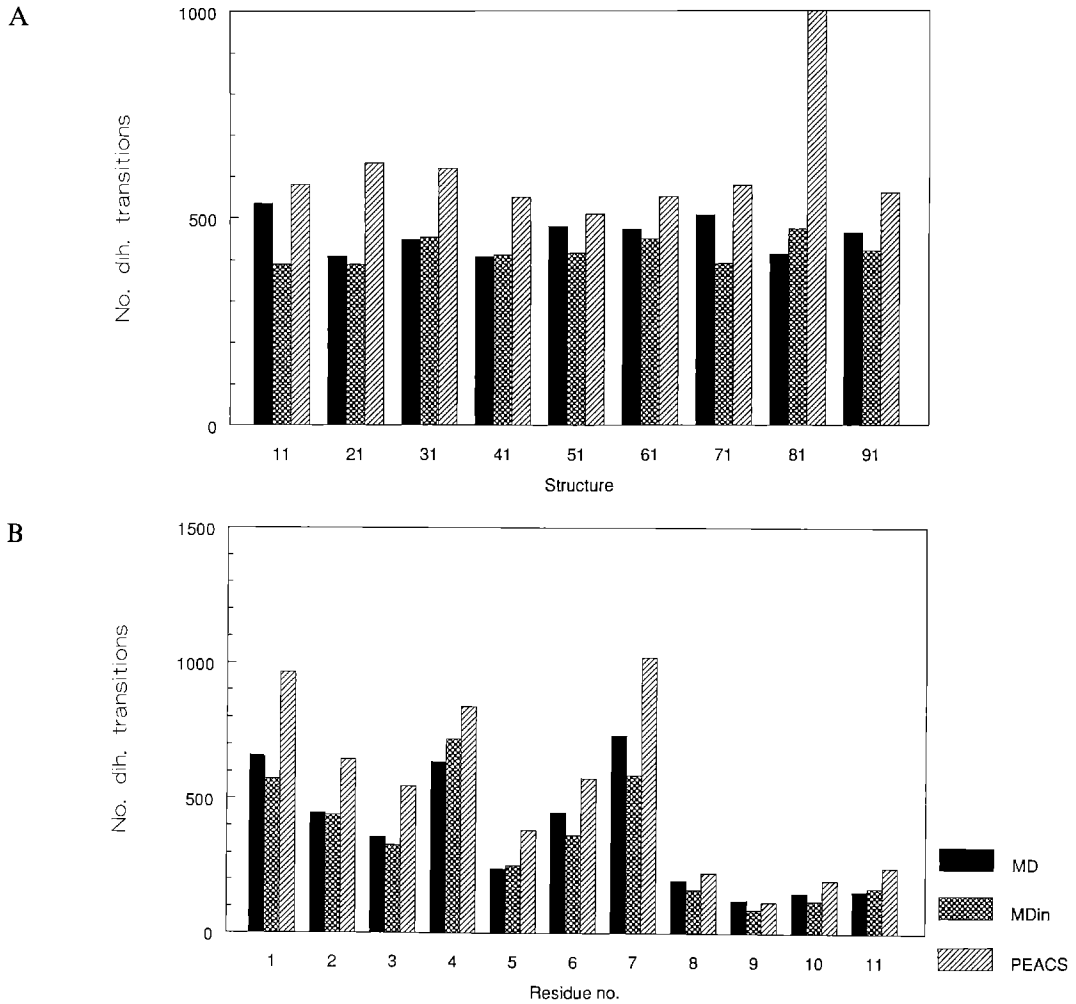


Fig. 4. Bar diagram with the number of backbone dihedral transitions as a function of structure number (A) and as a function of residue number (B). The dihedral transitions are calculated from the 40 ps trajectories from runs annealed from 1200 K.

makes the conventional approach of refining single structures using a quadratic penalty function on NOE distance-restraint violations basically wrong [28,29], because it may easily lead to structures with high internal strain in order to fulfill all measured NMR parameters simultaneously. One can successfully make use of time-dependent distance constraints to circumvent these problems [28]. However, this approach is only correct when (time) averaging over an ensemble of structures with roughly a Boltzmann distribution over the conformational states that are accessible at the temperature of the NMR experiments. The structures generated by PEACS are perfect starting structures for MD simulations with time-dependent distance restraints.

In summary, we conclude that the PEACS method is more useful than MD for refining model structures that are based on NMR data because it finds structures with lower potential energies

and lower sum of NOE distance violations, and it samples much more of the available conformational space.

ACKNOWLEDGEMENTS

We gratefully thank A. Torda for many helpful discussions. This work was carried out with the support of the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for Advancement of Pure Research (N.W.O.).

REFERENCES

1 Garey, M.R. and Johnson, D.S., Computers and Intractability: A guide to the Theory of NP-Completeness, Freeman, San Francisco, 1979.
2 Wiberg, K.B. and Boyd, R.H., J. Am. Chem. Soc., 94 (1972) 8426.
3 Lipton, M. and Still, W.C., J. Comput. Chem., 9 (1988) 343.
4 Griewank, A.O., J. Opt. Theor. Appl., 34 (1981) 11.
5 A recent monograph on Molecular Dynamics and Monte Carlo techniques is: Allen, M.P. and Tildesley, D.J., Computer Simulation of Liquids, Clarendon, Oxford, 1987.
6 Northrup, S.H. and McCammon, J.A., Biopolymers, 19 (1980) 1001.
7 DiNola, A., Berendsen, H.J.C. and Edholm, O., Macromolecules, 17 (1984) 2044.
8 Kirkpatrick, S., Gelatti Jr., C.D. and Vecchi, M.P., Science, 220 (1983) 671.
9 van Gunsteren, W.F., Kaptein, R. and Zuiderweg, E.R.P., In Olson, W.K. (Ed.) Nucleic Acid Conformation and Dynamics, Proceedings of a NATO/CECAM Workshop, CECAM, Orsay, 1984, pp. 79–92.
10 Brünger, A.T., Kuriyan, J. and Karplus, M., Science, 235 (1987) 458.
11 For a review see: Scheraga, H.A., Chemica Scripta, 29A (1989) 3.
12 Piela, L., Kostrowicki, J. and Scheraga, H.A., J. Phys. Chem., 93 (1989) 3339.
13 Nilges, M., Clore, G.M. and Gronenborn, A.M., FEBS Lett., 239 (1988) 129.
14 a. Crippen, G.M., J. Comput. Chem., 3 (1982) 471.
b. Crippen, G.M., J. Comput. Chem., 5 (1984) 548.
15 a. Purisma, E.O. and Scheraga, H.A., Proc. Natl. Acad. Sci. U.S.A., 83 (1986) 2782.
b. Purisma, E.O. and Scheraga, H.A., J. Mol. Biol., 196 (1987) 697.
16 Donnelly, R.A. and Rogers Jr., J.W., Int. J. Quantum Chem.: Quantum Chem. Symp., 22 (1988) 507.
17 Cotterill, R.M.J. and Madsen, J.U., In Characterizing Complex Systems, Bohr, H. (Ed.) World Scientific, Singapore, 1990, pp. 177–191.
18 Lautz, J., Kessler, H., Blaney, J.M., Scheek, R.M. and van Gunsteren, W.F., Int. J. Pept. Prot. Res., 33 (1989) 281.
19 Lautz, J., Kessler, H., Kaptein, R. and van Gunsteren, W.F., J. Comput.-Aided Mol. Design, 1 (1987) 219.
20 Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R., J. Chem. Phys., 81 (1984) 3684.
21 van Gunsteren, W.F. and Berendsen, H.J.C., Groningen Molecular Simulation (GROMOS) Library Manual, Biomos, Groningen, 1987.
22 Åqvist, J., van Gunsteren, W.F., Leijonmarck, M. and Tapia, O., J. Mol. Biol., 183 (1985) 461.
23 van Gunsteren, W.F. and Berendsen, H.J.C., In Hermans, J. (Ed.) Molecular Dynamics and Protein Structures, Polycrystal Books Service, Western Springs, IL., 1985, pp. 5–14.
24 Ryckaert, J.P., Ciccotti, G.H. and Berendsen, H.J.C., J. Comput. Phys., 23 (1977) 327.
25 van Gunsteren, W.F. and Berendsen, H.J.C., Mol. Phys., 34 (1977) 1311.
26 Loosli, H.R., Kessler, H., Oschkinat, H., Weber, H.P., Petcher, T.J. and Widmer, A., Helv. Chim. Acta, 68 (1985) 682.
27 Berendsen, H.J.C., van Gunsteren, W.F., Zwinderman, H.R. and Geurtsen, R.G., Ann. N.Y. Acad. Sci., 482 (1986) 264.
28 Torda, A.E., Scheek, R.M. and van Gunsteren, W.F., Chem. Phys. Lett., 157 (1989) 289.
29 Torda, A.E., Scheek, R.M. and van Gunsteren, W.F., J. Mol. Biol., 214 (1990) 223.