J. Mol. Biol. (1993) **234**, 751–762

A Structure Refinement Method Based on Molecular Dynamics in Four Spatial Dimensions

René C. van Schaik, Herman J. C. Berendsen

Department of Biophysical Chemistry
Rijksuniversiteit Groningen, Nijenborgh 4
9747 AG, Groningen, The Netherlands

Andrew E. Torda and Wilfred F. van Gunsteren

Laboratorium für Physikalische Chemie
ETH Zentrum, CH-8092 Zürich, Switzerland

(Received 5 February 1993; accepted 9 July 1993)

We have developed a method for structure refinement based on molecular dynamics in four spatial dimensions (4D-MD). The method was applied with success to the structure refinement of Cyclosporin A (CPA) and the *lac*-repressor headpiece (LAC) using atom-atom distance restraints derived from NMR data, two cases where conventional MD refinement methods failed. In the case of CPA we were able to refine seven out of nine substantially different structures, while conventional MD only refines three structures. Previously, it had appeared to be impossible to refine a given LAC structure with conventional MD methods, manual modifications were necessary in order to fulfil all the distance restraints. In this study we show that LAC can be refined without manual interference and using only 5000 steps (10 ps) of 4D-MD. The latter result is particularly interesting because it indicates that this method may be a very useful tool when modelling loops in proteins.

Keywords: structure refinement; molecular dynamics; NOE

1. Introduction

Since its introduction a few years ago (van Gunsteren *et al.*, 1985; Clore *et al.*, 1985; Brünger *et al.* 1987) molecular dynamics (MD†) has been the most important tool for refining molecular structures with respect to potential energy and consistency with experimental data. Unfortunately the method has limitations, and several improvements on the original MD method, such as temperature annealing (simulated annealing; Kirkpatrick *et al.*, 1983; Clore *et al.*, 1985) and potential energy annealing (PEACS; van Schaik *et al.*, 1992), have been proposed over the years, but all these methods suffer from the fact that the molecule may get trapped in local minima of the potential energy hyper surface, which do not necessarily correspond to structures that fulfil the experimental data set or structures with a low potential energy.

Introducing artificial degrees of freedom (that are later removed), is a less intuitive, but equally useful approach. This is based on the idea that by increasing the number of degrees of freedom of the system, the number of minima of the potential energy function is decreased and barriers that are present on the 3D potential energy hyper surface may be circumvented. The energy embedding and rotational energy embedding methods (Crippen, 1982, 1984, 1987; Crippen & Havel, 1990) take this to an extreme. Here, the molecular structure is first embedded in a N-dimensional space, where no local minima exist. A local minimizer takes the system to the high-dimensional global minimum and a projection to the next lower dimension is then performed. This cycle of minimization/projection is repeated until a structure with low energy in three dimensions is obtained. This can be done in different ways (Crippen, 1982, 1984, 1987; Crippen & Havel, 1990; Purisma & Scheraga, 1986, 1987). However, there is

† Abbreviations used: MD, molecular dynamics; PEACS, potential energy annealing conformational search; 3D, three-dimensional; 4D, four-dimensional; NOE, nuclear Overhauser effect; a.m.u., atomic mass unit; CPA, cyclosporin A; r.m.s., root-mean-square; EM, energy minimization; LAC, *lac*-repressor headpiece; NOESY, NOE spectroscopy.

0022-2836/93/230751-12 $08.00/0

© 1993 Academic Press Limited

no guarantee that low local minima in high-dimensional space do directly correspond to low local minima in three dimensions. Another drawback is that a potential energy function based on distances is used. Crippen (1987) showed that the energy embedding algorithm can be successfully applied to standard molecular mechanics force fields, while Purisma and Scheraga (1987) proposed a variation that also uses a potential energy function that incorporates both distance geometric and empirical force field terms. The potential energy function is minimized in distance space using a local minimization routine. During the energy minimization, certain penalty terms in the potential energy function are raised thereby forcing the system to relax in 3D. These methods require storage of the order $N^2/2$, and they lack the better search performance of dynamical methods.

In the context of distance geometry calculations, several groups have shown the utility of simply introducing one artificial dimension (Lautz *et al.*, 1989; van Nuland *et al.*, 1992; Weber *et al.*, 1988) and performing some kind of refinement in 4D-space. Our approach follows naturally from this previous work, we combine the search properties of molecular dynamics and the advantage of working in a space of a higher dimensionality. Again, the basic idea is that 3D atomic rearrangements are made easier by adding one extra spatial dimension. A requirement for the method to work properly is that low local minima in 4D-space should correspond directly with low local minima in 3D-space. The reduction to 3D-space is achieved by a slow dynamical back projection using a harmonic penalty function in the force field that penalizes non-zero co-ordinates in the 4th dimension. The weight of the penalty function is gradually increased relative to the potential energy function, thereby forcing the system to go back into 3D (Blaney & Crippen, 1986).

2. Methods and Materials

(a) *The 4D-space*

The 4D linear vector space is described by a set of 4 linearly independent basis vectors, $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ and $\mathbf{e}_w$, and an arbitrary vector $\mathbf{r}$ in this 4D-space is represented in terms of its 4 components:

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z + w\mathbf{e}_w. \tag{1}$$

When the basis set is taken to be orthonormal ($\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$) the scalar product between vectors $\mathbf{r}$ and $\mathbf{r}'$ is defined by:

$$\mathbf{r}\mathbf{r}' = xx' + yy' + zz' + ww', \tag{2}$$

and the length of vector $\mathbf{r}$, or more commonly the norm of vector $\mathbf{r}$ is given by:

$$\mathbf{r} = |\mathbf{r}| = (\mathbf{r}\cdot\mathbf{r})^{\frac{1}{2}} = \sqrt{x^2+y^2+z^2+w^2}. \tag{3}$$

This expression shows that any term in the potential energy function that is only a function of atom–atom distances can easily be generalized to 4 dimensions by using eqn (3) to evaluate a distance in 4D-space. This is an important property because almost all terms of standard force fields can be expressed in terms of distances between particles (Crippen, 1987). The distance $\mathbf{r}_{ij}$ between two particles is simply calculated as the norm of the difference in position vectors, $\mathbf{r}_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Further, we can obtain the angle γ between two vectors $\mathbf{r}_{ij}$ and $\mathbf{r}_{kj}$ using the scalar product and the norm of vectors *via*:

$$\cos(\gamma) = \frac{\mathbf{r}_{ij}\cdot\mathbf{r}_{kj}}{\mathbf{r}_{ij}\mathbf{r}_{kj}}. \tag{4}$$

(b) *The 4D force field*

The 4D potential energy function (eqn (5)) is very similar to the normal 3D potential energy function. It is composed of terms representing bond-stretching, bond-angle bending, harmonic improper dihedral (out-of-plane, out-of-tetrahedral configuration) bending, sinusoidal torsion, van der Waals, electrostatic and harmonic 3D-projection interactions. The latter term was added to the force field in order to be able to project the system back from 4D into 3D-space (Blaney & Crippen, 1986). This is achieved by harmonic coupling of each particle to the $w = 0$ plane with a force constant of size K_{3D}. When K_{3D} is increased during the refinement, the system is forced to go back to the 3D sub-space. Further, an extra term was added to the force field in order to incorporate experimental distance restraints originating from nuclear Overhauser effects (NOEs) obtained from NMR spectroscopy (van Gunsteren *et al.*, 1985).

$$\begin{aligned} V_{4D}(\{\mathbf{r}_i\}) = & \sum_{\text{all bonds}}^{N_{\text{bonds}}} \tfrac{1}{2}K_b(b-b_0)^2 \\ & + \sum_{\text{all angles}}^{N_{\text{angles}}} \tfrac{1}{2}K_\theta(\theta-\theta_0)^2 \\ & + \sum_{\text{improper dihedrals}}^{N_{\text{improper}}} \tfrac{1}{2}K_\xi(\xi-\xi_0)^2 \\ & + \sum_{\text{torsional angles}}^{N_{\text{dihedral}}} K_\varphi[1+\cos(n\varphi-\delta)] \\ & + \sum_{i<j}^{N_{\text{nonbon}}} \left[\frac{C_{12}(i,j)}{r_{ij}^{12}} - \frac{C_6(i,j)}{r_{ij}^{6}} + \frac{q_iq_j}{4\pi\varepsilon_0\varepsilon_r r_{ij}}\right] \\ & + \sum_{i=1}^{N_{\text{at}}} \tfrac{1}{2}K_{3D}w_i^2 + \sum_{\text{distance restraints}}^{N_{\text{dr}}} \tfrac{1}{2}K_{\text{dr}}(r_{ij}-r_{ij}^0)^2 \end{aligned} \tag{5}$$

Using the equations given in the Appendix, the implementation of the potential energy function eqn (5) and its negative gradient (forces on the particles) is straightforward. Only the implementation of the improper dihedral angle term and distance restraining term resulting from stereo-specific NOE assignments involving virtual atoms needs some special attention. Due to their chiral properties it appeared to be impossible to properly evaluate these terms in 4D. Therefore they are evaluated using the 3D-projection of the particles involved (see the Appendix).

(c) *The equations of motion in 4D-space*

We assume that Newtons equations of motion are valid in 4D-space:

$$\frac{d^2\mathbf{r}_i(t)}{dt^2} = -\frac{1}{m_i}\frac{\partial V_{4D}(\{\mathbf{r}_j\})}{\partial \mathbf{r}_i}. \tag{6}$$

The mass of particle i is denoted by m_i, its 4D-position vector by $\mathbf{r}_i$ and t is the time.

(d) *Temperature coupling*

Two "temperatures" were defined for our 4D-system, the normal 3D-temperature (eqn (7)), and an additional w-temperature (eqn (8)).

$$T_{3D} = \frac{2}{(3N - N_{cons})k_B} \sum_{i=1}^{N} \tfrac{1}{2} m_i (v_x^2 + v_y^2 + v_z^2)_i, \quad (7)$$

$$T_w = \frac{2}{Nk_B} \sum_{i=1}^{N} \tfrac{1}{2} m_i v_{w_i}^2, \quad (8)$$

where N is the number of particles, N_{cons} is the number of constraints, k_B is Boltzmann's constant, and v_x, v_y, v_z and v_w are the velocity components of the particles. By defining 2 temperatures we are able to control the magnitude of the velocities in the 4th dimension almost independently from the normal 3D-velocities. This allows us to decrease the w-velocities during the back projection period independently from the xyz-velocities. When forcing the system out of the 4th dimension by increasing K_{3D}, the kinetic energy of the 4th dimension must be reduced to zero simultaneously. Both temperatures are weakly coupled to a temperature bath using the Berendsen thermostat (Berendsen *et al.*, 1984):

$$\frac{dT(t)}{dt} = \frac{1}{\tau_T}[T_0 - T(t)], \quad (9)$$

where T_0 is the reference temperature, and τ_T the temperature relaxation constant. These 2 values may be different for both temperatures T_{3D} and T_w.

(e) *Projection optimization*

Before starting the back projection and at the end of the back projection, the structure is rotated such that the 3D-projection is optimal. This avoids high strain as a result of the 3D-projection potential energy term at the beginning of the back projection, and bad van der Waals contacts at the beginning of the 3D-energy minimization. The rotation that gives an optimal 3D-projection is calculated using the inertia tensor (Crippen, 1982):

$$\tilde{\mathbf{I}} = \sum_{i=1}^{N} (\mathbf{r}_i^2 \tilde{\mathbf{I}} - \mathbf{r}_i \mathbf{r}_i^+). \quad (10)$$

This is the inertia tensor of a system of particles of unit mass. The inertia tensor is diagonalized such that the eigenvalues appear in ascending order on the diagonal:

$$\tilde{\mathbf{I}} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix} \text{with } \lambda_1 < \lambda_2 < \lambda_3 < \lambda_4. \quad (11)$$

and the matrix of corresponding eigenvectors (column-vector Λ_i reads:

$$\tilde{\mathbf{R}} = (\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4). \quad (12)$$

When the transpose of this matrix is used to perform a 4D-rotation of the structure (principal axis transformation),

$$\mathbf{r}_i' = \tilde{\mathbf{R}}^T \mathbf{r}_i, \quad (13)$$

and for the velocities

$$\mathbf{v}_i' = \tilde{\mathbf{R}}^T \mathbf{v}_i, \quad (14)$$

the inertia tensor will be diagonal with the eigenvalues in ascending order. This means that the total 4D-component of the structure $\{\mathbf{r}_i'\}$ is minimal. The 4D-rotation may result in a structure with the wrong chirality. In order to obtain the conformer with the correct chirality, the improper dihedral energy is calculated for the projected structure and its xy-plane reflection mirror image. The improper dihedral energy is sensitive to the chirality of the system, and by selecting the structure with the lowest improper dihedral energy, the structure with the correct overall chirality is obtained. In practice the excursions into the 4th dimension are small enough to guarantee also correct local chirality.

3. Experiments and Results

All structure refinement studies were carried out using software from the GROMOS suite of programs (van Gunsteren & Berendsen, 1987), with the adaptations to the original GROMOS87 potential energy function as described above. Since we are not interested in correct dynamics of the system, the mass of hydrogens in the system was set to 10 a.m.u., allowing a time step size of two femtoseconds without numerical problems in the leap frog integration algorithm.

(a) *Refinement of cyclosporin A*

For the refinement of cyclosporin A, 57 atom-atom distance restraints originating from nuclear Overhauser effects (NOEs) obtained from NMR spectroscopy (Loosli *et al.*, 1985; Lautz *et al.*, 1987, 1989) were included. Recently, the NOEs of cyclosporin A have been measured again, which resulted in a slightly different set of atom–atom distance constraints (Kessler *et al.*, 1990). In order to be able to compare our results with those of earlier refinement studies (Lautz *et al.*, 1989; van Schaik *et al.*, 1992) we use the same set of distance constraints as used by van Schaik *et al.* (1992). A force constant of 9000 kJ·mol^{-1}·nm^{-2} for the harmonic distance restraining term was used. Non-bonded interactions were evaluated using a twin-range cut-off criterion. Short range interactions (<0·8 nm) were calculated every time step using a pairlist, while long-range coulombic interactions (<1·2 nm) were calculated every tenth step when the pairlist was updated. Lautz *et al.* (1989) repeatedly applied the distance geometry program DGEOM to CPA using the original set of 57 NOE distance constraints, which resulted in 27 structures. The mutual root-mean-square (r.m.s.) 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 a starting structure for the various refinement runs discussed below. We chose the following protocol. The first 5000 steps of 4D-MD were performed while cooling the system linearly down from 1200 K to 300 K, with temperature relaxation constants $\tau_{T_{3D}} = \tau_{T_w} = 0{\cdot}01$ ps. The 3D-projection force constant K_{3D} was kept fixed at 100 kJ·mol^{-1}·nm^{-2}. Before starting the back projection the structure with optimal 3D-projection was obtained by a 4D-rotation using eqns (13) and (14). The back projection was performed by

Table 1

Results of the refinement of Cyclosporin A, using 4D-MD, MD and PEACS

	4D-MD annealed from 1200 K			MD annealed from 1200 K[a]			PEACS annealed from 1200 K[a]		
Structure[b]	E_{pot}[c] (kJ/mol)	Σ_{viol}[d] (Å)	Δr_{MDS1}[e] (Å)	E_{pot} (kJ/mol)	Σ_{viol} (Å)	Δr_{MDS1} (Å)	E_{pot} (kJ/mol)	Σ_{viol} (Å)	Δr_{MDS1} (Å)
xclass11	−39	0·5	0·4	57	3·2	3·1	0	1·4	2·0
xclass21	−70	0·2	0·3	51	4·0	1·7	−65	0·9	0·3
xclass31	17	1·5	1·6	35	3·9	2·9	−66	1·1	0·4
xclass41	−52	0·5	0·4	−11	0·9	1·8	−26	1·2	1·9
xclass51	−25	1·8	0·4	73	3·6	2·8	−65	1·0	0·6
xclass61	−62	0·5	0·6	73	4·5	3·5	−8	2·0	1·9
xclass71	134	4·5	2·3	−1	3·2	1·9	27	1·1	1·3
xclass81	−47	0·8	0·4	31	2·5	2·5	150	1·2	2·1
xclass91	−75	0·8	0·2	−82	1·4	0·3	−50	2·5	0·5

[a] From van Schaik *et al.* (1992).
[b] The names of the structures are taken from Lautz *et al.* (1989).
[c] Potential energy of the system after energy minimization. Distance restraint energy is not included.
[d] Sum of all distance violations in excess of the reference NOE distances.
[e] r.m.s. positional differences of the C^{α} atoms with respect to the MDS1 structure.

increasing K_{3D} linearly from 100 to 10,000 kJ · mol^{-1} · nm^{-2} in 15,000 steps. At the same time T_{3D} was decreased linearly from 300 K to 150 K and T_w from 300 K to 0 K, with $\tau_{T_{3D}} = \tau_{T_w} = 0{\cdot}002$ ps. Before starting energy minimization (EM) the optimal 3D-projection was calculated. The resulting structures were subjected to 500 steps of conjugate gradients energy minimization in 3D-space, whereby no cut-off on the evaluation of the non-bonded interactions was applied. The results of these refinements are shown in Table 1. For comparison, the results of MD refinement in 3D and PEACS refinement (van Schaik *et al.*, 1992) are also given in this Table. Table 1 clearly shows the improved performance of our 4D-refinement method over standard MD-refinement. Not only the energies are lower, also the sum of all distance violations in excess of the reference NOE distances is lower and the refined structures tend to converge to the previously found solution indicated by the symbol MDS1 (Lautz *et al.*, 1987). This convergence is also shown in Figure 1. The variation between the structures at the bottom of the molecules (residues 2Abu and 3Sar) is a reflection of the relatively low number of distance constraints to atoms in this part of the molecule. When no cut-off on the evaluation of non-bonded forces was applied, it was possible to refine structures xclass31 and xclass71 with 4D-MD as well. However, in that case the overall performance of the algorithm is slightly worse (6 structures refined, and 3 having high energies). In general, the 4D-MD algorithm performs even better than the PEACS algorithm: especially the sum of violations is lower and the new structures are closer to the MDS1 structure. Some minima found by PEACS (xclass11, 41 and 61) were not found with the

Table 2

Results of the refinement of Cyclosporin A, using 4D-MD with different back projections

Structure[a]	Before back projection Σ_{viol}[b] (Å)	Direct energy minimization (EM) E_{pot}[c] (kJ/mol)	Dynamically in 15,000 steps followed by EM E_{pot} (kJ/mol)	Dynamically in 45,000 steps followed by EM E_{pot} (kJ/mol)
xclass11	0·8	69	−39	−30
xclass21	0·5	−17	−70	−45
xclass31	1·5	235	17	35
xclass41	0·2	−21	−52	−59
xclass51	0·7	+32	−25	−26
xclass61	0·5	−31	−62	−48
xclass71	1·9	497	134	75
xclass81	0·8	126	−47	−59
xclass91	0·4	−59	−75	−57

[a] The names of the structures are taken from Lautz *et al.* (1989).
[b] Sum of all distance violations in excess of the reference NOE distances, calculated in 4D-space.
[c] Potential energy of the system after energy minimization. Distance restraint energy is not included.

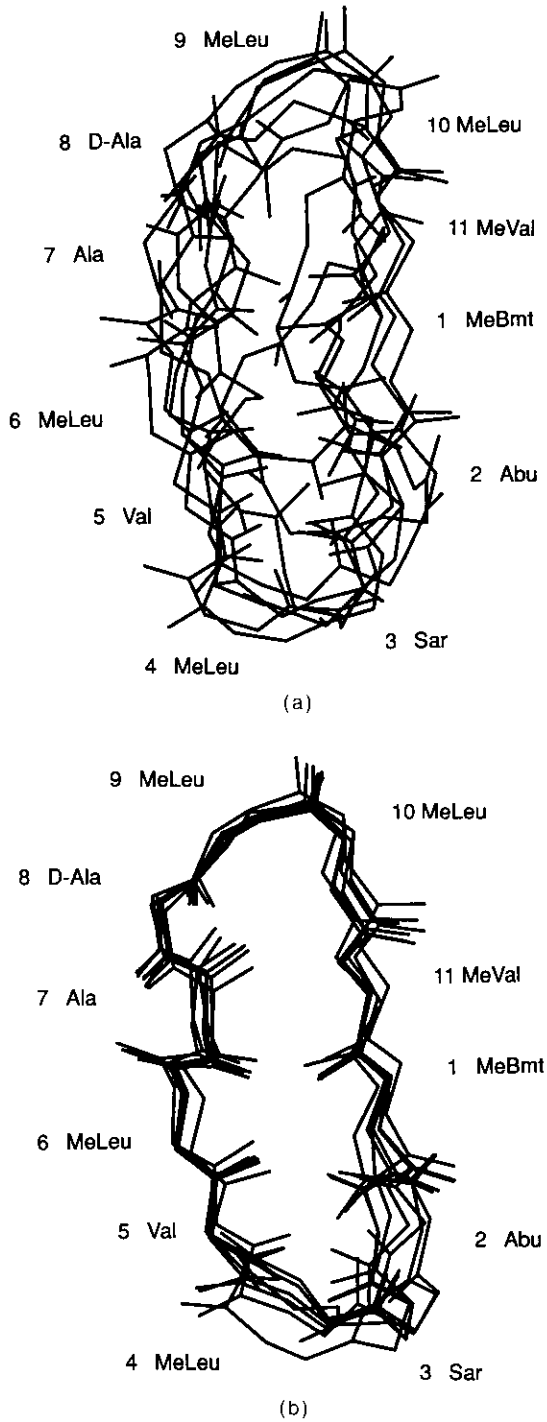


Figure 1. (a) Distance geometry structures. (b) Structures after 4D-MD refinement. Only structures of xclass11, 21, 41, 51, 61, 81 and 91 are shown.

4D-MD method, and are probably local minima in which the PEACS method occasionally gets trapped.

(b) *Slow back projection*

In order to investigate the effect of the back projection rate on the results, we performed two additional refinement studies. As a starting point of these runs the structures after the first 5000 4D-MD steps were taken. In the first run we performed an energy minimization in 4D-space, during which K_{3D} was increased from 100 to 10,000 kJ·mol^{-1}·nm^{-2}. The second run was almost identical to the original one, except that back projection now was carried out over 45,000 steps instead of the original 15,000 steps. The results are shown in Table 2. Direct energy minimization is clearly inferior to dynamic back projection. However, taking more steps during the back projection does not always guarantee a better result. This originates from the fact that the last structure of a 4D-MD back projection run does not necessarily lead to a structure with the lowest energy after energy minimization. The use of temperature annealing will redress this problem and will yield a structure with low energy after EM. During the back projection period every 250 steps a structure was stored, resulting in total in 180 structures. Each structure was 3D-energy minimized, and the potential energy of the system as a function of the number of 4D-MD refinement steps is plotted in Figure 2. These graphs indicate that the energy of the system is fairly stable after 15,000 steps, and that for runs leading to a refined structure performing back projection over 5000 steps already leads to a reasonably stable energy. We conclude from this that using about 5000 steps is sufficient for a good dynamic back projection and that using 45,000 steps is overkill. Another important observation that can be made from Table 2 is that the two structures that have a high energy at the end of the back projection also have a high sum of violations at the beginning of the back projection. We conclude from this that it is necessary to start the back projection with a structure that has a low sum violations, which means that the most important structure rearrangements should take place before the back projection starts.

(c) *LAC*

For refinement of the *lac*-repressor headpiece a similar approach was used. As a starting point we took the energy minimized model described by de Vlieg *et al.* (1986) from just before their manual modification. A set of 169 distance restraints (Zuiderweg *et al.*, 1985) was used in order to force the molecule to satisfy NMR data originating from NOESY spectra, with a distance restraining force constant of 4000 kJ·mol^{-1}·nm^{-2}. Non-bonded interactions were evaluated using a twin-range cut-off criterion. Short range interactions (<0·8 nm) were calculated every time step using a pairlist, while long-range coulombic interactions (<1·2 nm) were calculated every tenth step when the pairlist was updated. The first 1250 steps were performed at constant temperature of 300 K, with $\tau_{T_{3D}} = \tau_{T_w} =$ 0·01 ps., and $K_{3D} = 0{\cdot}0$ kJ·mol^{-1}·nm^{-2}. Again at the beginning of the back projection the structure was rotated such that the structure with optimal

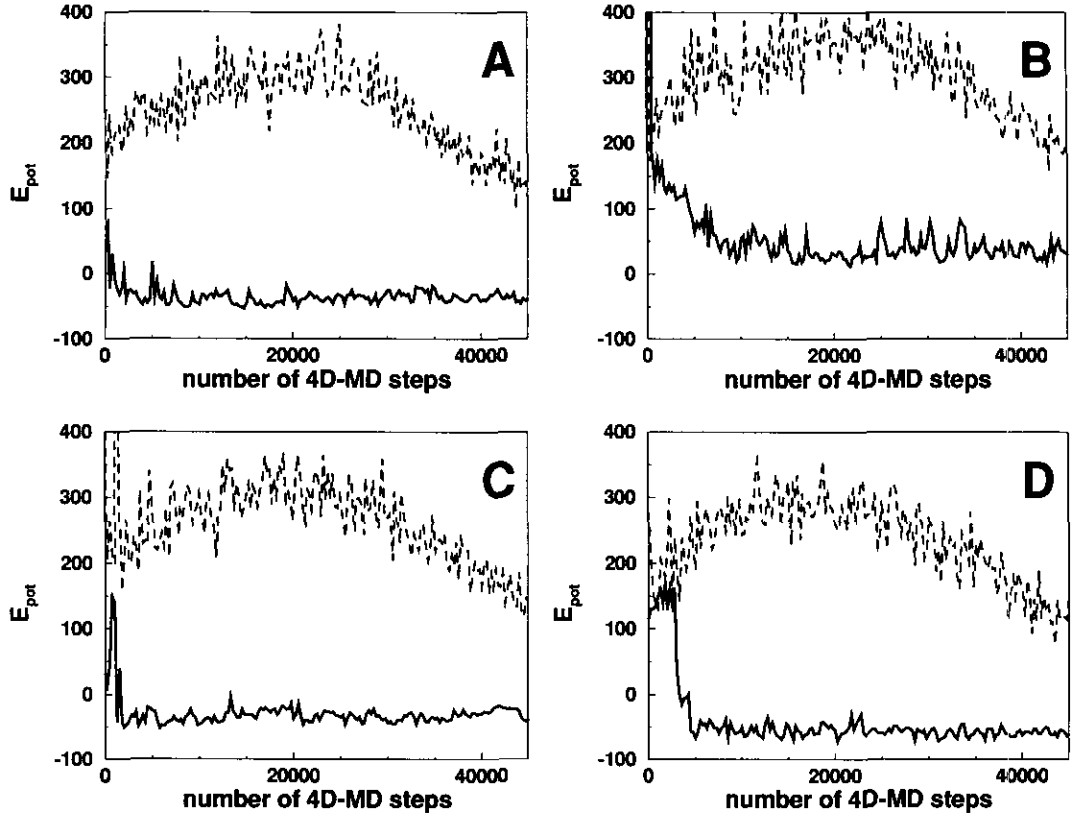


Figure 2. Potential energy of the system (in kJ/mol) during the 4D-MD back projection. Broken line, before energy minimization; continuous line, after 3D-energy minimization. A, xclass11; B, xclass31; C, xclass51; D, xclass81.

3D-projection was obtained. The back projection was performed in 3750 steps, by increasing K_{3D} from 10 to 5000 kJ · mol^{-1} · nm^{-2}. At the same time T_w is decreased from 300 to 0 K and T_{3D} kept at 300 K, with $\tau_{T_{3D}} = 0{\cdot}005$ ps., and $\tau_{T_w} = 0{\cdot}01$ ps. After projection optimization of the last structure, the resulting structure was 3D-energy minimized. The results are shown in Table 3. As in the case of CPA, the structure obtained by 4D-MD fulfils the NMR data set better, and is lower in energy than the 3D-MD refined or manually refined ones. We note that the results obtained with 4D-MD took only 5000 steps, while 3D-MD took 15,000 steps and the combined model building and a 3D-MD approach took 20,000 steps. In the original refinement by de Vlieg *et al.* (1986) it appeared to be necessary to make manual modifications to the structure (loop region residues 26 to 33), because MD (at 300 K, $K_{dr} = 250$ kJ · mol^{-1} · nm^{-2}) was not able to surmount the barrier corresponding with this loop rearrangement. When the temperature was increased to 1200 K, and, K_{dr} to 9000 kJ · mol^{-1} · nm^{-2}, MD could surmount this barrier (van Schaik, unpublished results), however, the resulting structure has still a large sum of NOE violations. By allowing the system to go into 4D-space the loop transition can be performed easily. It is no longer necessary to pass over the (3D) barrier because there exists a 4D-path around the barrier. The resulting loop structure is shown in Figure 3. The new loop structure is quite different from the loop structure obtained by de Vlieg *et al.* (1986), apparently the number of distance restraints is insufficient to determine the loop structure unambiguously.

Table 3

Results of the refinement of the lac*-repressor headpiece, using 4D-MD, and a combined model building/MD approach*

Method	E_{pot}[a] (kJ/mol)	Σ_{viol}[b] (Å)	Δr_{lac}[c] (Å)
Initial structure[d]	−3128	32·4	1·8
After 4D-MD[e]	−3197	1·8	1·7
After MD[f]	−3022	7·0	2·0
After model building + MD[g]	−3122	5·8	0·0

[a] Potential energy of the system after energy minimization. Distance restraint energy is not included.

[b] Sum of all distance violations in excess of the reference NOE distances.

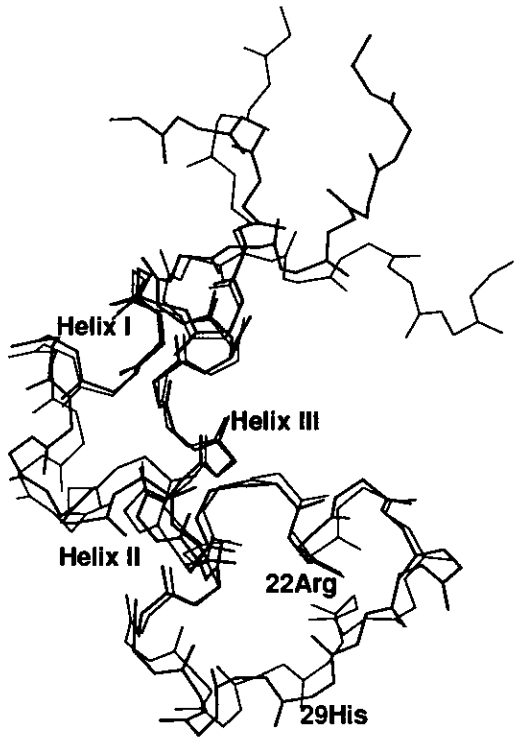
[c] r.m.s. positional differences of the C$^{\alpha}$ atoms with respect to the *lac*-repressor headpiece structure (de Vlieg *et al.*, 1986).

[d] Energy minimized starting structure from de Vlieg *et al.* (1986), from just before their manual modification.

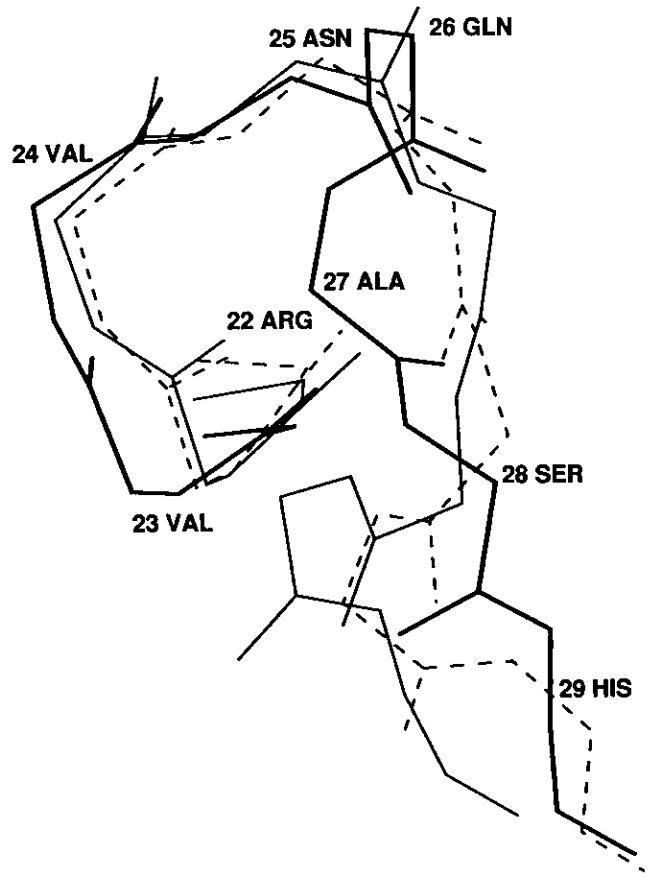
[e] See text.

[f] 15,000 steps MD, using temperature annealing from 1200 K (van Schaik, unpublished results).

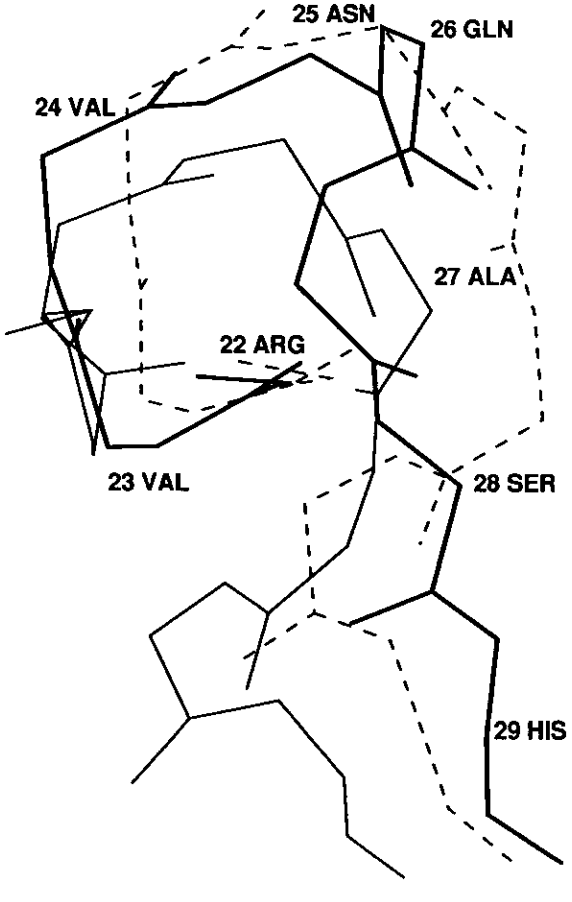
[g] Manual modification + 20,000 steps of MD at 300 K, de Vlieg *et al.* (1986).



(a)



(c)



(b)

Figure 3. (a) Back bone structure of the *lac*-repressor headpiece. Continuous line, 4D-MD refined structure; thick continuous line, reference structure from de Vlieg *et al.* (1986). (b) Loop structure of the *lac*-repressor headpiece residues 22 to 29. Broken line, starting structure; continuous line, 3D projection of the structure before back projection; thick continuous line, reference structure from de Vlieg *et al.* (1986). (c) Loop structure of the *lac*-repressor headpiece residues 22 to 29. Broken line, MD refined structure; continuous line, 4D-MD refined structure; thick continuous line, reference structure from de Vlieg *et al.* (1986).

4. Discussion

The refinement procedure presented here is a useful tool for refinement of structures based on NMR data, especially when the initial structures are such that large conformational changes have to be made in order to fulfil the NMR data set. The examples given here show that 4D-MD is a more efficient search method than 3D-MD or the PEACS algorithm. It would be interesting to investigate whether force field modifications such as soft-core or reduced van der Waals interactions (Levitt, 1983; Nilges *et al.*, 1988) deliver similar benefits as 4D-MD. Although we have no direct comparison with very high-dimensional approaches, (energy embedding, Crippen, 1982, 1984, 1987; rotational embedding, Crippen & Havel, 1990; relaxing dimensionality, Purisma & Scheraga, 1986, 1987),

some points are worth noting. First, energy embedding requires storage of $O(N^2)$ whereas the 4D-MD method is still $O(N)$. This is a consideration when treating very large structures. 4D-MD also has a potential advantage in computational speed compared with higher dimensional spaces when using non-bonded forces and a distance based cut-off criterion. Going from 3D to 4D structure refinements, we experienced a speed penalty of 1·3 to 1·7. This is partly because we are calculating with $4N$ rather than $3N$ variables, but also because more particles pack within a given distance as dimensionality increases. For example, one can pack six discs around a central disc in 2D, but 12 spheres around a central sphere in 3D. It is possible to conceive of conformations in very high dimensional spaces where each particle is within the interaction distance of every other particle.

It is also interesting to contrast the kind of results one can expect from this 4D-MD and known n-D approaches. Although it may not be appropriate to refer to a trajectory, one can certainly refer to the path taken by a system during energy embedding. Even if one has a numerically perfect minimizer, there will still be bifurcations in the path at each projection step. For a given projection scheme, however, the method will produce one result for a specified system. In contrast, 4D-MD still produces a time based trajectory and, in principle, will benefit from calculations of increasing length. It also has the ability to produce an ensemble of structures rather than a single answer.

4D-MD structure refinement also has a natural connection to metric matrix based distance geometry calculations. Presently, one usually generates 3D co-ordinates by extracting the eigenvectors with the three largest eigenvalues from the metric matrix. These co-ordinates are then refined in 3D. This scheme involves a truncation of information since a metric matrix constructed from experimental information will rarely correspond to a set of co-ordinates embedded in 3D (Crippen & Havel, 1988). A more attractive approach may be to generate 4D-co-ordinates directly for 4D refinement and slow back projection to 3D rather than abrupt truncation (Havel, 1991; Weber *et al.*, 1988). The advantage of the dynamical back projection is clear, the structures obtained are much better, resulting in low energies and low sums of violations. The 4D-MD method may also prove to be very useful when modelling loops in proteins. In the refinement of the *lac*-repressor headpiece a rather large loop transformation was made quite easily.

Appendix

Implementation of the Four-dimensional Force Field

In this Appendix the potential energy function is given (van Gunsteren & Berendsen, 1987) and it is shown how to calculate the four-dimensional forces (Crippen, 1987).

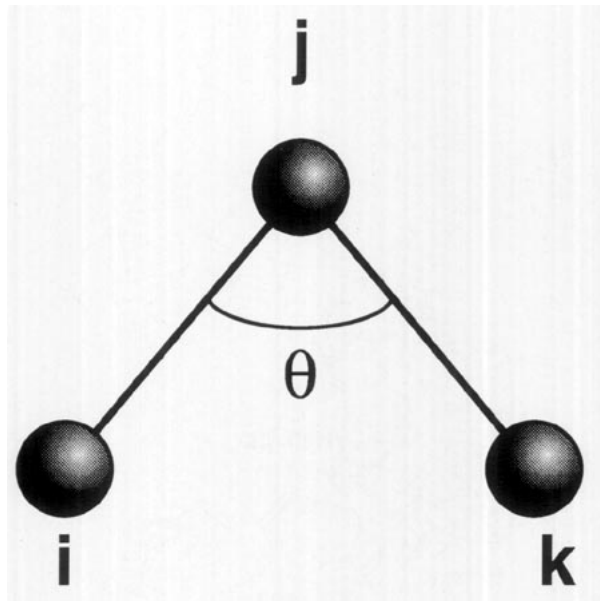


Figure A1.

(a) *Bond stretch potential energy term*

For the bond between particles i and j the bond stretch potential energy term is defined as:

$$E(b) = \tfrac{1}{2}K_b(b-b_0)^2, \tag{A1}$$

where $b = r_{ij} = (\mathbf{r}_{ij}\cdot\mathbf{r}_{ij})^{1/2}$ and $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. The vector $\mathbf{r}$ can be either three-dimensional with components (x, y, z) or four-dimensional with components (x, y, z, w). The forces $\mathbf{f}$ can be computed using:

$$\begin{aligned}\mathbf{f}_i &= -\frac{\partial E(b)}{\partial \mathbf{r}_i} = -\frac{dE}{db}\frac{\partial b}{\partial \mathbf{r}_i} = -K_b(b-b_0)\frac{\mathbf{r}_{ij}}{b}\\ \mathbf{f}_j &= -\frac{\partial E(b)}{\partial \mathbf{r}_j} = -\frac{dE}{db}\frac{\partial b}{\partial \mathbf{r}_j} = K_b(b-b_0)\frac{\mathbf{r}_{ij}}{b}.\end{aligned} \tag{A2}$$

(b) *Bond angle bending potential energy term*

The bond angle between particles i, j and k (see Fig. A1) can be computed in terms of the lengths of four-dimensional vectors using the law of cosines:

$$\begin{aligned}&r_{ik}^2 = r_{ij}^2 + r_{kj}^2 - 2r_{ij}r_{kj}\cos(\theta)\\ &\Rightarrow \theta = \arccos\left(\frac{r_{ij}^2 + r_{kj}^2 - r_{ik}^2}{2r_{ij}r_{kj}}\right) = \arccos\left(\frac{\mathbf{r}_{ij}\cdot\mathbf{r}_{kj}}{r_{ij}r_{kj}}\right).\end{aligned} \tag{A3}$$

The potential energy is given by:

$$E(\theta) = \tfrac{1}{2}K_\theta[\theta-\theta_0]^2, \tag{A4}$$

and the forces by:

$$\begin{aligned}\mathbf{f}_m &= -\frac{\partial E(\theta)}{\partial \mathbf{r}_m} = -\frac{dE(\theta)}{d\theta}\frac{d\theta}{d\cos(\theta)}\frac{\partial\cos(\theta)}{\partial \mathbf{r}_m}\\ &= \frac{K_\theta[\theta-\theta_0]}{\sqrt{1-\cos^2(\theta)}}\frac{\partial\cos(\theta)}{\partial \mathbf{r}_m}, \quad m = i, j, k.\end{aligned} \tag{A5}$$

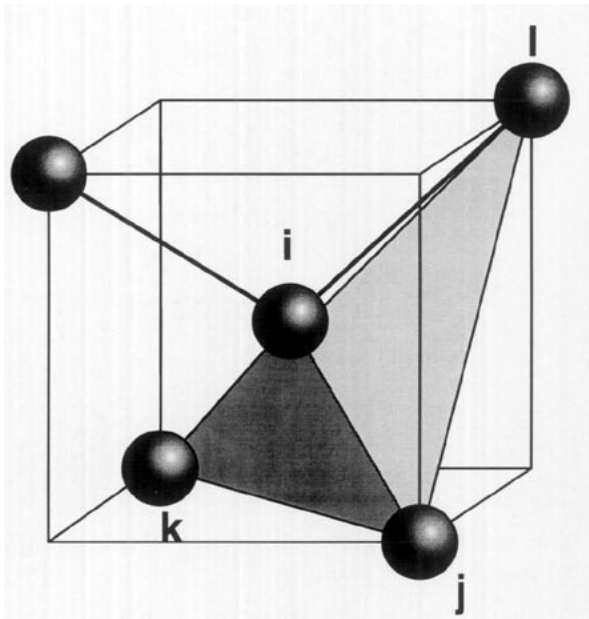


Figure A2.

Using eqn (A3) we obtain:

$$\frac{\partial \cos(\theta)}{\partial \mathbf{r}_i} = \frac{r_{ij}^2 \mathbf{r}_{kj} - (\mathbf{r}_{ij} \cdot \mathbf{r}_{kj}) \mathbf{r}_{ij}}{r_{ij}^3 r_{kj}},$$

$$\frac{\partial \cos(\theta)}{\partial \mathbf{r}_k} = \frac{r_{kj}^2 \mathbf{r}_{ij} - (\mathbf{r}_{ij} \cdot \mathbf{r}_{kj}) \mathbf{r}_{kj}}{r_{ij} r_{kj}^3}, \qquad \text{(A6)}$$

$$\frac{\partial \cos(\theta)}{\partial \mathbf{r}_j} = -\frac{\partial \cos(\theta)}{\partial \mathbf{r}_i} - \frac{\partial \cos(\theta)}{\partial \mathbf{r}_k}.$$

The expressions for the forces are numerically instable for $\theta \to 0, \pi$. This problem can be easily solved by defining $\mathbf{f}_m = 0$ for $\theta = 0, \pi$ and setting a maximum on $(1 - \cos(\theta))^{-1/2}$.

(c) *Improper dihedral angle potential energy term*

The improper dihedral angle is defined in order to keep structures in planar or tetrahedral conformation (see Figs A2 and A3). This is accomplished by applying a harmonic potential energy term to the angle ξ between the planes *i*-*j*-*k* and *j*-*k*-*l*:

$$E(\xi) = \tfrac{1}{2} K_\xi [\xi - \xi_0]^2. \qquad \text{(A7)}$$

In this formula we need the sign of ξ, which cannot be obtained from scalar products. The sign of ξ plays an important role in the definition of the tetrahedral conformation, because it defines the chirality of the group of atoms considered. Since a set of four points (e.g. a tetrahedron) has no well-defined chirality in four dimensions, there is no way to distinguish between the two forms and hence we cannot obtain the sign of ξ. We have therefore chosen to evaluate the improper dihedral only in *xyz*-space, i.e. the 3D-projection of the molecule should have the correct chirality. The improper dihedral interaction can now be evaluated in the usual way. Let us define two extra vectors:

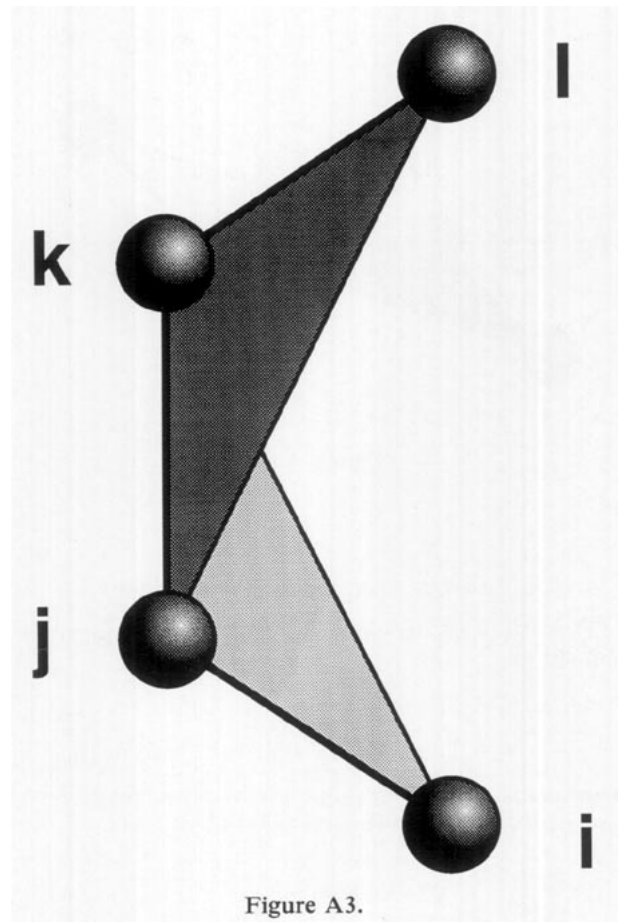


Figure A3.

$$\mathbf{r}_{mj} = \mathbf{r}_{ij} \times \mathbf{r}_{kj}$$
$$\mathbf{r}_{nk} = \mathbf{r}_{kj} \times \mathbf{r}_{kl}. \qquad \text{(A8)}$$

Using these two vectors we can obtain ξ *via*:

$$\xi = \arccos\left(\frac{\mathbf{r}_{mj} \cdot \mathbf{r}_{nk}}{r_{mj} r_{nk}}\right). \qquad \text{(A9)}$$

The forces on particle *i*, *j*, *k* and *l* can be derived from:

$$\mathbf{f}_q = -\frac{dE(\xi)}{d\xi} \frac{\partial \xi}{\partial \mathbf{r}_q} = -K_\xi [\xi - \xi_0] \frac{\partial \xi}{\partial \mathbf{r}_q}, \quad q = i, j, k, l. \qquad \text{(A10)}$$

Using eqn (A9) one can obtain (H. Bekker, personal communication):

$$\frac{\partial \xi}{\partial \mathbf{r}_i} = \frac{r_{kj}}{r_{mj}^2} \mathbf{r}_{mj},$$

$$\frac{\partial \xi}{\partial \mathbf{r}_l} = -\frac{r_{kj}}{r_{nk}^2} \mathbf{r}_{nk},$$

$$\frac{\partial \xi}{\partial \mathbf{r}_j} = \left(\frac{(\mathbf{r}_{ij} \cdot \mathbf{r}_{kj})}{r_{kj}^2} - 1\right) \frac{\partial \xi}{\partial \mathbf{r}_i} - \frac{(\mathbf{r}_{kl} \cdot \mathbf{r}_{kj})}{r_{kj}^2} \frac{\partial \xi}{\partial \mathbf{r}_l}, \qquad \text{(A11)}$$

$$\frac{\partial \xi}{\partial \mathbf{r}_k} = \left(\frac{(\mathbf{r}_{kl} \cdot \mathbf{r}_{kj})}{r_{kj}^2} - 1\right) \frac{\partial \xi}{\partial \mathbf{r}_l} - \frac{(\mathbf{r}_{ij} \cdot \mathbf{r}_{kj})}{r_{kj}^2} \frac{\partial \xi}{\partial \mathbf{r}_i}.$$

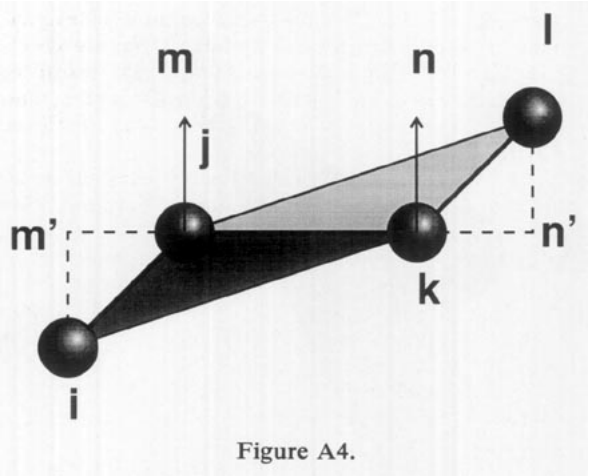


Figure A4.

(d) *Dihedral angle potential energy term*

The dihedral angle potential energy term is defined as:

$$\begin{aligned} E(\varphi) &= K_\varphi[1+\cos(n\varphi-\delta)] \\ &= K_\varphi[1+\cos(n\varphi)\cos(\delta)-\sin(n\varphi)\sin(\delta)], \end{aligned} \tag{A12}$$

where φ is defined as the angle between plane i-j-k and j-k-l (see Fig. A4). φ is calculated from:

$$\varphi = \arccos\left(\frac{\mathbf{r}_{im'}\cdot\mathbf{r}_{ln'}}{r_{im'}r_{ln'}}\right), \tag{A13}$$

where:

$$\begin{aligned} \mathbf{r}_{im'} &= \mathbf{r}_{ij}-\mathbf{r}_{m'j} = \mathbf{r}_{ij}-\frac{(\mathbf{r}_{ij}\cdot\mathbf{r}_{kj})}{r_{kj}^2}\mathbf{r}_{kj}, \\ \mathbf{r}_{ln'} &= -\mathbf{r}_{kl}-\mathbf{r}_{n'k} = -\mathbf{r}_{kl}+\frac{(\mathbf{r}_{kl}\cdot\mathbf{r}_{kj})}{r_{kj}^2}\mathbf{r}_{kj}. \end{aligned} \tag{A14}$$

In the GROMOS force field the phase shift δ equals 0 or π, hence $\sin(\delta) = 0$ and the angle multiplicity n is in the GROMOS force field limited to the values 1, 2, 3, 4 or 6. We express $\cos(n\varphi)$ in terms of powers of $\cos(\varphi)$ by using de Moivre's theorem:

$$\begin{aligned} \cos(2\varphi) &= 2\cos^2(\varphi)-1, \\ \cos(3\varphi) &= 4\cos^3(\varphi)-3\cos(\varphi), \\ \cos(4\varphi) &= 8\cos^4(\varphi)-8\cos^2(\varphi)+1, \\ \cos(6\varphi) &= 32\cos^6(\varphi)+48\cos^4(\varphi)+18\cos^2(\varphi)-1. \end{aligned} \tag{A15}$$

Now we can calculate the forces from:

$$\begin{aligned} \mathbf{f}_q &= -\frac{dE(\varphi)}{d\cos(n\varphi)}\frac{d\cos(n\varphi)}{d\cos(\varphi)}\frac{\partial\cos(\varphi)}{\partial\mathbf{r}_q} \\ &= -K_\varphi\cos(\delta)\frac{d\cos(n\varphi)}{d\cos(\varphi)}\frac{\partial\cos(\varphi)}{\partial\mathbf{r}_q}, \\ q &= i, j, k, l, \end{aligned} \tag{A16}$$

with:

$$\begin{aligned} \frac{\partial\cos(\varphi)}{\partial\mathbf{r}_i} &= \frac{1}{r_{im'}}\left(\frac{\mathbf{r}_{ln'}}{r_{ln'}}-\cos(\varphi)\frac{\mathbf{r}_{im'}}{r_{im'}}\right), \\ \frac{\partial\cos(\varphi)}{\partial\mathbf{r}_l} &= \frac{1}{r_{ln'}}\left(\frac{\mathbf{r}_{im'}}{r_{im'}}-\cos(\varphi)\frac{\mathbf{r}_{ln'}}{r_{ln'}}\right), \\ \frac{\partial\cos(\varphi)}{\partial\mathbf{r}_j} &= \left(\frac{(\mathbf{r}_{ij}\cdot\mathbf{r}_{kj})}{r_{kj}^2}-1\right)\frac{\partial\cos(\varphi)}{\partial\mathbf{r}_i} \\ &\quad -\frac{(\mathbf{r}_{kl}\cdot\mathbf{r}_{kj})}{r_{kj}^2}\frac{\partial\cos(\varphi)}{\partial\mathbf{r}_l}, \\ \frac{\partial\cos(\varphi)}{\partial\mathbf{r}_k} &= \left(\frac{(\mathbf{r}_{kl}\cdot\mathbf{r}_{kj})}{r_{kj}^2}-1\right)\frac{\partial\cos(\varphi)}{\partial\mathbf{r}_l} \\ &\quad -\frac{(\mathbf{r}_{ij}\cdot\mathbf{r}_{kj})}{r_{kj}^2}\frac{\partial\cos(\varphi)}{\partial\mathbf{r}_i}. \end{aligned} \tag{A17}$$

(e) *Non-bonded potential energy term*

The non-bonded potential energy term and the resulting forces are given by:

$$E(r_{ij}) = \frac{C_{12}(i,j)}{r_{ij}^{12}}-\frac{C_6(i,j)}{r_{ij}^6}+\frac{q_iq_j}{4\pi\varepsilon_0\varepsilon_r r_{ij}}, \tag{A18}$$

$$\begin{aligned} \mathbf{f}_i &= -\frac{\partial E(r_{ij})}{\partial\mathbf{r}_i} = \left[\frac{12C_{12}(i,j)}{r_{ij}^{14}}\right. \\ &\quad \left.-\frac{6C_6(i,j)}{r_{ij}^8}+\frac{q_iq_j}{4\pi\varepsilon_0\varepsilon_r r_{ij}^3}\right]\mathbf{r}_{ij}, \end{aligned} \tag{A19}$$

$$\begin{aligned} \mathbf{f}_j &= -\frac{\partial E(r_{ij})}{\partial\mathbf{r}_j} = \left[-\frac{12C_{12}(i,j)}{r_{ij}^{14}}\right. \\ &\quad \left.+\frac{6C_6(i,j)}{r_{ij}^8}-\frac{q_iq_j}{4\pi\varepsilon_0\varepsilon_r r_{ij}^3}\right]\mathbf{r}_{ij}. \end{aligned}$$

(f) *Distance restraining potential energy term*

The distance restraining potential energy term is incorporated in the force field to enforce the system to fulfil experimental distance restraints. A simple way to achieve this is by applying a simple harmonic restraining potential energy term with respect to distance restraint violations.

$$\begin{aligned} E(r_{ij}) &= \tfrac{1}{2}K_{dr}(r_{ij}-r_{ij}^0)^2 \qquad r_{ij} > r_{ij}^0 \\ &= 0 \qquad\qquad\qquad\qquad r_{ij} \leq r_{ij}^0. \end{aligned} \tag{A20}$$

This potential energy term only depends on the distance r_{ij}, a quantity that can be easily calculated in 4D-space. However there is a small complication, the GROMOS force field is a united atom force field which means that only polar hydrogens are explicitly represented in the force field. Since NOE distance restraints are acting on hydrogen atoms, hydrogen positions are generated during the run using the positions of neighbouring atoms (van Gunsteren *et al.*, 1985; van Gunsteren & Berendsen, 1987). This is almost always possible to do in 4D-space except for those cases where the assignment is stereo specific and the virtual hydrogen construction involves a 3D-vector product that cannot be generalized to 4D-space. In this special

case we decided to generate the hydrogen position on the basis of the 3D projection of its neighbouring atoms and give it a w-co-ordinate 0.0, i.e. we apply the distance restraint on the 3D projection of this hydrogen atom. Taking this into account, we can write for the forces:

$$\mathbf{f}_i = -\frac{\partial E(r_{ij})}{\partial \mathbf{r}_i} = -K_{dr}(r_{ij}-r_{ij}^0)\frac{\mathbf{r}_{ij}}{r_{ij}},$$
$$\mathbf{f}_j = -\frac{\partial E(r_{ij})}{\partial \mathbf{r}_j} = K_{dr}(r_{ij}-r_{ij}^0)\frac{\mathbf{r}_{ij}}{r_{ij}}. \quad \text{(A21)}$$

Using an all atom force field, this problem will disappear.

(g) *3D-projection potential energy term*

This harmonic potential energy term is added to the force field to obtain the possibility to slowly project back into three dimensions. This term only acts on the w-co-ordinate of the particles.

$$E(w) = \tfrac{1}{2}K_{3D}w^2, \quad \text{(A22)}$$

$$f_w = -\frac{dE(w)}{dw} = -K_{3D}w. \quad \text{(A23)}$$

We are grateful to Ruud Scheek for valuable comments and 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

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.

Blaney, J. M. & Crippen, G. M. (1986). DGEOM, distance geometry program. Quantum Chemistry Program Exchange (QCPE), Indiana University, Bloomington, Indiana 47405, U.S.A.

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science*, **235**, 458–460.

Clore, G. M., Gronenborn, A. M., Brünger, A. T. & Karplus, M. (1985). Solution conformation of a heptadecapeptide comprising the DNA binding helix F of the cyclic AMP receptor protein *Escherichia coli*. *J. Mol. Biol.* **186**, 435–455.

Crippen, G. M. (1982). Conformational analysis by energy embedding. *J. Comput. Chem.* **3**, 471–476.

Crippen, G. M. (1984). Conformational analysis by scaled energy embedding. *J. Comput. Chem.* **5**, 548–554.

Crippen, G. M. (1987). Why energy embedding works. *J. Phys. Chem.* **91**, 6341–6343.

Crippen, G. M. & Havel, T. F. (1988). Distance geometry and molecular conformation, Research Studies Press (Wiley), New York.

Crippen, G. M. & Havel, T. F. (1990). Global energy minimization by rotational energy embedding. *J. Chem. Inf. Comput. Sci.* **30**, 222–227.

de Vlieg, J., Boelens, R., Scheek, R. M., Kaptein, R. & Gunsteren, W. F. (1986). Restrained molecular dynamics procedure for protein tertiary structure determination from NMR data: A *lac* repressor headpiece structure based on information on J-coupling and from presence and absence of NOE's. *Israel J. Chem.* **27**, 181–188.

Havel, T. F. (1991). An evaluation of computational strategies for use in the computational determination of protein structure from distance constraints obtained by nuclear magnetic resonance. *Progr. Biophys. Mol. Biol.* **56**, 43.

Kessler, H., Köck, M., Wein, T. & Gehrke, M. (1990). Reinvestigation of the conformation of Cyclosporin A in chloroform. *Helv. Chim. Acta*, **72**, 1818–1832.

Kirkpatrick, S., Gelatt, C. D., Jr & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, **220**, 671–680.

Lautz, J., Kessler, H., Kaptein, R. & van Gunsteren, W. F. (1987). Molecular dynamics simulations of cyclosporin A: the crystal structure and dynamic modelling of the solution structure based on NMR data. *J. Comp.-Aided Mol. Design*, **1**, 219–241.

Lautz, J., Kessler, H., Blaney, J. M., Scheek, R. M. & van Gunsteren, W. F. (1989). Calculating three-dimensional molecular structure from atom-atom distance information: cyclosporin A. *Int. J. Peptide Protein Res.* **33**, 281–288.

Levitt, M. (1983). Protein folding by restrained energy minimization and molecular dynamics. *J. Mol. Biol.* **170**, 723–764.

Loosli, H. R., Kessler, H., Oschkinat, H., Weber, H. P., Petcher, T. J. & Widmer, A. (1985). The conformation of cyclosporin A in the crystal and in solution. *Helv. Chim. Acta*, **68**, 682–704.

Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988). Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. *FEBS Letters*, **239**, 129–136.

Purisma, E. O. & Scheraga, H. A. (1986). An approach to the multiple-minima problem by relaxing the dimensionality. *Proc. Nat. Acad. Sci., U.S.A.* **83**, 2782–2786.

Purisma, E. O. & Scheraga, H. A. (1987). An approach to the multiple-minima problem in protein folding by relaxing the dimensionality. *J. Mol. Biol.* **196**, 697–709.

van Gunsteren, W. F. & Berendsen, H. J. C. (1987). *Groningen Molecular Simulation* (GROMOS) Library Manual, Biomos: Groningen.

van Gunsteren, W. F., Boelens, R., Kaptein, R., Scheek, R. M. & Zuiderweg, E. R. P. (1985). An improved restrained molecular dynamics technique to obtain protein tertiary structure from nuclear magnetic resonance data. In *Molecular Dynamics and Protein Structure* (Herman, J., ed.), pp. 92–99, Polycrystal Books Service, Western Springs, IL.

van Nuland, N. A. J., Grötzinger, J., Dijkstra, K., Scheek, R. M. & Robillard, G. T. (1992). Determination of the three-dimensional solution structure of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* using NMR spectroscopy. *Eur. J. Biochem.* **210**, 881–891.

van Schaik, R. C., van Gunsteren, W. F. & Berendsen, H. J. C. (1992). Conformational search by potential energy annealing: algorithm and application to cyclosporin A. *J. Comp.-Aided Mol. Design*, **6**, 97–112.

Weber, P. L., Morrison, R. & Hare, D. (1988). Determining stereo-specific ^{1}H nuclear magnetic

resonance assignments from distance geometry calculations. *J. Mol. Biol.* **203**, 483–487.

Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., van Gunsteren, W. F. & Kaptein, R. (1985). Determination of protein structures from nuclear magnetic resonance data using a restrained molecular dynamics approach: the *lac* repressor DNA binding domain. *Biochimie*, **67**, 707–715.

Edited by R. Huber