



## Validation of the GROMOS force-field parameter set 45A3 against nuclear magnetic resonance data of hen egg lysozyme

T.A. Soares<sup>a</sup>, X. Daura<sup>b</sup>, C. Oostenbrink<sup>a</sup>, L.J. Smith<sup>c</sup> & W.F. van Gunsteren<sup>a,\*</sup>

<sup>a</sup>Laboratory of Physical Chemistry, ETH Hönggerberg Zürich, 8093 Zürich, Switzerland; <sup>b</sup>Institució Catalana de Recerca i Estudis Avancats and Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, Bellaterra, Spain; <sup>c</sup>Oxford Centre for Molecular Sciences, Central Chemistry Laboratory, University of Oxford, Oxford, United Kingdom

Received 14 July 2004; Accepted 21 October 2004

**Key words:** generalized order parameters, GROMOS force field, hen egg white lysozyme, NOE inter-proton distances, spin–spin coupling constants

### Abstract

The quality of molecular dynamics (MD) simulations of proteins depends critically on the biomolecular force field that is used. Such force fields are defined by force-field parameter sets, which are generally determined and improved through calibration of properties of small molecules against experimental or theoretical data. By application to large molecules such as proteins, a new force-field parameter set can be validated. We report two 3.5 ns molecular dynamics simulations of hen egg white lysozyme in water applying the widely used GROMOS force-field parameter set 43A1 and a new set 45A3. The two MD ensembles are evaluated against NMR spectroscopic data: NOE atom–atom distance bounds,  $^3J_{\text{NH}\alpha}$  and  $^3J_{\alpha\beta}$  coupling constants, and  $^{15}\text{N}$  relaxation data. It is shown that the two sets reproduce structural properties about equally well. The 45A3 ensemble fulfills the atom–atom distance bounds derived from NMR spectroscopy slightly less well than the 43A1 ensemble, with most of the NOE distance violations in both ensembles involving residues located in loops or flexible regions of the protein. Convergence patterns are very similar in both simulations: atom-positional root-mean-square differences (RMSD) with respect to the X-ray and NMR model structures and NOE inter-proton distances converge within 1.0–1.5 ns while backbone  $^3J_{\text{HN}\alpha}$ -coupling constants and  $^1\text{H}$ – $^{15}\text{N}$  order parameters take slightly longer, 1.0–2.0 ns. As expected, side-chain  $^3J_{\alpha\beta}$ -coupling constants and  $^1\text{H}$ – $^{15}\text{N}$  order parameters do not reach full convergence for all residues in the time period simulated. This is particularly noticeable for side chains which display rare structural transitions. When comparing each simulation trajectory with an older and a newer set of experimental NOE data on lysozyme, it is found that the newer, larger, set of experimental data agrees as well with each of the simulations. In other words, the experimental data converged towards the theoretical result.

### Introduction

The structure, flexibility and dynamic behaviour of proteins in solution are of fundamental interest

to their function. In the past 25 years, molecular dynamics (MD) simulation has emerged as a powerful tool to characterize biomolecular structure and dynamics at the atomic level and has helped to shed light on complex molecular processes associated with protein conformational changes (van Gunsteren and Berendsen, 1990; Karplus

\*To whom correspondence should be addressed. E-mail: wfvgn@igc.phys.chem.ethz.ch

and McCammon, 2002). It has become an essential technique even for experimentalists as most of the structures obtained by X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are refined with MD simulation protocols (van Gunsteren et al., 1999). The increased use of MD simulation techniques and results to interpret experimental data is chiefly due to the improvement in the quality of the biomolecular force fields used, which provide more realistic descriptions of microscopic forces on the one hand, and to the continuing growth of computer power, which allows for longer simulations and so for a more thorough sampling of the conformational space of a protein, on the other.

Classical molecular dynamics makes use of Newton's laws of motion to generate a trajectory, which describes the positions and velocities of a set of particles or atoms as a function of time. The forces that govern the atomic motions are derived from atom-atom interactions usually given in the form of an empirical potential energy function or a so-called force field. The choice of interaction function or force field in a biomolecular modeling study is an important one, since it will ultimately determine the quality of the results of the study.

Over the past decades a variety of force fields for biomolecular simulations have been developed. An extensive review of force field development, its principles and procedures can be found in Hünenberger and van Gunsteren (1997). Typical examples of condensed-phase biomolecular force fields are AMBER (Weiner and Kollman, 1981; Pearlman et al., 1995), CHARMM (Brooks et al., 1983; MacKerell Jr. et al., 1995), CHARMM (Momany and Rone, 1992), ECEPP/3 (Nemethy et al., 1992), ENCAD (Levitt, 1983, Levitt et al., 1995), GROMOS (van Gunsteren and Berendsen, 1987; van Gunsteren et al., 1996), and OPLS (Jorgensen and Tirado-Rives, 1988). These biomolecular force fields have a similar functional form of the interaction function, yet they differ considerably in their parameterization philosophy and parameter values. Since the latter can be obtained in a variety of ways, by fitting of a range of molecular properties (geometric, energetic, dynamic, dielectric, etc.) of small molecules against different sets of quantum-mechanical and experimental data regarding these molecules, different parameter sets may

yield widely different results when applied to large, complex biomolecular systems.

Since the early 1980s, the Groningen Molecular Simulation (GROMOS) software package for computer simulation has been developed in conjunction with an interatomic interaction function for MD simulation of biomolecular systems. Major versions of the GROMOS software are GROMOS87 (van Gunsteren and Berendsen, 1987) and GROMOS96 (van Gunsteren et al., 1996). The first set of (non-bonded) GROMOS force-field parameters can be found in Hermans et al., (1984). Since then the force field has continuously been improved and refined (van Gunsteren and Berendsen, 1987; Smith et al., 1995; van Gunsteren et al., 1996, 1998; Daura et al., 1998; Schuler and van Gunsteren, 2000; Schuler et al., 2001; Chandrasekhar and van Gunsteren, 2001, 2002; Chandrasekhar et al., 2003). The most widely used versions of the GROMOS force field are the GROMOS 37C4 force field (van Gunsteren and Berendsen, 1987) of 1985, an improved version (Smith et al., 1995) of it, the GROMOS 43A1 force field (van Gunsteren et al., 1996; Daura et al., 1998) of 1996, and the GROMOS 45A3 force field (Schuler et al., 2001) of 2001. The main features and particularities of the GROMOS force field have been summarized in van Gunsteren et al. (1998) and its functional form can be found in van Gunsteren et al. (1996, 1998) and Scott et al. (1999).

The question of the general quality of a particular force field cannot be easily answered. It will depend on the type of property and molecular system. An impression can be obtained from the literature concerning the application of a particular force field to biomolecular systems for which ample experimental data at the atomic level is available, e.g. from NMR spectroscopy or X-ray diffraction. Over the years, the successive GROMOS force-field parameter sets have been tested by simulations of a variety of proteins, nucleotides, sugars, and lipids. The most recent tests of the 43A1 parameter set involve a series of  $\beta$ -peptides (Daura et al., 1997, 1999, 2002; van Gunsteren et al., 2001), the proteins fatty-acid binding protein (Bakowies and van Gunsteren, 2002), hen egg white lysozyme (Smith et al., 1995, Stocker and van Gunsteren, 2000; Stocker et al., 2000),  $\alpha$ -lactalbumin (Smith et al., 1996, 1999), photo-active yellow protein (Antes et al., 2002), the estrogen receptor

ligand-binding domain (Oostenbrink et al., 2000) and DNA duplexes (Bonvin et al., 1998; Czechitzky et al., 2001). For a review of validation procedures we refer to van Gunsteren and Mark (1998). Although the 43A1 parameter set performed well for peptides and proteins, which contain only short chains of aliphatic carbons, it produced a too high density (at constant experimental pressure) or a negative pressure (at constant experimental density) when applied to lipids or alkane molecules consisting of larger chains of 7–20 CH<sub>2</sub> units (Schuler et al., 2001). This observation led to a reparametrization of the non-bonded interaction parameters of the aliphatic united atom carbons CH<sub>*n*</sub> (*n* = 0–4). The resulting 45A3 parameter set reproduces the heat of vaporization and density of a series of alkanes and at the same time their solvation free energy in water (Schuler et al., 2001).

Force-field development proceeds in tandem steps. Each new version is parametrized against selected physical properties of a given number of small compounds and then characterized and validated for systems of interest, i.e., biologically relevant molecules. The new 45A3 parameter set differs from the previous set 43A1 with respect to the van der Waals coefficients describing the interactions between the oxygen atom type OW of the SPC water model and aliphatic carbons of atom type CH<sub>*n*</sub> (*n* = 1–3) and between these aliphatic carbons and all other atom types. In the present study, we evaluate the performance of the GROMOS parameter set 45A3 when applied to a protein. We chose the protein hen egg white lysozyme (HEWL), because high-quality experimental data at the atomic level are available from X-ray diffraction (Carter et al., 1997) and NMR spectroscopy (Schwalbe et al., 2001), and because it has been used in tests (Smith et al., 1995; Stocker and van Gunsteren, 2000; Stocker et al., 2000) of previous versions of the GROMOS force field. The comparison of simulated with experimental data is made on the basis of analysis of atom-positional root-mean-square deviation (RMSD) from the X-ray and NMR structures and root-mean-square atom-positional fluctuations (RMSF), <sup>1</sup>H NOE inter-proton distances, spin–spin coupling constants (<sup>3</sup>*J*<sub>NHα</sub> and <sup>3</sup>*J*<sub>αβ</sub>) and <sup>15</sup>N relaxation data of the backbone amide groups and the side-chain NH groups. In addition, we have also verified convergence prop-

erties of these quantities as a function of the force-field parameter set and different time periods along the simulation.

## Methods

MD simulations were performed using the GROMOS suite of programs (van Gunsteren et al., 1996; Scott et al., 1999) and two united-atom parameter sets 43A1 (van Gunsteren et al., 1996) and 45A3 (Schuler et al., 2001). Initial coordinates were taken from the X-ray structure of hen egg white lysozyme solved at a resolution of 0.15 nm (Carter et al., 1997). Ionization states of residues were chosen according a pH of 7.0. The amino acids Asp, Glu, and the C-terminus were deprotonated and Lys, Arg, and the N-terminus were protonated; His was protonated at one of the imidazole nitrogens, He2. This led to a total charge of +8 electrons. The SPC water model (Berendsen et al., 1981) was used to describe the solvent molecules. Water molecules were added around the protein within a truncated octahedron with a distance of 8.12 nm between the square surfaces. After energy minimization and a short MD equilibration of the solvent, counterions were added by replacing water molecules with the highest electrostatic potential at their oxygen atom by chlorine atoms. The system achieved overall neutrality with the addition of 8 Cl<sup>-</sup> ions, resulting in a total number of 1321 solute atoms, 8 ions and 8357 water molecules. Non-bonded interactions were treated using a triple-range approach (Scott et al., 1999) with short and long range cutoffs of 0.8 and 1.4 nm, respectively. The pair list for short-range non-bonded interactions and the intermediate-range non-bonded forces and energies were updated every 10 fs. Interactions beyond the long-range cutoff were represented by a dielectric continuum with a relative dielectric permittivity ε<sub>r</sub> of 66 (Glättli et al., 2002). All bond lengths and the bond angle of the water molecules were constrained using the SHAKE algorithm (Ryckaert et al., 1977) with a tolerance of 10<sup>-4</sup>.

Equilibration of the solute was performed in a stepwise fashion at a constant temperature of 300 K and pressure of 1 atm. Initial velocities

were taken from a Maxwell distribution at 300 K. All protein atoms were harmonically restrained using a force constant  $f = 5.0 \times 10^4 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , which was gradually reduced to  $f = 3.75, 2.50$  and  $1.25 \times 10^4 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  at intervals of 20 ps. The whole system (solute, ions plus solvent) was equilibrated for an additional 20 ps, totalizing 100 ps of equilibration at 300 K. A time step of 2 fs was used during the MD simulations. Data production was carried out for the next 3.5 ns at 300 K and 1 atm. The temperatures of solute (including ions) and solvent were controlled by coupling them separately to a Berendsen thermostat (Berendsen et al., 1984) with a relaxation time of 0.1 ps. The pressure was maintained by means of isotropic coordinate scaling (Berendsen et al., 1984) with a coupling constant of 0.5 ps and a compressibility of  $4.575 \times 10^{-4} (\text{kJ mol}^{-1} \text{ nm}^{-3})^{-1}$ . Configurations of the trajectory were recorded every 0.4 ps. All NMR quantities were calculated as implemented in the program GROMOS (van Gunsteren et al., 1996; Scott et al., 1999).

## Results and discussion

### *Root-mean-square deviation from the X-ray and NMR model structures*

Atom-positional root-mean-square deviations were determined for backbone atoms in the MD trajectories with respect to their positions in the X-ray (dashed line) and NMR lowest-energy (solid line) structures (Figure 1). The parameter sets 43A1 and 45A3 display the same average RMS deviation of 0.18 nm from the X-ray structure, and 0.21 nm from the lowest-energy NMR model structure. The RMSD values for the parameter set 45A3 reach a first plateau value of 0.15 nm from the X-ray structure and 0.19 nm from the NMR model structure right after the equilibration period, whereas the RMSD values for the parameter set 43A1 reach the same plateau values after a period of 600 ps. After that both MD trajectories present similar RMSD profiles, fluctuating around 0.2 nm beyond 2.0 ns (Figure 1). Both MD trajectories stayed closer to the X-ray structure, in terms of RMSD, than to the lowest-energy NMR conformer.

### *Atom-positional root-mean-square fluctuations and secondary structure*

RMS atom-positional fluctuations of alpha-carbons were calculated for the final 2 ns of simulation to ensure full convergence (Hünenberger et al., 1995) (Figure 2). The 45A3 trajectory displays larger atomic fluctuations than the 43A1 trajectory. Figure 3 shows that there is good agreement in terms of secondary structure pattern between the two MD trajectories and the set of 50 NMR structures. The only exception is the region corresponding to the third  $\beta$ -strand of the anti-parallel  $\beta$ -sheet (residues Ile58 and Asn59) which is present in the X-ray structure and in the MD simulations but could not be identified in the set of 50 NMR model structures used in this study. In fact, Schwalbe et al. (2001) have shown that the inclusion of the dipolar coupling restraints in the NMR structure calculations induces slight alterations in the relative orientations of the amide and carbonyl pairs in the  $\beta$ -strand region when compared to the model structures obtained by using the same NOE restraints but in the absence of dipolar coupling restraints. These alterations result in hydrogen bonds no longer being recognized by the Kabsch and Sander criteria (Kabsch and Sander, 1983). In the NMR model structures obtained without dipolar coupling data, a  $\beta$ -bridge involving Asn59 in the third strand can be observed (Schwalbe et al., 2001).

### *NOE inter-proton distances*

Two data sets of inter-proton distances corresponding to NOE intensities have been determined for lysozyme (Schwalbe et al., 2001, Smith et al., 1993). One set (set1) is constituted of 1079 NOE bounds (Smith et al., 1993) and a more recent one (set2) comprises 1630 NOE bounds (Schwalbe et al., 2001). These bounds include pseudo-atom distance corrections as given by Wüthrich et al. (1983). Inter-proton distances were calculated for the structures in the two MD trajectories and were averaged using  $r^{-3}$  averaging (Tropp, 1980). Comparisons of the  $r^{-3}$ -averaged inter-proton distances with the NMR derived upper bounds (including pseudo-atom corrections) are summarized in Table 1 and their distribution is shown in Figure 4.

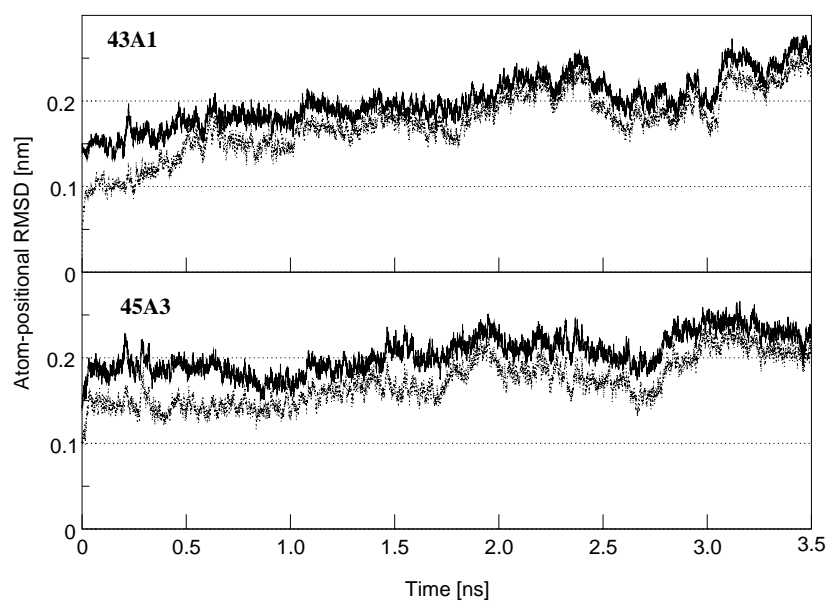


Figure 1. Root-mean-square deviations (RMSD) of backbone atoms N, C, CA and O of HEWL from the X-ray structure (1AKI) (dashed lines) and from the lowest-energy NMR structure (1E8L) (solid lines) as function of time and for the two GROMOS force-field parameter sets 43A1 (upper panel) and 45A3 (lower panel). Rotational and translational fitting of pairs of structures was applied using all CA atoms.

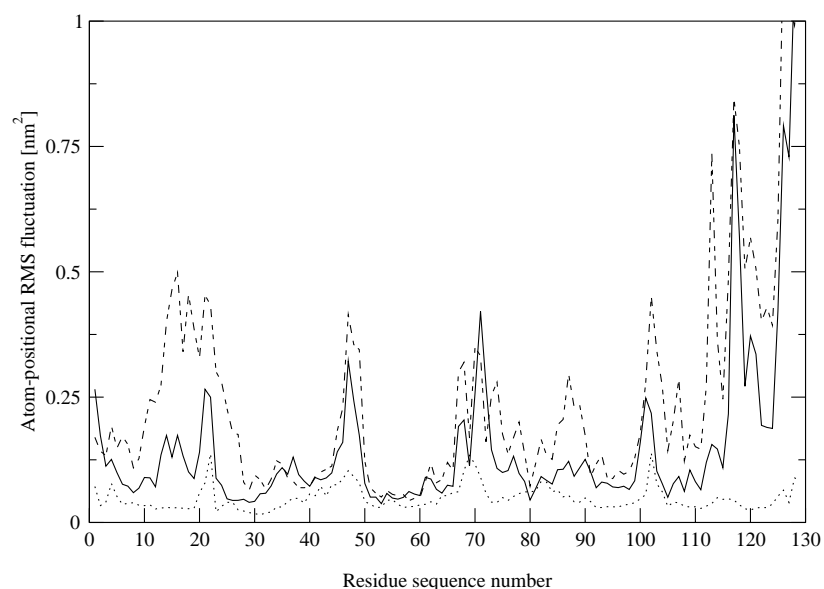


Figure 2. Root-mean-square atom-positional fluctuations (RMSF) of CA atoms of HEWL as function of residue sequence number, calculated for the final 2 ns of the two MD trajectories (parameter set 43A1: solid line; parameters set 45A3: dot-dashed line) and from the set of 50 NMR model structures (1E8L) (dotted line).

They are presented as distance bound violations, i.e., distances averaged over the ensemble minus the corresponding NMR derived upper distance

bound. This difference can also adopt negative values, which means that in the MD simulations the inter-proton distance is on average shorter

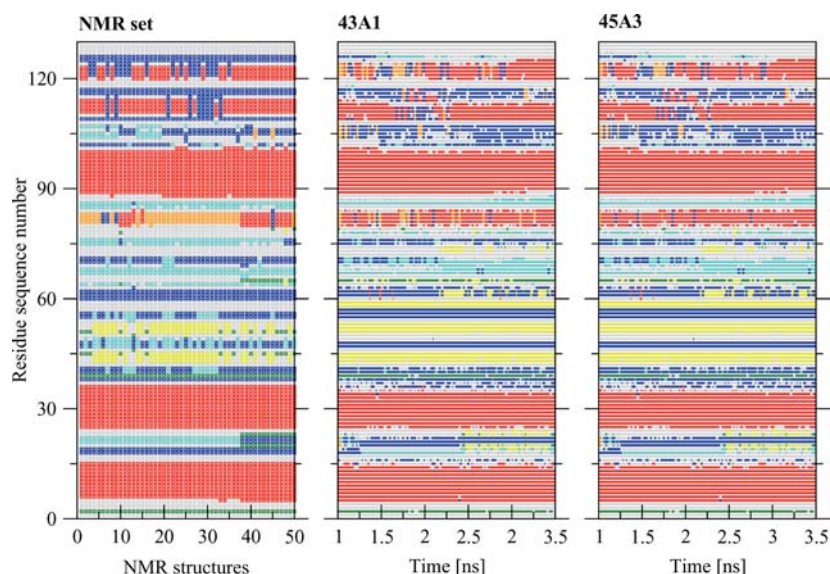


Figure 3. Secondary structure of HEWL as function of time for the two MD simulations using different force-field parameter sets (older 43A1, newer 45A3) and as function of structure sequence number for a set of 50 NMR model structures (Schwalbe et al., 2001). Secondary structure definition according to (Kabsch and Sander, 1983):  $\alpha$ -helix (red),  $3_{10}$ -helix (orange),  $\beta$ -bridge (green),  $\beta$ -strand (yellow), bend (cyan), turn (blue), undefined secondary structure (grey). In the MD trajectories each dot represents a period of 4 ps.

than the upper bound derived from the NMR experiment (Figure 4). The very negative values smaller than  $-0.4$  nm are due to atoms that are close to each other and for which the upper bound is relatively large, in the range of 0.64–1.18 nm. All inter-proton distances were larger than 0.18 nm, the standard value adopted for the lower bound of inter-proton distances. The total number of NOE violations corresponding to different time intervals is consistently higher in the 45A3 ensemble than in the 43A1 one. This indicates that the newer force-field parameter set 45A3 produces slightly more and a few larger (see Figure 4) violations than the older 43A1 set. However, the observed violations may not be very significant, since experimental NOE bounds may contain sizeable uncertainties. In this regard we note that when comparing the violations of a single MD simulation with respect to the two sets of experimental NOE bounds, the MD simulations (for both force-field parameter sets) show slightly lower average violations for the newer, larger set (set2) of experimental bounds than with the older, smaller one (set1). Between the years 1993 (set1) and 2001 (set2) the experimental data thus slightly converged towards the simulated ones.

#### Spin–spin coupling constants

Three-bond  ${}^3J_{\text{HN}\alpha}$ - and  ${}^3J_{\alpha\beta}$ -coupling constants have been calculated from the two MD ensembles and from the set of 50 NMR model structures and compared against 95  ${}^3J_{\text{HN}\alpha}$  and 100 (57 + 43)  ${}^3J_{\alpha\beta}$  experimental coupling constants (Smith et al., 1991). The convergence properties of these coupling constants along different time periods are presented in Figure 5 and comparisons of experimental and calculated  ${}^3J_{\text{HN}\alpha}$ - and  ${}^3J_{\alpha\beta}$ -coupling constants are presented in Figure 6 and Table 2, respectively. In both simulations the  ${}^3J_{\text{HN}\alpha}$ -coupling constants converge within about 1.5 ns (Figure 5a). The largest differences are observed between coupling constants averaged over the initial time period of 0.0–0.5 ns and the remaining intervals, for which the averaged  ${}^3J_{\text{HN}\alpha}$  values change very little. In contrast, several average  ${}^3J_{\alpha\beta}$ -coupling constants display large variations when considering different averaging time periods (Figure 5b). For a number of residues convergence is not reached in the time scale simulated. Still longer simulations are required to accurately predict the side-chain conformational distribution probed by coupling constants, which

Table 1. NOE bound violations computed from two MD trajectories (43A1/45A3) against two sets of experimental NOE distance bounds from Smith et al. (1993), (set1, 1079 bounds) and from Schwalbe et al. (2001), (set2, 1630 bounds)

Averaging period (ns)	Number of violations (set1)			Mean violation $\langle R_E - R_0 \rangle$	
	$> 0.1$ nm	$> 0.2$ nm	$> 0.3$ nm		
0.0–0.5	19/37	7/10	1/3	0.006/0.009	
0.5–1.5	28/31	5/15	1/2	0.007/0.009	
1.5–3.5	29/46	7/20	2/11	0.008/0.012	
0.0–3.5	21/34	4/12	1/3	0.006/0.009	
	Number of violations (set2)				
	$> 0.1$ nm	$> 0.2$ nm	$> 0.3$ nm		
0.0–0.5	26/51	7/14	0/2	0.005/0.008	
0.5–1.5	35/50	6/20	0/4	0.007/0.009	
1.5–3.5	42/63	13/26	2/11	0.008/0.011	
0.0–3.5	31/47	4/15	0/3	0.006/0.008	
Comparison of the largest NOE violations of the experimental upper bound distances from set1 (1079)					
NOE pair	43A1		45A3		
	Exp. $R_0$	$R_E$	$R_E - R_0$	$R_E$	$R_E - R_0$
<i>8LeuH<math>\delta</math>-17LeuH<math>\gamma</math></i>	0.54	0.956	0.416	0.991	0.451
<i>23TyrHN-28TrpH<math>\zeta</math>2</i>	0.30	0.419	0.119	0.660	0.360
<i>35GluH<math>\beta</math>-38PheH<math>\epsilon</math></i>	0.75	1.013	0.263	1.117	0.367
Comparison of the largest NOE violations of the experimental upper bound distances from set2 (1630)					
NOE pair	43A1		45A3		
	Exp. $R_0$	$R_E$	$R_E - R_0$	$R_E$	$R_E - R_0$
<i>23TyrHN-28TrpH<math>\zeta</math>2</i>	0.30	0.419	0.119	0.660	0.360
<i>35GluH<math>\beta</math>-38PheH<math>\epsilon</math></i>	0.75	1.013	0.263	1.117	0.367
<i>25LeuH<math>\beta</math>-18AspH<math>\beta</math>1</i>	0.55	0.588	0.038	0.881	0.331

Results are obtained for time periods of 0.0–3.5 ns. The values for the 43A1 and 45A3 trajectories are separated by the symbol /. Largest NOE violations: values  $\geq 0.3$  nm.  $R_0$  is the experimentally derived distance bound including pseudo-atom correction (Wüthrich et al., 1983).  $R_E$  is the distance obtained from the MD trajectory using  $r^{-3}$  averaging.  $R_E - R_0$  indicates a violation. A total of 1079 (set1) (Smith et al., 1993) and 1630 (set2) (Schwalbe et al., 2001) inter-proton NOE bounds have been used in the analysis and when calculating the mean ( $\langle \dots \rangle$ ). Distances are in nm.

are experimentally averaged over millisecond time periods. The 95  $^3J_{\text{HN}\alpha}$ -coupling constants calculated from the 43A1 (solid circles) and 45A3 (open squares) ensembles reproduce the experimental values within an RMSD of 1.6 Hz ( $R$ -values of 19% and 18%, respectively, see Figure 6). Only a few residues present coupling constants which differ by more than 1.5 Hz between the two parameter sets, Tyr20 (5.5 Hz), Ser36 (9.6 Hz), Ile78 (8.0 Hz), Asn103 (8.2 Hz), Trp108 (9.6 Hz), where the experimental values are given between parentheses.

In the case of the 100  $^3J_{\alpha\beta}$ -coupling constants, the average  $^3J$ -values calculated from the MD

simulations show a fairly large deviation from the experimental ones with RMSD values of 2.0 Hz for both ensembles 43A1 and 45A3 ( $R$ -values of 24% and 23%, respectively, see Table 2). In hen egg white lysozyme, two classes of residues can be distinguished based on their side-chain conformations (Smith et al., 1991): residues where the torsional angle  $\chi_1$  exhibits only one staggered conformation and residues where  $\chi_1$  takes up multiple conformations. The values of  $^3J_{\alpha\beta}$  for the two  $\beta$  protons in the (staggered) conformations of the dihedral angle  $\chi_1$  are either both 3.4 Hz (for  $\chi_1 = 60^\circ$ ) or 12.9 and 3.4 Hz (for  $\chi_1 = 180^\circ, -60^\circ$ ) (Karplus, 1959),

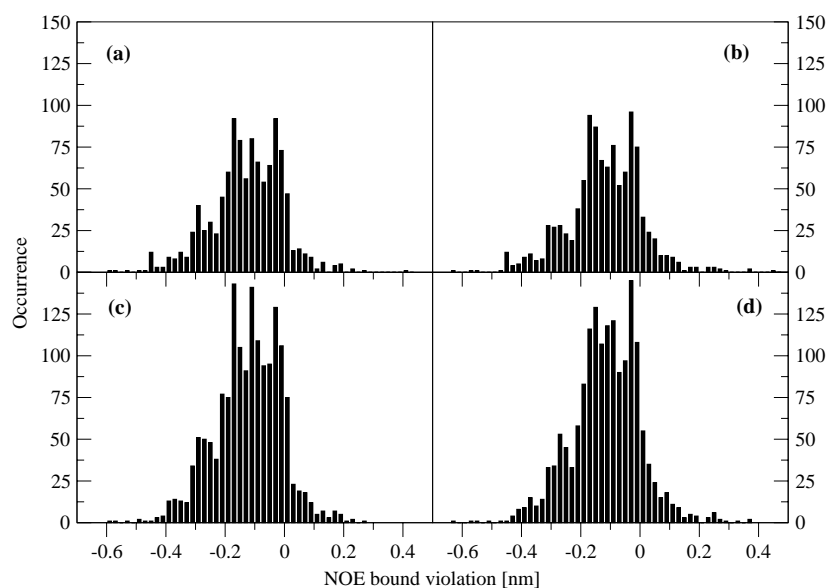


Figure 4. Occurrence of  $r^{-3}$ -averaged  $^1\text{H}$ - $^1\text{H}$  NOE distances in two simulations of HEWL with respect to two sets of NMR derived NOE distance bounds. Positive values represent violations of the experimental NOE upper distance bounds. NOE  $r^{-3}$ -averaged distances were calculated from two 3.5 ns MD trajectories. Top panels: 1079 NOE distances (experimental set1) (Smith et al., 1993) calculated from the MD simulations based on the force-field parameter sets 43A1 (a) and 45A3 (b). Bottom panels: 1630 NOE distances (experimental set2) (Schwalbe et al., 2001) calculated from the MD simulations based on the force-field parameter sets 43A1 (c) and 45A3 (d). Distances in nm.

and for  $^3J_{\alpha\beta}$ -values originating from multiple conformations in the range 6–8 Hz (Smith et al., 1991). Predictably, a better agreement between experimental and calculated  $^3J_{\alpha\beta}$ -values from the MD simulations is observed for  $^3J_{\alpha\beta}$ -values in the range 3.5–5.0 Hz and in the range 10.0–12.0 Hz (Table 2) where averaging effects are very restricted or even absent for some residues. Most of the  $^3J_{\alpha\beta}$ -coupling constants that diverge significantly between the two parameter sets are in the range 5–11 Hz. Several of the side chains involved are experimentally identified as not occupying a single staggered conformation about  $\chi_1$  (Smith et al., 1991), e.g. Glu7, Asn19, Arg45, Ser85. The remaining ones do display different behaviour for the torsional angle involved as function of time between the two simulations, e.g. Tyr20, Val29, Thr69, Leu75, Thr89 (Figure 7). Over the time period of 3.5 ns shown, too few  $\chi_1$  torsional angle transitions are occurring for a proper averaging of the  $^3J_{\alpha\beta}$ -coupling constant. This means that the large discrepancies found in Table 2 between the two parameter sets are mainly due to insufficient sampling

of the  $\chi_1$  torsional angle degree of freedom. Table 3 shows that the number of  $^3J$ -coupling constants that differ by more than 1 Hz from the experimental values is not very different between the two MD ensembles 43A1 and 45A3 and the set of NMR model structures. The deviation between the experimental  $^3J$ -value data and the 50 NMR model structures reflects the fact that broad torsional angle ranges ( $\geq 60^\circ$ ), rather than the coupling constant values themselves, were used as restraints in the NMR structure calculations. In addition,  $\chi_1$  torsional angle restraints were not used for residues whose side chain did not occupy a single staggered conformation.

#### $^{15}\text{N}$ relaxation data

Backbone  $^1\text{H}$ - $^{15}\text{N}$  order parameters calculated from the MD simulations along different time periods are presented in Figure 8. In both trajectories, convergence of  $S^2$  is faster in regions of secondary structure and is achieved within about 1 ns. This behaviour is expected for N–H bond



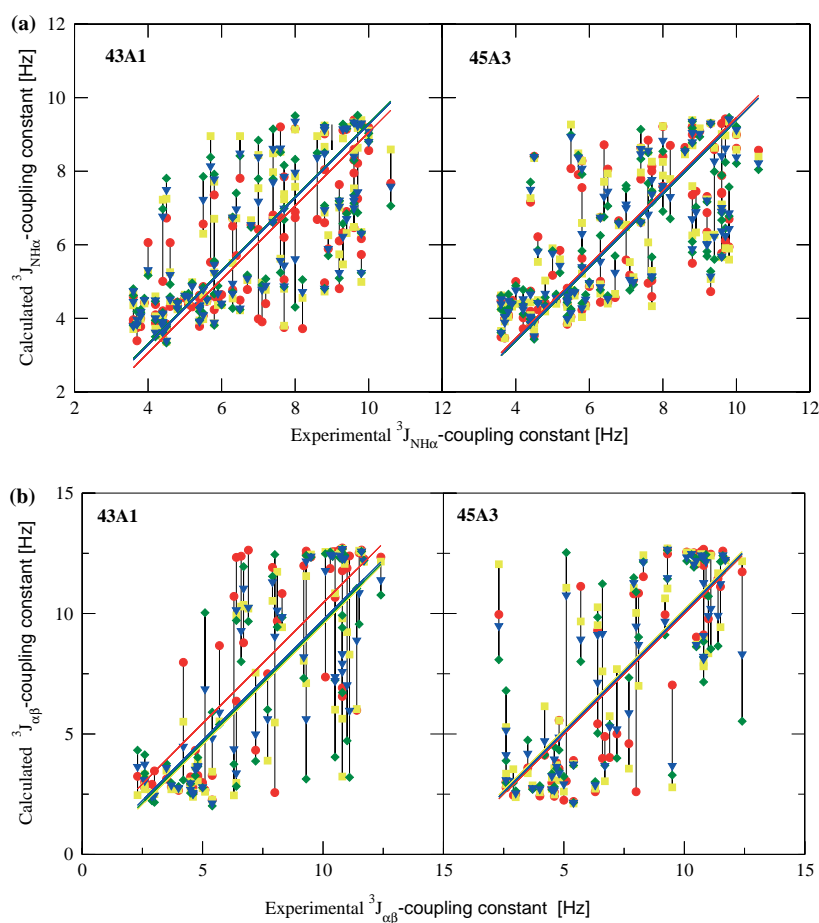


Figure 5.  $^3J$ -coupling constants of HEWL calculated from the MD trajectories for the force-field parameter sets 43A1 (left) and 45A3 (right), and averaged over different time windows: 0.0–0.5 ns (red circles), 0.5–1.5 ns (yellow squares) and 1.5–3.5 ns (green diamonds) and 0.0–3.5 ns (blue triangles). (a) Top panels correspond to  $95\ ^3J_{\text{NH}\alpha}$ -coupling constants (see Figure 6). (b) Bottom panels correspond to  $100\ ^3J_{\alpha\beta}$ -coupling constants (see Table 2). The calculated coupling constants belonging to the same data set are connected by a line.

vectors that sample approximately the same region of the conformational space during the different time periods considered. In our simulations, the regions containing loops and termini of helices required longer convergence times, up to a few nanoseconds.

In Figure 9 the simulated  $^1\text{H}$ – $^{15}\text{N}$  backbone order parameters are compared against experimental  $^1\text{H}$ – $^{15}\text{N}$  order parameters (Buck et al., 1995). The two MD ensembles (solid line 43A1, dot-dashed line 45A3) exhibit similar amplitudes and profiles with respect to backbone N–H mobility. On average, the experimental order parameters (dashed line) are larger than the values calculated from the MD simulations. Experimentally, Gly22 presents an unusually high  $S^2$  value (1.0) (Buck

et al., 1995). However, Gly22 exhibits motions on the microsecond to millisecond timescale with a  $T_1/T_2$  ratio of 4.18 as compared to an average  $T_1/T_2$  ratio of  $3.32 \pm 0.13$  for HEWL (Buck et al., 1995). The order parameter calculated from our simulations suggests that Gly22 is more flexible, in agreement with these results. Experimental  $^1\text{H}$ – $^{15}\text{N}$  order parameters are also available for the side-chain NH groups of 28 residues in hen egg white lysozyme (Buck et al., 1995), NH2 of Asn and Gln, N $\epsilon$ H of Arg, and N $\epsilon$ 1H of Trp residues. In general, the MD ensembles slightly underestimate the disorder of the N–H vectors (Figure 10). This can be explained from insufficient sampling of side-chain conformations on the nanosecond time scale considered. There is a clear improvement in

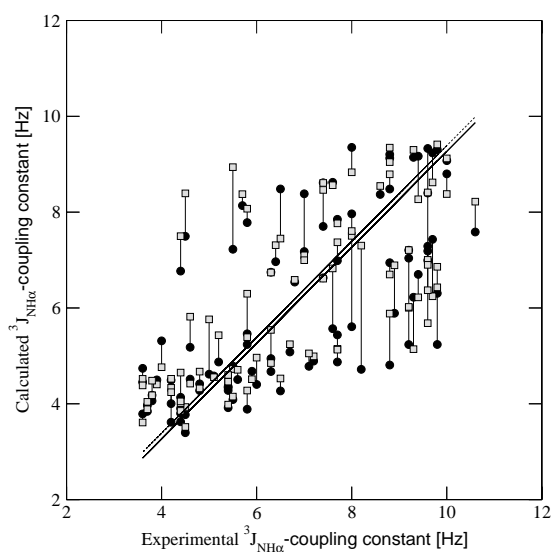


Figure 6. Comparison of experimental and calculated  $^3J$ -coupling constants for HEWL. The force-field parameter set 43A1 is represented by circles and the set 45A3 by squares.  $95\ ^3J_{\text{HN}\alpha}$ -coupling constants were calculated from the two 3.5 ns MD trajectories using the Karplus relationship (Karplus, 1959) with parameters  $a = 6.4$  Hz,  $b = -1.4$  Hz,  $c = 1.9$  Hz (Pardi et al., 1984). Experimental data were taken from Smith et al. (Smith et al., 1991). The values for the glycine residues were omitted, since they were not stereospecifically assigned.

Table 2.  $\text{C}_\alpha\text{H}-\text{C}_\beta\text{H}$   $^3J_{\alpha\beta}$ -coupling constants for hen egg white lysozyme as measured by NMR, and computed from two MD trajectories (43A1, 45A3) and from a set of 50 NMR model structures

Amino acid residue	Experiment		43A1		45A3		NMR structures	
	$\alpha\beta_A$	$\alpha\beta_B$	$\alpha\beta_2$	$\alpha\beta_3$	$\alpha\beta_2$	$\alpha\beta_3$	$\alpha\beta_2$	$\alpha\beta_3$
2Val	10.8		7.9		8.1		12.7	
3Phe	3.0	10.0	10.7	2.4 <sup>a</sup>	12.0	2.5 <sup>a</sup>	12.1	2.4 <sup>a</sup>
6Cys	3.5	11.5	12.4	3.6 <sup>a</sup>	11.0	4.2 <sup>a</sup>	11.3	6.5 <sup>a</sup>
7Glu	6.7	6.4	4.1	11.0 <sup>a</sup>	11.8 <sup>a</sup>	3.7	7.0 <sup>a</sup>	5.8
13Lys	9.2	5.1	6.0	8.2 <sup>a</sup>	4.5	9.7 <sup>a</sup>	2.4	10.3 <sup>a</sup>
15His	2.6	11.2	11.1	3.7 <sup>a</sup>	9.2	5.1 <sup>a</sup>	11.0	1.8 <sup>a</sup>
18Asp	11.0	4.2	3.2	7.0 <sup>a</sup>	4.5	9.9 <sup>a</sup>	6.6	11.2 <sup>a</sup>
19Asn	6.4	7.3	3.4 <sup>a</sup>	11.7	7.1 <sup>a</sup>	7.5	3.2 <sup>a</sup>	8.1
20Tyr	2.3	11.7	3.6 <sup>a</sup>	9.3	9.5	5.0 <sup>a</sup>	2.5 <sup>a</sup>	12.6
23Tyr	10.9	2.7	12.3 <sup>a</sup>	2.9	12.1 <sup>a</sup>	3.6	11.9 <sup>a</sup>	2.0
27Asn	10.3	2.4	12.4 <sup>a</sup>	2.9	12.5 <sup>a</sup>	3.1	12.6 <sup>a</sup>	4.2
28Trp	10.7	4.1	12.4 <sup>a</sup>	3.8	12.4 <sup>a</sup>	3.2	12.9 <sup>a</sup>	3.3
29Val	11.1		6.0		10.2		12.9	
30Cys	10.8	5.3	3.7	12.7 <sup>a</sup>	2.9	12.3 <sup>a</sup>	2.1	12.1 <sup>a</sup>
34Phe	5.0	10.7	12.3	2.8 <sup>a</sup>	12.1	2.9 <sup>a</sup>	10.1	1.8 <sup>a</sup>
37Asn	8.1	4.2	10.1 <sup>a</sup>	4.8	8.7 <sup>a</sup>	4.2	7.1 <sup>a</sup>	5.1
39Asn	10.8	4.5	2.8	12.3 <sup>a</sup>	3.5	11.3 <sup>a</sup>	2.4	12.0 <sup>a</sup>
40Thr	4.5		3.0		2.7		7.7	
43Thr	3.7		2.9		2.8		6.8	
45Arg	6.9	6.7	4.5	10.2 <sup>a</sup>	9.7 <sup>a</sup>	5.3	4.4	6.5 <sup>a</sup>
46Asn	4.7	11.2	12.6	3.5 <sup>a</sup>	12.6	3.6 <sup>a</sup>	11.5	6.3 <sup>a</sup>
47Thr	2.6		3.1		4.1		3.0	
48Asp	2.6	3.7	4.0	3.1 <sup>a</sup>	4.0	3.1 <sup>a</sup>	2.9 <sup>a</sup>	4.2

Table 2. Continued.

Amino acid residue	Experiment		43A1		45A3		NMR structures	
	$\alpha\beta_A$	$\alpha\beta_B$	$\alpha\beta_2$	$\alpha\beta_3$	$\alpha\beta_2$	$\alpha\beta_3$	$\alpha\beta_2$	$\alpha\beta_3$
51Thr	9.3		12.2		11.5		11.6	
52Asp	11.6	3.6	12.6 <sup>a</sup>	3.7	12.4 <sup>a</sup>	3.2	12.7 <sup>a</sup>	3.5
53Tyr	10.4	3.0	12.4 <sup>a</sup>	2.6	12.0 <sup>a</sup>	2.3	12.8 <sup>a</sup>	3.6
59Asn	5.4	11.3	2.1 <sup>a</sup>	11.6	2.1 <sup>a</sup>	12.0	3.2 <sup>a</sup>	12.9
61Arg	10.8	5.7	3.7	8.3 <sup>a</sup>	3.9	8.2 <sup>a</sup>	1.9	11.2 <sup>a</sup>
64Cys	2.7	4.6	4.6	2.6 <sup>a</sup>	4.4	2.7 <sup>a</sup>	7.0	1.8 <sup>a</sup>
65Asn	11.4	4.5	3.9	8.9 <sup>a</sup>	5.0	9.9 <sup>a</sup>	2.7	12.2 <sup>a</sup>
66Asp	5.1	4.5	3.8	6.9 <sup>a</sup>	3.1	10.8 <sup>a</sup>	4.5 <sup>a</sup>	2.5
68Arg	4.8	6.5	11.4	3.6 <sup>a</sup>	9.0	4.9 <sup>a</sup>	8.9	4.1 <sup>a</sup>
69Thr	9.3		5.6		12.7		11.1	
72Ser	5.4	7.6	4.8 <sup>a</sup>	10.6	3.7 <sup>a</sup>	10.9	5.5 <sup>a</sup>	7.5
74Asn	10.5	3.9	3.3	7.2 <sup>a</sup>	2.6	12.2 <sup>a</sup>	1.9	11.2 <sup>a</sup>
75Leu	12.4	2.1	11.4 <sup>a</sup>	3.6	8.3 <sup>a</sup>	6.1	9.8 <sup>a</sup>	1.9
77Asn	8.3	5.9	9.9 <sup>a</sup>	4.6	12.1 <sup>a</sup>	3.4	2.7	9.8 <sup>a</sup>
85Ser	5.7	7.4	9.3	5.9 <sup>a</sup>	5.6 <sup>a</sup>	8.9	5.8	5.3 <sup>a</sup>
86Ser	6.4	4.1	10.1 <sup>a</sup>	4.5	9.1 <sup>a</sup>	4.0	5.7	7.7 <sup>a</sup>
87Asp	11.5	5.1	3.5	10.9 <sup>a</sup>	4.2	11.2 <sup>a</sup>	2.4	12.4 <sup>a</sup>
88Ile	4.5		2.7		3.3		1.8	
89Thr	9.5		12.4		3.7		12.2	
92Val	10.1		11.8		12.3		12.7	
93Asn	10.8	3.5	7.6 <sup>a</sup>	7.6	9.1 <sup>a</sup>	5.7	9.6 <sup>a</sup>	5.5
94Cys	4.0	12.2	2.8 <sup>a</sup>	12.5	2.7 <sup>a</sup>	12.4	1.8 <sup>a</sup>	10.5
99Val	6.3		4.4		2.9		2.1 <sup>a</sup>	
100Ser	7.7	4.0	5.6 <sup>a</sup>	4.8	5.9	6.7 <sup>a</sup>	4.9	8.3 <sup>a</sup>
101Asp	6.6	5.6	5.6	9.3 <sup>a</sup>	5.1	9.2 <sup>a</sup>	4.9	6.7 <sup>a</sup>
106Asn	10.5	3.6	7.4 <sup>a</sup>	5.6	8.7 <sup>a</sup>	5.7	10.1 <sup>a</sup>	5.1
109Val	8.0		9.0		9.5		9.1	
118Thr	4.2		4.5		4.7		3.3	
119Asp	11.7	4.9	2.5	12.2 <sup>a</sup>	2.8	12.2 <sup>a</sup>	1.8	10.3 <sup>a</sup>
123Trp	2.9	10.6	11.3	2.4 <sup>a</sup>	11.9	3.1 <sup>a</sup>	10.3	1.8 <sup>a</sup>
124Ile	4.6		2.5		4.0		2.3	
125Arg	7.9	6.1	11.3 <sup>a</sup>	3.3	11.3 <sup>a</sup>	3.3	2.7	9.4 <sup>a</sup>
127Cys	4.8	11.6	11.4	3.7 <sup>a</sup>	11.3	3.4 <sup>a</sup>	10.2	2.2 <sup>a</sup>
128Arg	7.2	7.9	9.5	5.0 <sup>a</sup>	9.5	5.2 <sup>a</sup>	6.5	5.9 <sup>a</sup>

The experimental values were taken from Table 3 of Smith et al. (1991). There the  $\beta_A$  and  $\beta_B$  values were defined as the proton attached to the  $\beta$  carbon whose resonance has the higher and lower chemical shift, respectively. In the simulations the stereospecific assignments are known and the  $\beta_2$  and  $\beta_3$  protons are defined according to standard rules (Markley et al., 1998). For residues with only one  $\beta$  proton the values are listed in the respective left-hand columns. The  $^3J_{\alpha\beta}$ -values were calculated using the Karplus relationship (Karplus, 1959) with parameters  $a = 9.5$ ,  $b = -1.6$ ,  $c = 1.8$  Hz (DeMarco et al., 1978). Simulation values were obtained for time periods of 0.0–3.5 ns using the two force-field parameter sets 43A1 and 45A3.  $^3J$ -values are in Hz.

<sup>a</sup> These simulated values were assumed to correspond to the experimental values in the column  $\alpha\beta_A$  and the remaining simulated values to the experimental ones in column  $\alpha\beta_B$  when calculating the differences between simulated and experimental  $^3J_{\alpha\beta}$ -values (Table 3).

the correlation between experimental and calculated  $S^2$  values for the new parameter set 45A3 ( $R = 0.74$ ) compared to the old parameter set 43A1 ( $R = 0.69$ ).

## Conclusions

This paper presents the results of two 3.5 ns molecular dynamics simulations of hen egg white

Table 3. Number of  ${}^3J$ -coupling constants for which the absolute difference between the experimentally measured  ${}^3J$ -coupling constants on the one hand and the calculated  ${}^3J$ -coupling constants averaged (time window of 0.0–3.5 ns) over the MD trajectories generated using the force-field parameter sets 43A1 and 45A3 and averaged over the set of 50 NMR model structures (PDB code: 1E8L) on the other hand, is larger than a given value

$ {}^3J_{\text{exp}} - {}^3J_{\text{calc}} $	${}^3J_{\text{HN}\alpha}$ coupling constants (Hz)			${}^3J_{\alpha\beta}$ coupling constants (Hz)		
	Simulations		NMR set	Simulations		NMR set
	43A1	45A3		43A1	45A3	
$> 1$ Hz	47	39	35	61	61	59
$> 2$ Hz	22	20	12	32	31	31
$> 3$ Hz	5	11	5	14	11	10
$> 4$ Hz	1	1	1	5	4	1
$> 5$ Hz	0	0	0	1	3	0

A total of 95  ${}^3J_{\text{HN}\alpha}$ - and 100  ${}^3J_{\alpha\beta}$ -coupling constants have been used in the comparison of experimental (Smith et al., 1991) and simulated data. The identification of simulated with experimental  ${}^3J_{\alpha\beta}$ -values is specified in Table 2.  ${}^3J$ -coupling constants are in Hz.

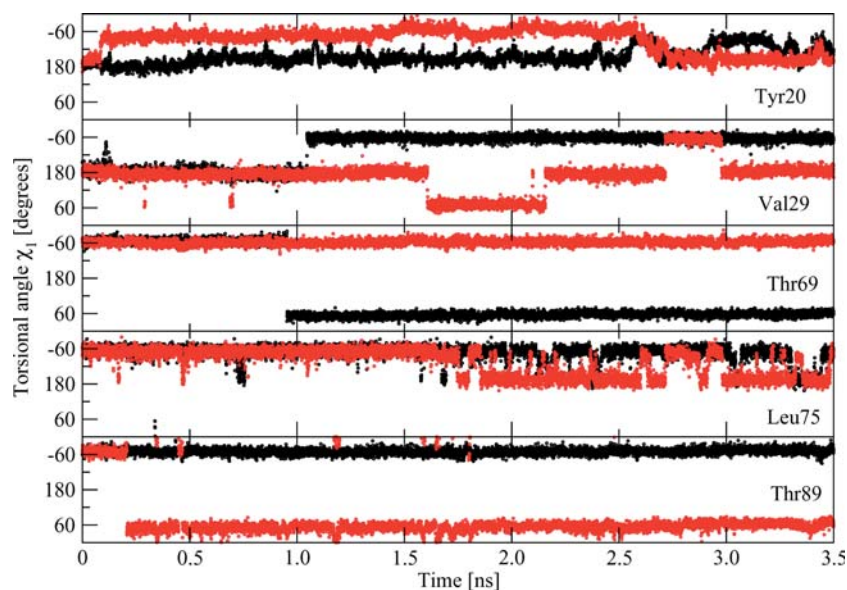


Figure 7. Time series of the torsional angle  $\chi_1$  in residues for which the two MD simulations yield different averaged  ${}^3J_{\alpha\beta}$ -coupling constants (Tyr20, Val29, Thr69, Leu75, Thr89). Black lines: MD simulation using the force-field parameter set 43A1. Red lines: MD simulation using the force-field parameter set 45A3.

lysozyme, one using the standard GROMOS96 parameter set 43A1 and the other the newer set 45A3. Comparison of the simulated and measured NMR parameters has been undertaken to evaluate the performance of the newer force-field parameter set 45A3 for a protein. The convergence of properties as function of time is very similar between the two simulations. Quantities such as the backbone atom-positional differences with respect to the X-ray structure or the NMR model

structure, and  $r^{-3}$ -averaged NOE inter-proton distances converge within 1.5 ns, while backbone  ${}^3J_{\text{HN}\alpha}$ -coupling constants and  ${}^1\text{H}$ - ${}^{15}\text{N}$  order parameters take longer, up to a few nanoseconds. For a number of residues, side-chain  ${}^3J_{\alpha\beta}$ -coupling constants and side-chain  ${}^1\text{H}$ - ${}^{15}\text{N}$  order parameters have not reached full convergence in the time scale simulated, due to the dependence of these properties on relatively rare structural transitions, e.g. torsional angle transitions. The need for more

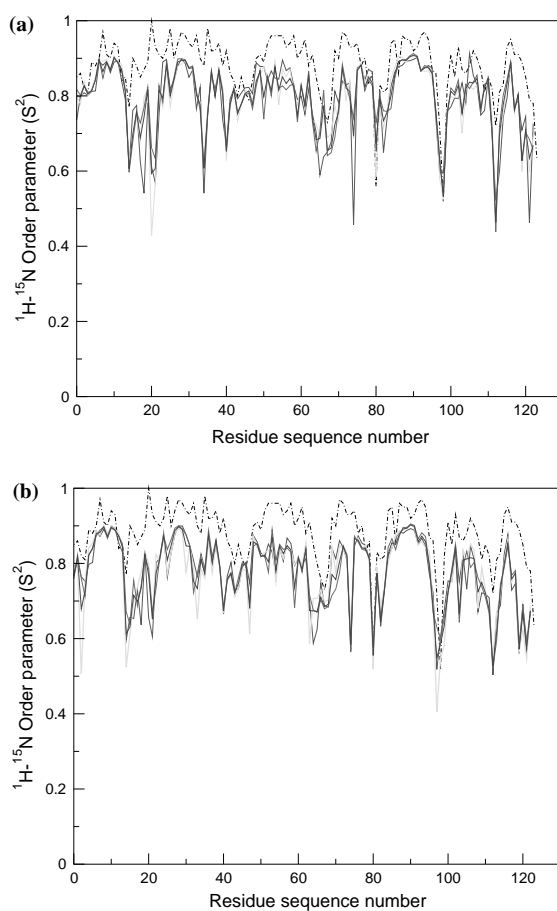


Figure 8. Backbone  $^1\text{H}$ - $^{15}\text{N}$  order parameter ( $S^2$ ) of HEWL calculated from the MD trajectories for the force-field parameter sets 43A1 and 45A3 and averaged (using a 200 ps window) over different time periods: 0.0–0.5 ns (red), 0.5–1.5 ns (yellow), 1.5–3.5 ns (green) and 0.0–3.5 ns (blue). Experimental values (dot-dashed line) were taken from (Buck et al., 1995). (a) Corresponds to the force-field parameter set 43A1. (b) Corresponds to the force-field parameter set 45A3.

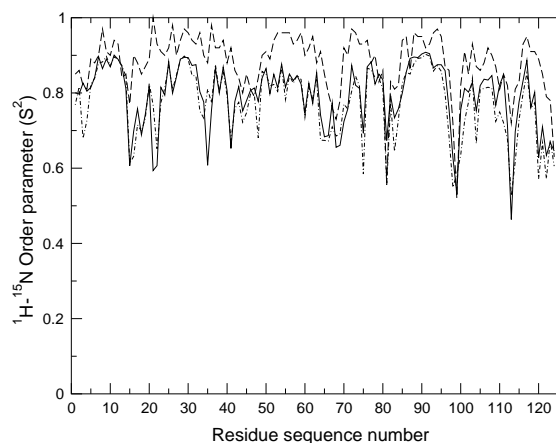


Figure 9. Backbone  $^1\text{H}$ - $^{15}\text{N}$  order parameter ( $S^2$ ) as function of residue sequence number. For the MD ensembles, averages were calculated using a 200 ps window over 3.5 ns of simulation using two force-field parameter sets: 43A1 (solid line) and 45A3 (dot-dashed line). Experimental data (dashed line) were from (Buck et al., 1995). No experimental values are available for Lys1, Ser50, Pro70, Pro79 and Ala110.

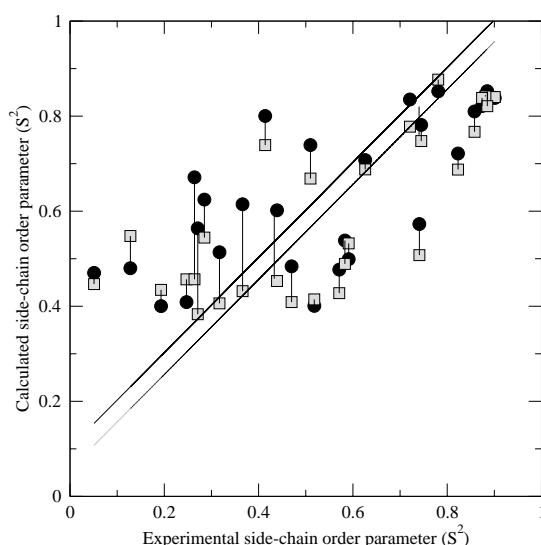


Figure 10. Comparison of 28 experimental  $^1\text{H}$ - $^{15}\text{N}$  order parameters ( $S^2$ ) of side-chain NH groups with values averaged over the MD trajectories (200 ps window over 3.5 ns of simulation). Solid circles: MD simulation using the force-field parameter set 43A1. Open squares: MD simulation using the force-field parameter set 45A3. Corresponding calculated  $S^2$  values in the two sets of data are connected by a line. The two diagonal lines represent a least-squares fit of the simulated against the experimental data (solid: 45A3, dashed 43A1).

extensive sampling of side-chain motion is underlined by the observation that the MD simulations slightly overestimate the side-chain order parameters compared to experiment. Generally, the various quantities are reproduced comparably well by the two sets. Most of the NOE distance violations identified in both MD trajectories involve residues located in loops or flexible regions of the protein.

Traditionally, the accuracy of MD simulations has been restricted by three factors: (i) the quality of the force field used, (ii) the extent of sampling of the conformational space and (iii) the sensitivity of the different quantities to the precise shape of the conformational distributions. Current advances in force-field parametrization and increased computational resources have allowed for more realistic MD simulations of a mounting variety of biomolecular systems. The validation (van Gunsteren and Mark, 1998) of improved force-field parameter sets through MD simulations of proteins is an essential step towards the understanding of biomolecular behaviour. The older GROMOS parameter set 43A1 did perform well for proteins, but not for lipids, which were generally simulated using a special set of parameters tailored for membranes (Egberts et al., 1994). The new parameter set

45A3 was calibrated including long (up till 18 carbon atoms) aliphatic chains (Schuler et al., 2001) and should, therefore, be suitable not only for proteins, but for lipid aggregates too. Here we have shown that this parameter set 45A3 has not lost significant accuracy in simulations of proteins. It seems suitable to simulate both proteins and lipid aggregates (Chandrasekhar and van Gunsteren, 2001, 2002; Chandrasekhar et al., 2003) thereby sustaining the aim of developing a single force field for a variety of biomolecular systems.

As was done here, the possible improvement of a force field or other technical aspects of MD simulation can be tested by comparison of simulation trajectories obtained with older and newer force fields or techniques against one set of experimental data. Inversely, one could test the improvement of experimental techniques by comparison of older and newer experimental data on the same system with one simulation trajectory. Comparing two different sets of experimental NOE data on hen egg white lysozyme (Smith et al., 1993; Schwalbe et al., 2001) with each of our two simulation trajectories we found comparable agreement for the newer set of 1630 NOE bounds (set2) as for the older set of 1079 bounds

(set1), despite the larger size of the newer set. This shows that over time experimental data may converge towards the theoretical ones, and should serve as a cautionary note when drawing conclusions about the (insufficient) quality of simulation results from observed discrepancies between simulated and measured data.

### Acknowledgements

Financial support was obtained through the National Center of Competence in Research (NCCR) in Structural Biology of the Swiss National Science Foundation, which is gratefully acknowledged.

### References

- Antes, I., Thiel, W. and van Gunsteren, W.F. (2002) *Eur. Biophys. J.*, **31**, 504–520.
- Bakowies, D. and van Gunsteren, W.F. (2002) *J. Mol. Biol.*, **315**, 713–736.
- Berendsen, H.J.C., Postma, J.P.M., Van Gunsteren, W.F., Dinola, A. and Haak, J.R. (1984) *J. Chem. Phys.*, **81**, 3684–3690.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F. and Hermans, J. (1981) In *Intermolecular Forces*, Pullman, B. (Ed.), Reidel, Dordrecht, pp. 331–342.
- Bonvin, A.M., Sunnerhagen, M., Otting, G. and van Gunsteren, W.F. (1998) *J. Mol. Biol.*, **282**, 859–873.
- Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comput. Chem.*, **4**, 187–217.
- Buck, M., Boyd, D.B., Redfield, C., MacKenzie, D.A., Jeenes, D.J., Archer, D.B. and Dobson, C.M. (1995) *Biochemistry*, **34**, 4041–4055.
- Carter, D., He, J., Ruble, J.R. and Wright, B. (1997) *Protein Data Bank, entry 1AKI*.
- Chandrasekhar, I. and van Gunsteren, W.F. (2001) *Curr. Sci.*, **81**, 1325–1327.
- Chandrasekhar, I. and van Gunsteren, W.F. (2002) *Eur. Biophys. J.*, **31**, 89–101.
- Chandrasekhar, I., Kastenholz, M.A., Lins, R.D., Oostenbrink, C., Schuler, L.D., Tieleman, D.P. and van Gunsteren, W.F. (2003) *Eur. Biophys. J.*, **32**, 67–77.
- Czechitzky, W., Daura, X., Vasella, A. and van Gunsteren, W.F. (2001) *Helv. Chim. Acta*, **84**, 2132–2145.
- Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunsteren, W.F. and Mark, A.E. (1999) *Angew. Chem. Ind. Ed.*, **38**, 236–240.
- Daura, X., Glättli, A., Gee, P., Peter, C. and van Gunsteren, W.F. (2002) *Adv. Protein Chem.*, **62**, 341–360.
- Daura, X., Mark, A.E. and van Gunsteren, W.F. (1998) *J. Comput. Chem.*, **19**, 535–547.
- Daura, X., van Gunsteren, W.F., Rigo, D., Jaun, B. and Seebach, D. (1997) *Chem.-Eur. J.*, **3**, 1410–1417.
- DeMarco, A., Llinas, M. and Wüthrich, K. (1978) *Biopolymers*, **17**, 617–636.
- Egberts, E., Marrink, S.-J. and Berendsen, H.J.C. (1994) *Eur. Biophys. J.*, **22**, 423–436.
- Glättli, A., Daura, X. and van Gunsteren, W.F. (2002) *J. Chem. Phys.*, **116**, 9811–9828.
- Hermans, J., Berendsen, H.J.C., van Gunsteren, W.F. and Postma, J.P.M. (1984) *Biopolymers*, **23**, 1513–1518.
- Hünenberger, P.H. and van Gunsteren, W.F. (1997) In *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*, Vol. 3, van Gunsteren, W.F., Weiner, P.K. and Wilkinson, A.J. (Eds.), Kluwer Academic Publishers, Dordrecht, pp. 3–82.
- Hünenberger, P.H., Mark, A.E. and van Gunsteren, W.F. (1995) *J. Mol. Biol.*, **252**, 492–502.
- Jorgensen, W.L. and Tirado-Rives, J. (1988) *J. Am. Chem. Soc.*, **110**, 1657–1666.
- Kabsch, W. and Sander, C. (1983) *Biopolymers*, **22**, 2577–2637.
- Karplus, M. (1959) *J. Chem. Phys.*, **30**, 11–15.
- Karplus, M. and McCammon, J.A. (2002) *Nat. Struct. Biol.*, **9**, 646–652.
- Levitt, M. (1983) *J. Mol. Biol.*, **168**, 595–620.
- Levitt, M., Hirshberg, M., Sharon, R. and Daggett, V. (1995) *Comput. Phys. Commun.*, **91**, 215–231.
- MacKerell Jr., A.D., Wiorkiewicz-Kuczera, J. and Karplus, M. (1995) *J. Am. Chem. Soc.*, **117**, 11946–11975.
- Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E. and Wüthrich, K. (1998) *J. Biomol. NMR*, **12**, 1–23.
- Momany, F.A. and Rone, R. (1992) *J. Comput. Chem.*, **13**, 888–900.
- Nemethy, G., Gibson, K.D., Palmer, K.A., Yoon, C.N., Paterlini, G., Zagari, A., Rumsey, S. and Scheraga, H. A. (1992) *J. Phys. Chem.*, **96**, 6472–6484.
- Oostenbrink, B.C., Pitera, J.W., Van Lipzig, M.M.H., Mermerman, J.H.N. and van Gunsteren, W.F. (2000) *J. Med. Chem.*, **43**, 4594–4605.
- Pardi, A., Billeter, M. and Wüthrich, K. (1984) *J. Mol. Biol.*, **180**, 741–751.
- Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham III, T.E., DeBolt, S., Ferguson, D., Seibel, G. and Kollman, P.A. (1995) *Comput. Phys. Commun.*, **91**, 1–41.
- Ryckaert, J.-P., Ciccotti, G. and Berendsen, H.J.C. (1977) *J. Comput. Phys.*, **23**, 327–341.
- Schuler, L.D. and van Gunsteren, W.F. (2000) *Mol. Sim.*, **25**, 301–319.
- Schuler, L.D., Daura, X. and van Gunsteren, W.F. (2001) *J. Comput. Chem.*, **22**, 1205–1218.
- Schwalbe, H., Grimshaw, S.B., Spencer, A., Buck, M., Boyd, J., Dobson, C.M., Redfield, C. and Smith, L. (2001) *Protein Sci.*, **10**, 677–688.
- Scott, W.R.P., Hünenberger, P.H., Tironi, I.G., Mark, A.E., Billeter, S.R., Fennen, J., Torda, A.E., Huber, P., Krüger, P. and van Gunsteren, W.F. (1999) *J. Phys. Chem. A*, **103**, 3596–3607.
- Smith, L.J., Dobson, C.M. and van Gunsteren, W.F. (1996) *J. Mol. Biol.*, **286**, 1567–1580.
- Smith, L.J., Dobson, C.M. and van Gunsteren, W.F. (1999) *Proteins*, **36**, 77–86.
- Smith, L.J., Mark, A.E., Dobson, C.M. and van Gunsteren, W.F. (1995) *Biochemistry*, **34**, 10918–10931.
- Smith, L.J., Sutcliffe, M.J., Redfield, C. and Dobson, C.M. (1991) *Biochemistry*, **30**, 986–996.

- Smith, L.J., Sutcliffe, M.J., Redfield, C. and Dobson, C.M. (1993) *J. Mol. Biol.*, **229**, 930–944.
- Stocker, U. and van Gunsteren, W.F. (2000) *Proteins*, **40**, 145–153.
- Stocker, U., Spiegel, K. and van Gunsteren, W.F. (2000) *J. Biomol. NMR*, **18**, 1–12.
- Tropp, J. (1980) *J. Chem. Phys.*, **72**, 6035–6043.
- van Gunsteren, W.F. and Berendsen, H.J.C. (1987) *Groningen Molecular Simulation (GROMOS) Library Manual*, Biomos, Groningen.
- van Gunsteren, W.F. and Berendsen, H.J.C. (1990) *Angew. Chem. Int. Ed.*, **29**, 992–1023.
- van Gunsteren, W.F. and Mark, A.E. (1998) *J. Chem. Phys.*, **108**, 6109–6116.
- van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott, W.R.P. and Tironi, I.G. (1996) *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Vdf Hochschulverlag AG an der ETH Zürich, Zürich.
- van Gunsteren, W.F., Bonvin, A.M.J.J., Daura, X. and Smith, L.J. (1999) In *Structure, Computation and Dynamics in Protein NMR. Biol. Magnetic Resonance*, Vol. 17, Krishna, K.N. and Berliner, L.J. (Eds.), Plenum Publishers, New York, pp. 3–35.
- van Gunsteren, W.F., Bürgi, R., Peter, C. and Daura, X. (2001) *Angew. Chem. Int. Ed.*, **40**, 351–355.
- van Gunsteren, W.F., Daura, X. and Mark, A.E. (1998) In *Encyclopedia of Computational Chemistry*, Vol. 2, von Ragué Schleyer, P. (Ed.), John Wiley & Sons, New York, pp. 1211–1216.
- Weiner, P.K. and Kollman, P.A. (1981) *J. Comput. Chem.*, **2**, 287–303.
- Wüthrich, K., Billeter, M. and Braun, W. (1983) *J. Mol. Biol.*, **169**, 949–961.