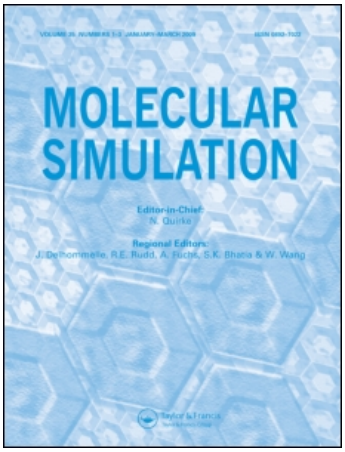

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

On the Use of United Atoms in Statistical Mechanical Simulations of Biomolecules

Olle Teleman[a]; Peter Ahlström[a]; Bo Jönsson[a]
[a] Physical Chemistry 2, Chemical Centre, Lund, Sweden

To cite this Article Teleman, Olle , Ahlström, Peter and Jönsson, Bo(1991) 'On the Use of United Atoms in Statistical Mechanical Simulations of Biomolecules', Molecular Simulation, 7: 3, 181 — 194
To link to this Article: DOI: 10.1080/08927029108022152
URL: http://dx.doi.org/10.1080/08927029108022152

Molecular Simulation, 1991, Vol. 7, pp. 181–194
Reprints available directly from the publisher
Photocopying permitted by license only

ON THE USE OF UNITED ATOMS IN STATISTICAL MECHANICAL SIMULATIONS OF BIOMOLECULES

OLLE TELEMAN, PETER AHLSTRÖM and BO JÖNSSON

Physical Chemistry 2, Chemical Centre, POB 124, S-22100 Lund, Sweden

(*Received May 1990, accepted October 1990*)

Three different Molecular Dynamics simulations of calbindin D_{9k}, each lasting at least 160 ps, are compared in order to assess the effects of the "united atom" approximation in statistic mechanical simulations of macromolecules. One simulation is performed with the united atom concept applied to aliphatic CH_n-groups, while a second simulation treats all atoms explicitly using carbon and hydrogen parameters derived from crystal data of alkanes. In a third force field the hydrogen parameters have been changed so as to give an artificially strong attraction. The applicability of the force fields is judged from a comparison between simulated and experimental fluorescence depolarization and nmr relaxation data. The second force field is shown to be in best agreement with experiment and does also show the smallest rms deviation from the protein crystal conformation. The united atom force field is only slightly inferior and the artificial force field is moderately inferior as regards the comparison with experimental data.

KEY WORDS: United atoms, molecular dynamics, calbindin.

INTRODUCTION

The application of statistical mechanical simulation techniques to biomolecular problems is a novel area of research. Monte Carlo and Molecular Dynamics simulations of simple liquids or of small molecule solvation are established techniques whose applicability and technical limitations are reasonably well understood. The appearance of supercomputers and the development of efficiently vectorised simulation codes have opened up new areas for numerical simulations and it is today possible to study for example small proteins by Molecular Dynamics (MD) simulations.

Simulations are of course limited in size and time — what is happening on a nanosecond time scale or slower is beyond reach of an MD simulation today. These limitations are set by the available computer resources and are changing. Other, more serious limitations are the lack of accurate intermolecular potentials for such complicated systems and the problems associated with the treatment of intramolecular degrees of freedom. For example, nobody has as yet been able to reproduce a secondary protein structure from a given force field and known primary sequence — and to obtain tertiary and quaternary structures seems even more futuristic. Hence, one should not use MD simulations to generate protein conformations from first principle, but rather as a tool for investigating local structures and fluctuations. In this case, the underlying assumption is that short-range interactions are well-described in simulations and that these together with the known covalent topology determine the local features. In this spirit MD simulations can be combined with distance constraints derived from nOe measurements in nmr spectroscopy. The constraints then guarantee the correct overall fold, while the simulation provides local structure and

dynamics. Stringent proof of this does not exist, but the simulation technique [1,2] has been compared [3] to distance geometry results [4,5] and found to converge with these.

Considering the large uncertainty in available potential models, the sensitivity of simulation results to small changes in the model is critical. Here, we try to assess this sensitivity from protein simulations with different force fields. We have chosen to study the effect of the united atom approximation, which is frequently applied to protein simulations and amounts to treating in particular non-polar hydrogen atoms and the heavy atom to which they are bound as one unit. For a typical biomolecule this reduces the total number of atoms by about 50 per cent, which entails considerable savings when simulating molecular systems *in vacuo*. In solution the gain is negligible as the main computational effort is spent on solvent-solvent interactions.

It has been noted in simulations of n-alkanes that a united atom approximation does not work as well for solid alkanes as it seems to do for the liquid state. An explicit representation of the hydrogen atoms as separate interaction sites seems to be a remedy [6]. Explicit hydrogen representation also appears necessary in order to obtain proper dynamic behaviour of the hydrocarbon tails in a micelle [7]. In the latter system the lack of explicit hydrogen atoms leads to a too fast rotational motion. The conclusion is based on rather indirect evidence, although it is not unreasonable considering the anisotropy of the repulsive part of a CH_n-group.

A different kind of weakness in a united atom approach appears in the interpretation of 2D nmr relaxation data. Here the coupling between hydrogen atoms is used to determine their spatial separation. It is evident that a united atom simulation offers a less stringent comparison with experimental data, since it requires the insertion of the implicit hydrogen atoms on idealized positions. However, the uncertainty added by use of idealized geometry or pseudoatoms is generally small compared to the uncertainty of the nOe measurement itself.

In this communication we report a comparison of three simulations of the intestinal calcium binding protein D_{9k}, generally called calbindin D_{9k}, in aqueous solution. The first simulation employs the united atom concept (referred to as the *U* simulation), while the two others treat all hydrogen atoms explicity. The force field was extended to explicit non-polar hydrogens atoms in two ways. Firstly, Lennard-Jones parameters were fitted directly to the force field derived by Williams from the lattice structure of hydrocarbon crystals [8] — referred to as the *A* force field. This is not a fully "consistent" force field, as it was not refitted to the data underlying the original *U* force field. That would go beyond the scope of this work. The second all atom force field (*B*) contains an artificially strong attractive interaction in all hydrogen terms and may tell us how structural and dynamical averages in a reasonably long protein simulation are affected by obvious faults in the potential. In the analysis we largely follow the treatment in an earlier paper [9] with emphasis on comparison with experimentally available properties so that simulation accuracy can be assessed. Calbindin was chosen for this study because of long-standing collaboration with nmr groups providing experimental data. In contrast to, say BPTI, the solution structure is not known, but work on it is in progress (W. Chazin, N. Skelton, personal communication).

INTERMOLECULAR POTENTIAL

The detailed features of the united atom potential have been discussed elsewhere [10]. The interaction between two atoms breaks down into covalent and noncovalent

interactions, the latter between atoms in the same molecule more than two bonds apart and between atoms in different molecules. The noncovalent potential, between two atoms i and j, consists of a Lennard-Jones (LJ) plus a pure electrostatic potential,

$$U_{ij} = 4\varepsilon_{ij}\left\{\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right\} + \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r r_{ij}}, \quad (1)$$

where ε_{ij} and σ_{ij} are the usual LJ parameters and q_i is the charge of atom i. The relative dielectric permittivity, ε_r, was set equal to unity. Parameters were taken from the literature [11–13] and the LJ parameters were calculated with the Slater-Kirkwood formula [13] under the assumption that the LJ potential minimum corresponds to the sum of the van der Waals radii of the two atoms. In particular we used the so called simple point charge (SPC) model for water [14] while including also intramolecular vibrations [15].

For the covalent part all internal bond and bond angle vibrations were treated explicitly according to

$$U_{\text{intra}} = \sum_{\text{bonds}} \frac{k_b}{2}(x_i - x_{i,\text{eq}})^2 + \sum_{\text{angles}} \frac{k_a}{2}(\theta_i - \theta_{i,\text{eq}})^2 , \quad (2)$$

in which x_i and $x_{i,\text{eq}}$ are the actual and equilibrium bond lengths and similarly for the bond angle with the force constants taken from the literature [11–12,14,16–17].

Internal rotations were handled with periodic dihedral potentials

$$U_{\text{int.rot.}} = \sum_k \sum_n C_{k,n}\{1\text{-}\cos(n\psi_k)\} \quad (3)$$

where ψ_k is a dihedral angle, $C_{k,n}$ an interaction parameter [12], k runs over all dihedral angles at each bond and n goes from 1 to 3. In practice only one of the three $C_{k,n}$ coefficients is non-zero.

There is no unique way to reduce the CH, CH_2 and CH_3 groups into single united atom centers. One approach is described by Dunfield *et al.* [18], another is to use the so called Slater-Kirkwood formula. The latter automatically leads to an increased dispersion interaction for the pseudoatoms due to the larger number of electrons, compensating for the hydrogen atom contributions. However, the modification of the repulsive part of the potential still remains arbitrary to some extent. In principle, the interatomic potential for the united atoms should also asymptotically contain an attractive r^{-4} contribution from thermally averaged charge-dipole interactions. However, it seems to be common practice to neglect this contribution and so has been done also in the present U potential. Some LJ parameters for the U potential are given in Table 1.

The net charges on -CH, $-CH_2$ and $-CH_3$ groups were kept at the same numbers as for the united atoms. The hydrogen atoms were given a charge of $+0.061$ e, which corresponds to a dipole moment for a C–H bond of 0.3 D, close to what can be estimated from theoretical calculations.

Our first model of non-polar hydrogens (the A force field) is based on a study of solid n-alkanes by Williams [8], who used the functional form

$$A\,r^6 + B\exp(-Cr),$$

where r is the distance and A, B and C are parameters. For consistency we have refitted Williams parameters (parameter set IV) to the Lennard-Jones form (equation

Table 1a Lennard-Jones distance parameter σ in Å for some sites. The CH, CH_2 and CH_3 sites are united atoms, the H site is a non-polar hydrogen atom. The top, middle and bottom numbers in each triplet belong to the *U*, *A* and *B* force fields, respectively. Note the extreme difference between *A* and *B* force fields in the H–H interaction.

	CH	CH_2	CH_3	C (sp^3)	CH (sp^2)	C (sp^2)	H
	3.296	3.341	3.385	3.252	3.341	3.252	-
CH	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
		3.385	3.430	3.296	3.385	3.296	-
CH_2		-	-	-	-	-	-
		-	-	-	-	-	-
			3.475	3.341	3.430	3.341	-
CH_3			-	-	-	-	-
			-	-	-	-	-
				3.207	3.296	3.207	-
C (sp^3)				3.457	-	3.421	2.900
				3.207	-	3.207	2.495
					3.385	3.207	-
CH (sp^2)					-	-	-
					-	-	-
						3.207	-
C (sp^2)						3.385	2.900
						3.207	2.495
							-
							2.976
H							1.782

Table 1b Lennard-Jones well depth parameter ε in kJ/mol for some sites. The CH, CH_2 and CH_3 sites are united atoms, the H site is a non-polar hydrogen atom. The top, middle and bottom numbers in each triplet belong to the *U*, *A* and *B* force fields, respectively. Note the extreme difference between *A* and *B* force fields in the H–H interaction.

	CH	CH_2	CH_3	C (sp^3)	CH (sp^2)	C (sp^2)	H
	0.568	0.667	0.741	0.681	0.717	0.681	-
CH	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
		0.784	0.873	0.801	0.846	0.801	-
CH_2		-	-	-	-	-	-
		-	-	-	-	-	-
			0.974	0.891	0.943	0.891	-
CH_3			-	-	-	-	-
			-	-	-	-	-
				0.825	0.869	0.825	-
C (sp^3)				0.398	-	0.574	0.205
				0.611	-	0.770	0.866
					0.918	0.869	-
CH (sp^2)					-	-	-
					-	-	-
						0.825	-
C (sp^2)						0.838	0.205
						0.984	1.084
							-
H							0.040
							1.782

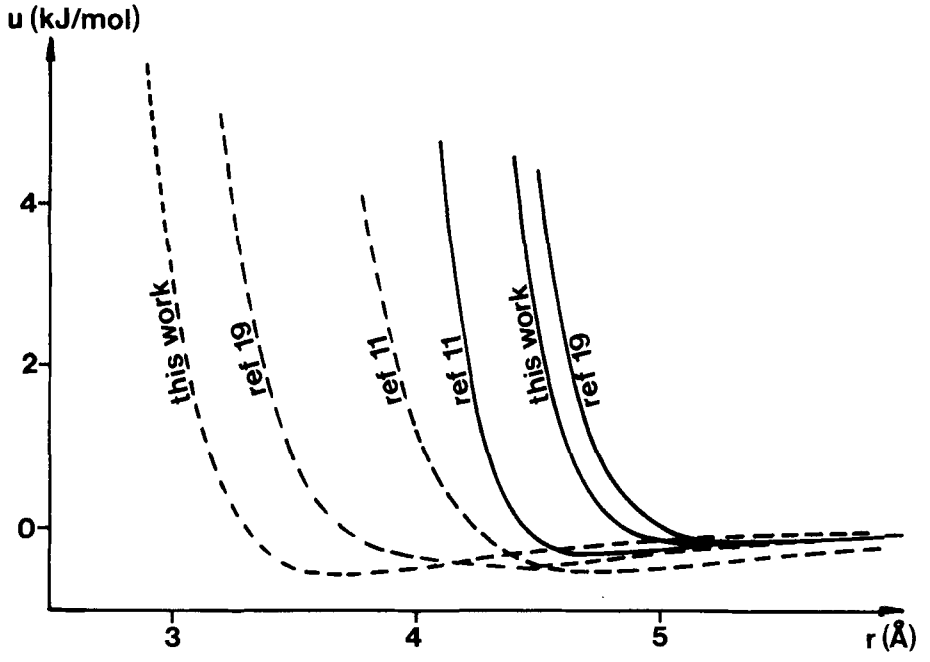

Figure 1 The Lennard-Jones interaction in kJ/mol as a function of C–C distance for a pair of co-linear aliphatic C–H groups. The solid curves are obtained with all atom force fields and the dashed curves with united atom force fields.

(1)). In the potential well the Lennard-Jones fit is very accurate, while slightly harder at small separations (0.04, 0.10 and 0.18 Å harder for C–C, C–H and H–H pairs respectively at an energy of kT at 300 K). Williams gives C–H data as well as C–C and H–H, which enables us to generate C–H parameters from first principles rather than by resorting to standard combination rules. The effect of this is probably largest for carbon–hydrogen pairs because they are numerous. The parameters for some sites are given in Table 1, and we note that $\sigma_{CH} \neq 0.5\ (\sigma_{CC} + \sigma_{HH})$ for the *A* force field.

In the *B* force field, we have introduced an artificially strong hydrogen interaction. For a pair of methyl groups, the interaction has a well depth of about 10 kJ/mol. The large interaction comes from an unphysically small r^{-12} coefficient, while the attractive r^{-6} term is close to theoretical estimates. Several force fields in the literature suffer from similarly small repulsive terms, but are helped by a combination with a too small dispersion term.

At this stage it seems appropriate to investigate the differences in interatomic potential between a united atom and an all atom model. We have also included some other common force fields in the comparison – the force field of Weiner *et al.* [19] used in AMBER and a force field due to Hermans *et al.* [11]. Figure 1 shows the interaction between two co-linear aliphatic C–H groups calculated with different force fields. The figure shows a huge difference between an all atom and a united atom calculation. The difference corresponds to more than a doubling of the C–H group volume. In a dense system consisting of pure hydrocarbons or mainly hydrocarbons it would lead to an enormous pressure. In the present system it is less dramatic, since the aliphatic

Table 2 Simulation parameters and averages.

	U	*A*	*B simulation*
Number of atoms	7468	7946	7946
Simulation box	43.4 × 40.3 × 46.5 Å in all simulations		
Length of simulation (ps)	39 + 124	42 + 122	40 + 127
Integration time steps (fs)	0.2 and 1.2	0.2 and 1.2	0.2 and 1.2
Temperature scaling interval (ps)	0.096	0.096	0.096
Total cpu time per ps (s)	6100[a]	6800[b]	7100[a]
Temperature (K)	302	304	304
Pressure (MPa)	− 36	59 ± 13	− 53
Radius of gyration (Å)	10.7 ± 0.1	12.0 ± 0.1	10.3 ± 0.1
Overall reorientation, τ_2 (ns)	0.9	1.0	ca. 0.4[c]
Diffusion coefficient (m^2/s)	$1.5 \cdot 10^{-10}$	$1.3 \cdot 10^{-10}$	$1.3 \cdot 10^{-10}$

[a] On a IBM 3090-170 S processor. [b] On one cpu of an IBM 3090-600 E. [c] This number is uncertain as fluctuations obscure the most long-lived component in the time correlation functions.

hydrogens only constitute a minor part, although the pressure is seen to increase between the *U* and the *A* simulation (Table 2). The depth of the minimum is similar in the two models. Hermans *et al.* [11] provide two choices for united atoms, one of which is in reasonable agreement with Williams potential. This choice is the one used in the GROMOS package, while the second is identical to our *U* potential.

For aromatic C-H groups a very accurate potential exists, based on *ab initio* quantum chemical calculations, which we have used as a reference in the comparison in Figure 2. The large discrepancy between all atom and united atom potentials persists also for aromatic groups. None of them agree with the "exact" calculation of Karlström *et al.* [20]. It should be stressed that the *ab initio* potential gives excellent results for liquid benzene [21] including an accurately reproduced second virial coefficient [20]. We also note that Williams potential obtained from solid n-alkane data differs significantly from the *ab initio* benzene potential. The discrepancies between semi-empirical and *ab initio* force fields are annoying but partly reflect the different conditions and species treated. Semi-empirical force fields include many-body terms in an average way, whereas most *ab initio* models only include pairwise polarizability etc. This may change, however, as a method now exists to obtain *ab initio* force fields with explicit polarizability [22].

The comparisons in Figure 1 and 2 do not include the Coulombic $1/r$-term. The inclusion of it will mainly affect the depth of the minimum in the all atom potentials, but has no significant effect on the repulsive part of a potential.

The large discrepancies seen in the above potentials, both between all atom and united atom models *and* between these and the most accurate potentials obtained from other sources, raises the question of their validity in protein simulations. Alternatively, how large faults can one afford to have in a potential and still expect realistic averages from the simulation? The motivation behind the *B* simulation was to gain some insight into the problem.

All simulations were started from the crystal structure of the protein [23]. In the two all atom simulations, hydrogen atoms were inserted at their idealized sp^2 or sp^3 positions. The protein in its crystal conformation, together with 36 crystal waters, was centered in a simulation box with the dimensions 43.4 × 40.3 × 46.5 Å. The box was then filled with 2212 more water molecules placed in a primitive cubic lattice.

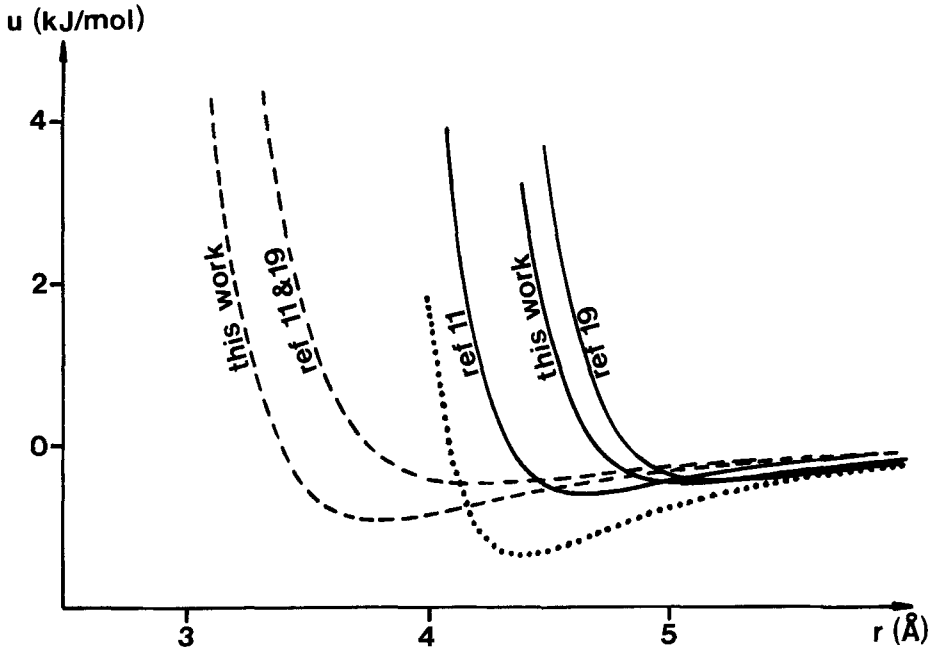

Figure 2 The Lennard-Jones interaction in kJ/mol as a function of C–C distance for a pair of co-linear aromatic C–H groups. The solid curves are obtained with all atom force fields and the dashed curves with united atom force fields. The dotted curve is obtained from quantum chemical calculations of ref. [20]. The united atom potentials of refs. [11] and [19] are virtually identical and drawn as one curve.

The MD simulations were performed with the program MUMOD [24], utilizing a double time step Gear algorithm. Rapidly varying degrees of freedom (bond lengths and bond angles) were integrated with a time step of $\Delta t = 0.2fs$ ($fs = 10^{-15}$ s) whereas for all other degrees of freedom a time step ΔT of $1.2fs$ was used (dihedral angles and non-covalent interactions). Periodic boundary conditions in all directions were used together with a spherical truncation at 10 Å for nonbonded interactions, and a neighbour list technique. In all simulations the initial x-ray conformation was allowed to equilibrate for 40 ps after which the simulation continued another 125 ps. In the following we have analysed the 125 ps part of the trajectories, and the equilibration trajectories were in no case used for analysis. Simulation parameter and thermodynamic averages are give in Table 2.

RESULTS AND DISCUSSION

The comparison of simulated and experimental quantities is obscured by the fact that we never obtain proper ensemble averages in protein simulations — that is, only a limited part of phase space is sampled. The three simulations were started from the same initial configuration (the x-ray structure), but they will, as an effect of differences in force field, loose coherence and *maybe* cover different parts of phase space. Clearly, a sub-nanosecond simulation can not provide a true ensemble average for such a complex system and we can not ascertain if a difference in a calculated quantity is due

to the force field or the lost coherence. However, it is a general belief that the x-ray structure is close to the solution structure and that the latter is a well-defined local free energy minimum. Determinations of solution conformations of tendamistat [25], the barley serine protease inhibitor 2 [26] and the basic pancreatic trypsin inhibitor [27] support this idea. Calbindin is a remarkably stable protein [28] and preliminary nmr data [31] as well as calcium binding studies [30] indicate that calbindin largely retains its crystal conformation in aqueous solution. It is an implicit assumption in the present work that the relevant part of phase space is visited during the course of the simulation and hence any differences in calculated averages are due to the force field.

The amount of experimental data to use in a validation of simulated averages is usually rather scarce. Quite often the rms deviation from the known x-ray structure is the only available comparison, which is of course unsatisfactory since the solution structure need not be the same as the crystal structure. Normally the rms deviation increases gradually during simulation — at least for the first 100–200 ps [10,29]. Fortunately, for calbindin fluorescence depolarisation data for the single tyrosine residue exist, as well as ^{43}Ca nmr relaxation measurements, which may be used to judge force fields.

Global Properties

In order to compare the simulated protein structures with the initial x-ray conformation we have superimposed simulation conformations on the x-ray structure and minimized the rms deviation, defined as

$$R_{A,C} = \left[\sum_{i=1}^{n} m_i \mid \mathbf{r}_{i,A} - \mathbf{r}_{i,C} \mid^2 / \sum_{i=1}^{N} m_i \right]^{1/2} \quad (4)$$

Here, m_i is the mass of atom i and $\mathbf{r}_{i,A}$ and $\mathbf{r}_{i,C}$ its coordinate in the simulated conformation and the x-ray structure, respectively. The rms deviations were monitored for the whole trajectory (Figure 3a). Although the simulation was preceded by 40 ps of equilibration the rms deviations have not stabilized. This is a typical behaviour seen in many protein simulations [10,29], and may be a consequence of removing the constraints imposed by the crystal or of force field inaccuracy. Also the *A* simulation show considerable fluctuations, even if the rms deviation decreases during the last 25 ps. We point out that the rms deviation against the crystal structure is not a criterion to judge force fields by, since the solution and crystal structures may be different. In fact, preliminary results on the solution structure indicate an rms deviation of about 2 Å (N. Skelton, personal communication).

We have also calculated the rms deviation as a function of time for pairs of trajectories in order to obtain a measure of their difference (cf. Figure 3b). Here, $R_{A,U}$ and $R_{A,B}$ are larger than $R_{U,B}$, 3.1, 2.8 and 2.1 Å respectively, and they are similar in magnitude to $R_{U,C}$ $R_{A,C}$ and $R_{B,C}$. Thus the changes the protein molecule undergoes in a 160 ps simulation are similar in extent to the discrepancy between simulations, which is caused by lost fluctuation coherence and/or by trajectories exploring different parts of phase space.

The radius of gyration is 11.3 Å in the crystal structure. This decreases to 10.7 Å and 10.3 Å in the *U* and *B* simulations, but increases to 12.0 Å in the *A* simulation (Table 2), due to the larger CH, CH_2 and CH_3 groups. These changes correspond to volume changes of approximately 10–20 per cent, which is much more than is to be expected

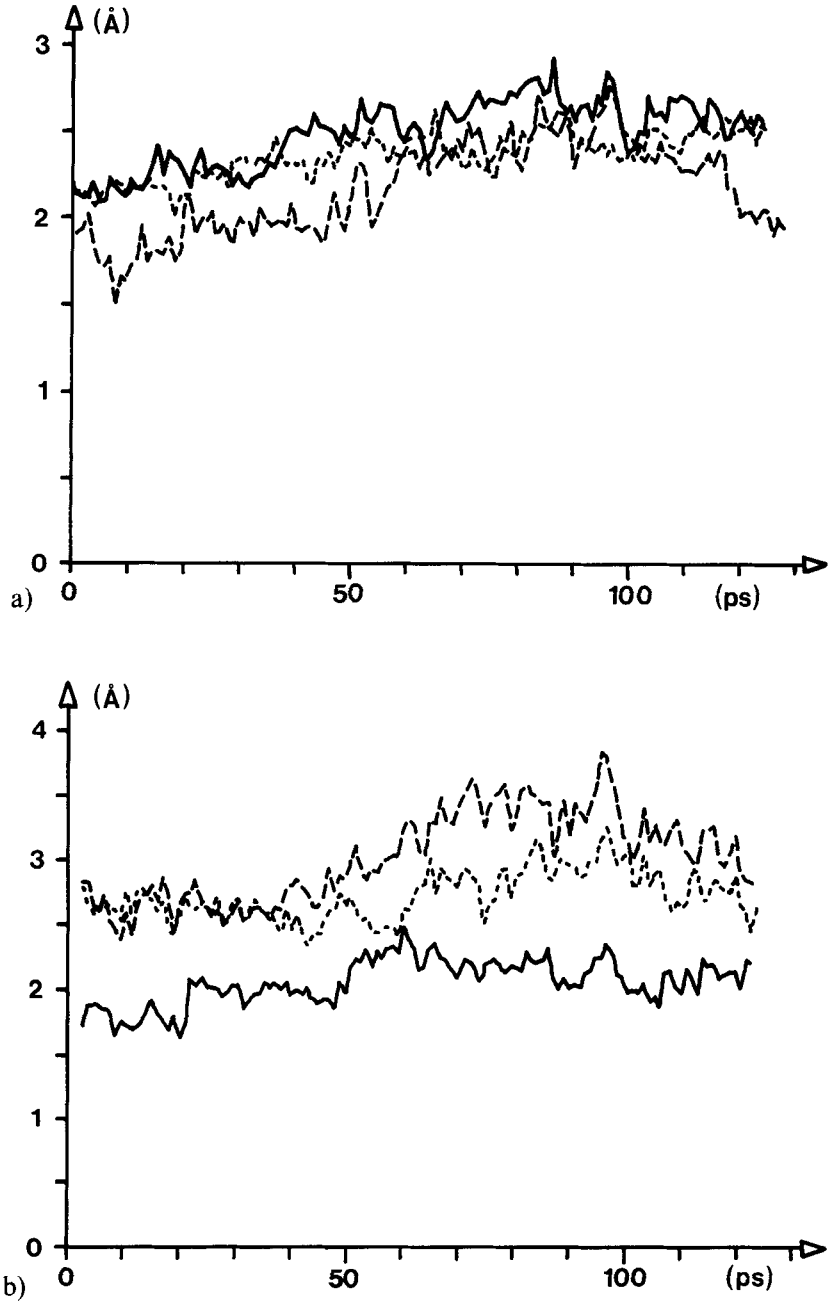

Figure 3 Root mean square deviations at optimal rigid body superposition for the heavy backbone atoms C_α, C, O and N. Figure 3a shows the deviation from the crystal conformation as a function of time in the *U* (full line), *A* (dashed line) and *B* (dotted line) simulations. Figure 3b compares pairs of simulations. Full line shows the rms deviation between *B* and *U* simulations, dashed line between *A* and *U* simulations and dotted line between *A* and *B* simulations.

for a protein going from the crystal to the solution state. It is unlikely that this increase is due to single residues protruding out into the solution, in particular as the protein largely retains its crystal structure in aqueous solution [30, 31]. In consequence the large deviation in radius of gyration in all three trajectories must be regarded as simulation artefacts.

The protein volume differs for the three force fields, which means that the volume accessible to water is different, and thus the pressure. By comparing two simulations we can calculate an approximate value for the isothermal compressibility of water. This is not relevant to the comparison of force fields, but we will pursue it nevertheless, since only seldomly several simulations of similar systems are available. We compare two approximate state points for water (p, V, T) represented by two simulations. The temperature is the same but p and $V = V_{box} - V_{protein}$ differ, thus

$$\kappa = -(\Delta V/V)(\Delta p)^{-1}$$

which was found to be $8 \cdot 10^{-10}$ Pa^{-1} = $8 \cdot 10^{-5}$ bar^{-1}, while the real compressibility of water is $4.5 \cdot 10^{-5}$ bar^{-1}. Now, the simulation value is only a rough estimate due to the indirect method of calculation.

The diffusion coefficient is roughly the same in all simulations. The global reorientation was evaluated by means of the total angular momentum [32], and is somewhat longer for the A than the U simulation, which is also consistent with the larger radius of gyration. Note that the overall τ_2 for the U simulation (0.9 ns, Table 2) is longer than the value given in reference [9]. That value was based on α-helix reorientation, some fluctuations in which can be averaged out in a more stringent treatment.

Local Properties

The presence of only one tyrosine sidechain, Tyr 13, in calbindin D_{9k} simplifies the interpretation of time-resolved fluorescence experiments. Rigler *et al.* [33] determined τ_2 = 4.6 ns for the global reorientation but also identified a local component with τ_2 = 0.36 ns. The directions of absorption and emission dipoles of the tyrosine ring are unknown, but the dynamic range of ring reorientation can be estimated from τ_2 for a vector from C_γ to C_ξ ($\tau_{2,\parallel}$) and from one across the ring, i.e. from $C_{\delta 1}$ to $C_{\delta 2}$ or from $C_{\varepsilon 1}$ to C_ε ($\tau_{2,\perp}$). The τ_2 values are given in Table 3. All three simulations provide time correlation functions with rather large fluctuations. The U simulation contains an identifiable motion with quite short τ_2 values, while slower components are present but cannot be quantitatively assessed. The A and B simulations both have a more

Table 3 Characteristic time in ns for the time correlation function based on the second order Legendre polynomial for the cosine of vectors from C_γ to C_ξ ($\tau_{2,\parallel}$) as well as from $C_{\delta 1}$ to $C_{\delta 2}$ and from $C_{\varepsilon 1}$ to $C_{\varepsilon 2}$ ($\tau_{2,\perp}$) in Tyr 13, as well as a correlation time determined by time resolved fluorescence anisotropy. Characteristic times were derived from correlation functions by least-squares linear fit to the logarithm between 5 and 25 ps. The $\tau_{2,\perp}$ value for the B simulation is given as a range because the time correlation function cannot be properly resolved into components between 5 and 25 ps. If more of the correlation function is used, characteristic times of 0.1 and 0.5 ns are obtained, so that this gives a range for the wanted property.

	$\tau_{2,\parallel}$ (ns)	$\tau_{2,\perp}$ (ns)
U	0.06	0.04
A	0.20	0.25
B	0.31	0.1–0.5
Experiment	0.36	

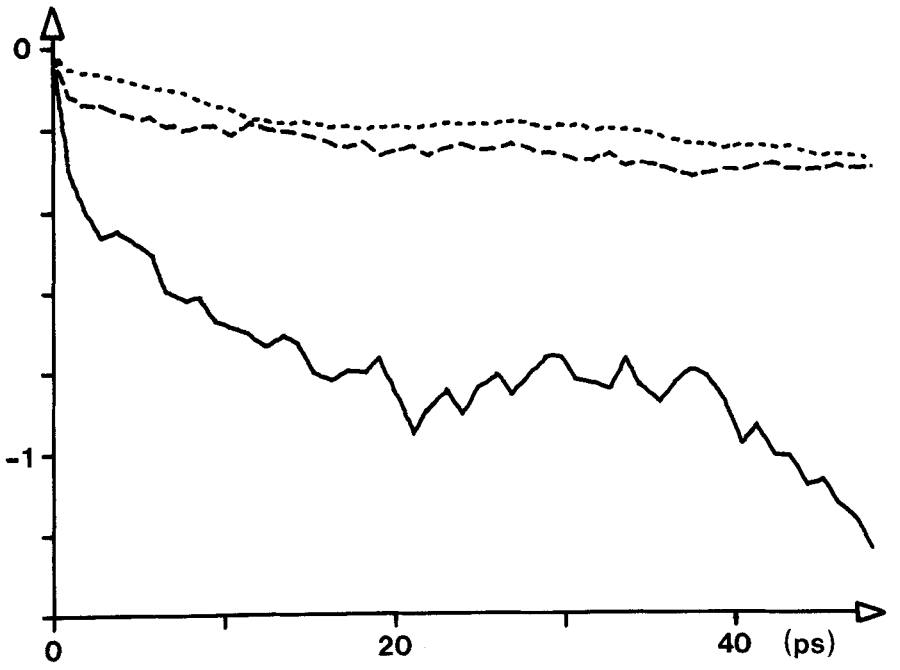

Figure 4 Logarithm of the time correlation function (C_2) based on the second order Legendre polynomial of the reorientation cosine for vectors across the Tyr 13 ring (from $C_{\delta 1}$ to $C_{\delta 2}$ and from $C_{\varepsilon 1}$ to $C_{\varepsilon 2}$). Full, dashed and dotted lines show the *U*, *A* and *B* simulations, respectively.

pronounced intermediate time range motion, with correlation times quite close to the experimental value. Statistics underlying $\tau_{2,\perp}$ are better in the *A* simulation. In no case is there a flip. It is noticeable that both the *A* and *B* simulations give similar correlation times, despite the large differences in the potential. Figure 4 gives an indication of the numerical difficulties met with when trying to extract a well-defined correlation time.

To further characterize local mobility, Table 4 lists the mean square fluctuations around mean positions of the atoms in certain residues in the hydrophobic protein interior. Translation and rotation of the molecule as a whole has been removed by optimal rigid body superposition. Except for a constant factor of $8\pi^2/3$ the quantity obtained is the direct counterpart of the temperature factors used in crystallography. The fluctuations are similar in the *U* and *A* trajectories, but far smaller in the *B* one, again an effect of the artificially large H–H attraction. The crystal temperature factors have not been published in detail [23].

The ^{43}Ca nmr spectrum of ^{43}Ca-loaded calbindin contains one signal at 7 ppm with a line width of 350 Hz, asigned to site II by Vogel *et al.* [34], and one at 9 ppm with a width of 1000 Hz (site I). For a quadrupolar nucleus such as ^{43}Ca the line width is determined by the interaction of the nuclear quadrupole moment with the electric field gradient. In an earlier paper [9] we reported line widths based on the *U* simulation and Table 5 compares the *A*, *B* and *U* simulations. Note here that the $\Delta\nu_{\mathrm{I}}$ and $\Delta\nu_{\mathrm{II}}$ for the *U* simulation differ from those given earlier, again because of the better evaluation of τ_2 for the overall rotation. From the ^{43}Ca line width ratio the *A* force field does very

Table 4 Internal flexibility as monitored by fluctuations around mean positions for 8 residues in the hydrophobic interior of calbindin. Mean positions of atoms were calculated after removal of translation and rotation of the entire molecule. The mean square deviations from mean positions have been averaged over the individual residues and correspond to the temperature factor used in X-ray crystallography if multiplied by $8\pi^2/3$. Numbers are in $Å^2$.

	U	*A*	*B*
Phe 10	1.14	1.56	0.27
Leu 23	0.86	0.65	0.77
Leu 28	0.93	0.31	0.08
Leu 31	0.88	0.86	0.25
Leu 32	0.76	0.55	0.10
Val 61	0.39	0.85	0.18
Phe 63	0.65	1.18	0.29
Phe 66	0.61	0.42	0.14
average of these	0.78	0.80	0.26

much better than the *U* force field. If we take the too fast dynamics of the applied water model into account, the *A* simulation provides also absolute line widths very close to the experimental results. In the *B* simulation the line width ratio is reversed, so that this trajectory cannot be used for assignment purposes. The absolute line widths are very small due to too fast overall motion and a too small average field gradient. For this property the *B* simulation is far inferior to the *U* and *A* ones. It should be pointed out that uncertainties in the Sternheimer shielding factor, in overall reorientation time and in nuclear quadrupole moment affect the accuracy of absolute line widths but not of the line width ratio, so that the latter is the better property to judge the force field by.

CONCLUSION

The all atom force field *B*, which exaggerates the dispersion attraction of the hydrogen atoms, provides slightly better agreement with experimental data than the united atom force filed *U* as regards Tyr 13 local motion. The *B* force field fails to reproduce the quadrupolar relaxation of the calcium ions. The rms deviation from the crystal form is less for calbindin in the *A* simulation than in the *U* or *B* simulations. The radius of gyration is too large in the *A* simualtion, which may be due to the fit to Williams' data. An adjustment of parameters to improve the fit for the repulsive part is a possible remedy. Attempts at avoiding the protein compaction in earlier force fields (e.g. the *U* simulation) have involved increasing united atom radii. This sometimes causes a repulsion between atoms separated by three covalent bonds and thus distorts the dihedral angle potential. In many united atom force fields this problem is circumvented by reducing the Lennard-Jones interactions for these so-

Table 5 Line widths in the ^{43}Ca nmr spectrum of calbindin D_{9k}.

	U	*A*	*B*	*Experiment*
$\Delta\nu_I$ (Hz)	490	320	40	1000
$\Delta\nu_{II}$ (Hz)	80	110	70	350
$\Delta\nu_I/\Delta\nu_{II}$	6	3	0.6	3

called 1–4 pairs, but the problem is entirely avoided by all atom force fields. We also note that it is only by careful comparison with experimental data, in this case nmr relaxation data, that we can discard the artifical force field *B*.

The better realism of the *A* force field makes it the model to prefer, especially in view of the moderate increase in computation.

Acknowledgement

We are grateful to Prof. W.F. van Gunsteren for valuable comments on this manuscipt.

References

[1] R. Kaptein, E.R.P. Zuiderweg, R.M. Scheek, R. Boelens and W.F.v. Gunsteren, "A Protein Structure from Nucelar Magnetic Resonance Data: lac Repressor Headpiece", *J. Mol. Biol.*, **182**, 179–182 (1985).

[2] T.A. Holak, S.K. Kearsley, Y. Kim and J.H. Prestegard, "Three-Dimensional Structure of Acyl Carrier Protein Determined by Nuclear Magnetic Resonance Pseudoenergy and Distance Geometry Calculations", *Biochemistry*, **27**, 6135–6142 (1988).

[3] G.M. Clore, M. Nilges, A.T. Brünger, M. Karplus and A.M. Gronenborn, "A Comparison of the Restrained Molecular Dynamics and Distance Geometry Methods for Determining Three-Dimensional Structure of Proteins of the Basis of Interproton Distances", *FEBS Lett.*, **213**, 269–277 (1987).

[4] W. Braun and N. Go, "Calculation of Protein Conformations by Proton-Proton Distance Constraints", *J. Mol. Biol.*, **186**, 611–626 (1985).

[5] G.M. Crippen and T.F. Havel, *"Distance Geometry and Molecular Conformation"* Research Studies Press, Taunton, Somerset, 1988.

[6] J.-P. Ryckaert and M.L. Klein, "Translational and rotational disorder in solid n-alkanes: Constant temperature-constant pressure molecular dynamics calculations using infinitely long flexible chains", *J. Chem. Phys.*, **85**, 1613–1620 (1986).

[7] B. Jönsson, O. Edholm and O. Teleman, "Molecular dynamics simulation of a sodium octanoate micelle in aqueous solution", *J. Chem. Phys.*, **85**, 2259–2271 (1986).

[8] D.E. Williams, "Nonbonded potential parameters derived from crystalline hydrocarbons", *J. Chem. Phys.*, **47**, 4680–4684 (1967).

[9] P. Ahlström, O. Teleman, J. Kördel, S. Forsén and B. Jönsson, "A molecular dynamics simulation of bovine calbindin D_{9k}. Molecular structure and dynamics", *Biochemistry*, **28**, 3205–3211 (1989).

[10] P. Ahlström, O. Teleman, B. Jönsson and S. Forsén, "Moleclar dynamics simulation of parvalbumin in aqueous solution", *J. Am. Chem. Soc.*, **109**, 1541–1551 (1987).

[11] J. Hermans, H.J.C. Berendsen, W.F. van Gunsteren and J.P.M. Postma., "A consistent empirical potential for water-protein interactions", *Biopolymers*, **23**, 1513–1518 (1984).

[12] W.F. van Gunsteren and M. Karplus, "Effects of constraints on the dynamics of macromolecules", *Macromolecules*, **15**, 1528–1544 (1982).

[13] M. Margenau and N.R. Kestner, *Theory of Intermolecular Forces*, Pergamon, New York, 1969.

[14] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, and J. Hermans, "Interaction Models for Water in Relation to Protein Hydration" in *Intermolecular forces*, B. Pullman, ed., D. Reidel, Dordrecht, 1981, p. 331–342.

[15] O. Teleman, B. Jönsson and S. Engström, "Molecular dynamics simulation of a water model with intramolecular degrees of freedom", *Mol. Phys.*, **60**, 193–203 (1987).

[16] D. Dolphin and A.E. Wick, *Tabulation of Infrared Spectral Data*, Wiley-Interscience, New York, 1977.

[17] G. Herzberg, *Molecular Spectra and Molecular Structure: Infrared and Raman Spectra of Polyatomic Molecules*, van Nostrand, Princeton, 1945.

[18] L.G. Dunfield, A.W. Burgess and H.A. Scheraga, "Energy Parameters in Polypeptides. 8. Empirical Potential Energy Algorithm for the Conformational Analysis of Large Molecules", *J. Phys. Chem.*, **82**, 2609–2616 (1978).

[19] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta Jr. and P. Weiner, "A new force field for molecular mechanical simulation of nucleic acids and proteins", *J. Am. Chem. Soc.*, **106**, 765–784 (1984). S.J. Weiner, P.A. Kollman, D.T. Nguyen and D.A. Case, "An all atom force field for simulations of proteins and nucleic acids", *J. Comp. Chem.*, **7**, 230–252 (1986).

[20] G. Karlström, P. Linse, A. Wallquist and B. Jönsson, "Intermolecular potentials for the H_2O-C_6H_6

and the C_6H_6-C_6H_6 systems calculated in an ab initio SCF-CI approximation", *J. Am. Chem. Soc.*, **105**, 3777–3782 (1983).

[21] P. Linse, S. Engström and B. Jönsson, "Molecular dynamics simulations of solid and liquid benzene", *Chem. Phys. Letters*, **115**, 95–100 (1985).

[22] A. Wallqvist and G. Karlström, "A New Non-Empirical Force Field for Computer Simulations", *Chemica Scripta*, **29A**, 131–137 (1989).

[23] D.M.E. Szebenyi and K. Moffat, "The Refined Structure of Vitamin D-dependent Calcium-Binding Protein from Bovine Intestine", *J. Biol. Chem.*, **261**, 8761–8777 (1986).

[24] O. Teleman and B. Jönsson, "Vectorising a general purpose molecular dynamics simulation program", *J. Comput. Chem.*, **7**, 58–66 (1986).

[25] A.D. Kline, W. Braun and K. Wüthrich, "Determination of the Complete Three-Dimensional Structure of the α-Amylase Inhibitor Tendamistat in Aqueous Solution by Nuclear Magnetic Resonance and Distance Geometry", *J. Mol. Biol.*, **204**, 675–724 (1988).

[26] G.M. Clore, A.M. Gronenborn, M.N.G. James, M. Kjaer, C.A. McPhalen and F.M. Poulsen, "Comparison of the Solution and X-ray Structures of the Barley Serine Protease Inhibitor 2", *Protein Engin.*, **1**, 313–318 (1987).

[27] G. Wagner, W. Braun, T.F. Havel, T. Schaumann, N. Go and K. Wüthrich, "Protein Structures in Solution by Nuclear Magnetic Resonance and Distance Geometry", *J. Mol. Biol.*, **195**, 611–641 (1987).

[28] B. Wendt, T. Hoffmann, S.R. Martin, P. Bayley, P. Brodin, T. Grundström, E. Thulin, S. Linse and S. Forsen, "Effect of amino acid substitutions and deletions on the thermal stability, the pH stability and unfolding by urea of bovine calbindin D_{9k}", *Eur. J. Biochem.*, **175**, 439–445 (1988).

[29] W.F. van Gunsteren, H.J.C. Berendsen, J. Hermans, W.G.J. Hol and J.P.M. Postma, "Computer simulation of the dynamics of hydrated protein crystals and its comparison with x-ray data", *Proc. Natl. Acad. Sci. USA*, **80**, 4315–4319 (1983). R.M. Levy, R.P. Sheridan, J.W. Keepers, G.S. Dubey, S. Swaminathan and M. Karplus, "Molecular dynamics of myoglobin at 298 K – Results from a 300-ps computer simulation", *Biophys. J.*, **48**, 509–518 (1985).

[30] B. Svensson, B. Jönsson and C. Woodward, "Electrostatic contributions to the binding of Ca^{2+} in calbindin mutants - A Monte Carlo study", *J. Biophys. Chem.*, **38**, 179–183 (1990).

[31] G.-J. Kördel and T. Drakenberg, personal communication.

[32] O. Teleman, "A Stringent Formulation of the Overall Rotational Diffusion in Molecules and Other Flexible Assemblies", *J. Comp. Chem.*, **11**, 64–66 (1990).

[33] R. Rigler *et al.*, personal communication.

[34] H.J. Vogel, T. Drakenberg, S. Forsen, J.D.J. O'Neil and T. Hofmann, "Structural Differences in the Two Calcium Binding Sites of the Porcine Intestinal Calcium-Binding Protein: A Multinuclear NMR Study", *Biochemistry*, **24**, 3870–3876 (1985).