How Sensitive Are Nanosecond Molecular Dynamics Simulations of Proteins to Changes in the Force Field?

Alessandra Villa,[†,‡] Hao Fan,[†,§] Tsjerk Wassenaar,[†,⊥] and Alan E. Mark*[,†,||]

Laboratory of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands, and School of Molecular and Microbial Sciences, and the Institute for Molecular Biosciences, University of Queensland, St Lucia, QLD 4072, Australia

Received: December 14, 2006; In Final Form: March 12, 2007

The sensitivity of molecular dynamics simulations to variations in the force field has been examined in relation to a set of 36 structures corresponding to 31 proteins simulated by using different versions of the GROMOS force field. The three parameter sets used (43a1, 53a5, and 53a6) differ significantly in regard to the nonbonded parameters for polar functional groups and their ability to reproduce the correct solvation and partitioning behavior of small molecular analogues of the amino acid side chains. Despite the differences in the force field parameters no major differences could be detected in a wide range of structural properties such as the root-mean-square deviation from the experimental structure, radii of gyration, solvent accessible surface, secondary structure, or hydrogen bond propensities on a 5 to 10 ns time scale. The small differences that were observed correlated primarily with the presence of charged residues as opposed to residues that differed most between the parameter sets. The work highlights the variation that can be observed in nanosecond simulations of protein systems and implications of this for force field validation, as well as for the analysis of protein simulations in general.

Introduction

Molecular dynamics (MD) simulation techniques are increasingly used to analyze and predict the structural and dynamic properties of biomolecular systems at an atomic level. The size and complexity of biomolecular systems, such as proteins, DNA, and RNA, together with the time scales that must be examined in general necessitates the use of classical Newtonian dynamics in conjunction with empirical force fields. In such force fields simple analytical functions are used to represent the potential energy surface of the system in terms of the Cartesian coordinates of the interacting atoms. The fidelity with which the force field represents the underlying potential energy surface is the primary determinant of the accuracy of a simulation. As such, there is much interest in both force field development and validation. The reason why both issues remain the focus of intense interest despite the fact that the first simulations of proteins were performed over 20 years ago is threefold: First, the parametrization of a biomolecular force field is an underdetermined problem. Many parameters are fitted based on a limited range of reference data. Even when using equivalent analytical functions to describe the potential energy surface, different sets of parameters can give similar fits to the available data. Second, the fidelity with which the intermolecular interactions can be represented by using what are relatively crude analytical functions is limited. By necessity higher weights must be given to specific aspects of the available reference data. Third, biomacromolecules are not parametrized directly. Instead, parametrization is based on the properties of small model compounds, with sets of parameters then transferred to other (larger) systems. For a general overview of these and other considerations when developing force fields for simulating biomolecular systems we refer the reader to separate reviews by Hünenberger, MacKerell, and Jorgensen.[1−3]

Currently, a range of force fields for the simulation of biomolecular systems are in common use. These include AMBER,[4,5] CHARMM,[6,7] GROMOS,[8,9] and OPLS.[10] All of these force fields use very similar functional forms to describe specific interactions, but differ significantly both in terms of the values of specific parameters and in the philosophy by which these parameters are obtained. In particular, the force fields vary in respect to the partial charges used to represent the electrostatic interactions between molecules and the Lennard-Jones terms used to represent the van der Waals interactions between atoms. This variation raises the challenging question of validation.

The validation of molecular force fields generally involves comparing the results of simulations to a range of experimental data, preferably data not used in the original parametrization. Validation requires that the properties derived from the simulations have converged, that a direct comparison with experiment is made, and that a consensus over a range of different properties (structural, energetic, and dynamical properties) as well as for a range of different types of molecules and environments can be obtained.[11] This is, however, rarely if ever possible. Instead, validation studies generally focus on a single or small range of properties for a given system.[12−19] Recent years have, however, seen a growing number of broader studies that have attempted

* Address correspondence to this author at the University of Queensland. E-mail: a.e.mark@uq.edu.au. Phone: +61-7-3365 4180. Fax: +61-7-3365 3872.

[†] University of Groningen.

[‡] Current address: Institute for Physical and Theoretical Chemistry, J.W. Goethe University, Max-von-Laue Strasse 7, 60439 Frankfurt am Main, Germany.

[§] Current address: Department of Biopharmaceutical Sciences, University of California, San Francisco, Mission Bay, Byers Hall, 1700 4th Street, Suite 501, San Francisco, California 94158-2330.

[⊥] Current address: Department of NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.

[||] University of Queensland.

10.1021/jp068580v CCC: $37.00
Published on Web 05/10/2007

to evaluate the effectiveness of different biomolecular force fields by comparing simulations of protein and peptide systems by using different force fields or different parameter sets of a given force field. For example, Price and Brooks performed a comparative analysis of the structural and dynamical properties of three proteins derived from 2 ns MD simulations using three different force fields: AMBER94, CHARMM22, and OPLS-AA.[20] Despite the very short time scale over which these tests were performed the authors concluded that there were no consistent differences in the behavior of proteins simulated in the different force fields and thus that the three force fields were equally good. Van der Spoel and Lindahl in contrast compared significantly longer (50 ns) simulations of a single system, the Villin head piece, using the GROMOS96 and OPLS-AA/L force fields in combination with different water models.[21] They also compared the effects of different protonation states and the effect of using virtual particles to represent the interactions to H atoms. Overall, the available NMR data were best reproduced when using the GROMOS96 force field that also gave rise to larger and more frequent conformational changes with respect to average NMR structure on the time scale sampled. Sorin and Pande compared an ensemble of folding trajectories for two helical peptides using several variants of the AMBER force field and concluded that a variant of the AMBER99 potential with modified backbone torsional potentials was superior to others in reproducing experimental thermodynamics and kinetics for helix−coil transition.[22] Patel and co-workers attempted to assess the validity of a fluctuating charge (FQ) force field by monitoring the stability of six proteins in a series of nanosecond simulations.[23] They found that the results obtained using the FQ force field were comparable to results obtained using a nonpolarizable model. In a similar manner Soares and co-workers performed nanosecond simulations of hen egg white lysozyme using two GROMOS96 parameter sets (43a1 and 45a3) and concluded that both sets reproduced NMR spectroscopic data equally well.[24] Finally, Oostenbrink and co-workers have compared simulations of hen egg white lysozyme, a DNA dodecamer, and a β-peptide performed with different versions of the GROMOS96 force field.[25] They concluded that in terms of protein stability, the most recent 53a6 parameter set behaved similarly to the previous 45a3 set[26] and reproduced experimental data equally well.

In each of the above studies the authors concluded either that a specific force field performed significantly better than the alternatives tested or that the force fields were equivalent to earlier versions and thus valid. However, these conclusions were based on very limited samples in terms of the range of systems investigated or the time scales investigated, or both. Thus, while all such comparisons must, to some degree, reflect the nature of the underlying parameters, it is difficult to determine whether the conclusions from any of these studies are objective or statistically meaningful. In particular, no attempt to evaluate the sampling errors or even the expected spread of the distribution in the properties investigated was presented in any of these studies.[27]

In the current study, we attempt to assess whether it is possible to detect the effects of changes in the force field parameters in simulations on a 5 to 10 ns time scale by comparing the outcomes of a large number of independent simulations using statistically robust methods.[27] Changes in the force field were represented by three different versions of the GROMOS96 force field: the 43a1,[8] the 53a5, and the 53a6 parameter sets.[9] Each of these parameter sets use the same form of the potential energy function and have essentially identical bonded parameters. They differ significantly, however, in terms of the partial charges and Lennard-Jones parameters used to describe the nonbonded interactions, and in their ability to reproduce critical molecular properties. The 43a1 parameter set was based on a combination of fitting to crystallographic data, atomic polarizabilities,[28] and the properties of a range of liquid alkanes.[29] This parameter set fails, however, to reproduce the liquid properties of certain polar organic compounds and fails to reproduce the solvation properties of a range of small molecular analogues of amino acid side chains.[12] The 53a5 parameter set was optimized to reproduce the liquid properties (density and heat of vaporization) of a series of small organic molecules. The 53a6 parameter set, on the other hand, was specifically parametrized to reproduce thermodynamic properties such as free enthalpies of solvation. In particular, the three parameter sets differ in their ability to reproduce the partition properties of amino acids. Only the 53a6 set reproduces correctly the free enthalpies of solvation both in water and in cyclohexane for analogues of the amino acids.[9,12]

The results presented in this paper are based on a comparison of 36 structures[30−65] corresponding to 31 different proteins ranging in size from 50 to 100 amino acid residues simulated in explicit water. Two independent simulations (5 and 10 ns in length) were performed for each structure in combination with each of the three parameter sets. The simulations were characterized by a number of instantaneous properties. These properties include the deviation from the experimentally determined structure, the radius of gyration, the stability of backbone hydrogen bonds, and the hydrophobic and hydrophilic solvent accessible surface areas. The results were then combined and subjected to a detailed statistical analysis.

The aims of the study were (1) to determine if the different parameter sets proposed by the authors of the force fields give rise to systematic and/or statistically significant differences in the outcomes of the simulations, (2) to determine if the differences observed (if any) could be readily correlated with the differences within the different parameter sets, and (3) to determine the degree of variability that might be expected between replicate simulations of the same protein or between similar proteins simulated under identical conditions and with the same force field. It should be stressed that the study was not designed to determine which version of the force field was most appropriate for simulating a given system nor has any attempt been made to use these simulations to refine the available force fields.

Methods

Protein Data Set. The 36 structures[30−65] (corresponding to 31 different proteins) used in this study were taken from the Protein Data Bank.[66] Some general characteristics of each of these structures are summarized in Table 1. Of the 36 structures, 17 were determined by X-ray crystallography, 19 by NMR spectroscopy. This set of structures is the same as that used previously to compare the relative stability of X-ray crystallographic and NMR derived structures in molecular dynamics (MD) simulations.[67] The proteins represent different secondary structure types. All range from 50 to 100 amino acid residues and are believed to be monomeric in solution. None contain cysteine−cysteine bridges.

MD Simulations. All simulations were performed with the Groningen Machine for Chemical Simulation (GROMACS) package.[68,69] The parameter sets 43a1,[8] 53a5, and 53a6[9] of the GROMOS96 force field were used to describe the proteins. The charges of ionizable groups were set appropriate for pH 7.0

TABLE 1: A Summary of the Structural Characteristics of the Proteins Used in This Study

ID	PDBID[a]	description	SSE[b]	N_{res}	N_{cha}[c]
1	1vif	dihydrofolate reductase	α/β	60	0
2	1tuc	the α-spectrin SRC homology 3 domain, circular permutant	α/β	61	0
3	1vcc	the N-terminal fragment of vaccinia virus DNA topoisomerase	α/β	77	−1
4	1ail	the unique RNA-binding domain of the influenza virus NS1 protein	α	70	2
5	1cei	colicin E7 immunity protein	α/β	85	−9
6	1ctf	the C-terminal domain of the ribosomal protein L7/L12	α/β	68	−2
7	1pgx	the B2 immunoglobulin-binding domain of streptococcal protein G	α/β	70	−4
8	1tif	the N-terminal domain of translational initiation factor IF3	α/β	76	5
9	2acy	acyl-phosphatase	α/β	98	1
10	2fxb	ferredoxin	α/β	81	−17
11	1r69	the N-terminal domain of phage 434 repressor	α	63	4
12	1bm8	the DNA-binding domain of Mbp1	α/β	99	6
13	2ci2	serine proteinase inhibitor CI-2	α/β	63	−1
14	1pgb	the B1 immunoglobulin-binding domain of streptococcal protein G	α/β	56	−4
15	1shg	a Src-homology 3 (SH3) domain	α/β	57	1
16	1ubi	ubiquitin	α/β	76	0
17	1a19	barstar	α/β	89	−6
18	1aoy	the DNA-binding domain of arginine repressor	α/β	78	3
19	1stu	the dsRNA-binding domain of Drosophila staufen protein	α/β	68	5
20	1sro	the S1 RNA-binding domain of polyribonucleotide phosphorylase, PNPase	α/β	76	1
21	1sap	the DNA-binding protein Sac7d	α/β	66	6
22	1afi	mercuric ion binding protein, MerP	α/β	72	3
23	1bb8	the DNA-binding domain of tn916 integrase	α/β	71	5
24	2bby	the DNA-binding domain of rap30	α/β	69	3
25	2fmr	the KH1 domain of Fragile X protein	α/β	65	−4
26	1a1z	the FADD (Mort1) death-effector domain	α	83	−3
27	1bw6	the DNA-binding domain of centromere binding protein B	α	56	6
28	1coo	the C-terminal domain of RNA polymerase alpha subunit	α	81	−3
29	1lea	the LexA repressor DNA binding domain	α/β	72	2
30	2af8	the actinorhodin polyketide synthase acyl carrier protein	α	86	−12
31	2ezh	the I γ subdomain of the Mu end DNA-binding domain of phage Mu transposase	α	65	−2
32	3ci2	serine proteinase inhibitor CI-2	α/β	63	0
33	2gb1	the B1 immunoglobulin-binding domain of streptococcal protein G	α/β	56	−4
34	1aey	a Src-homology 3 (SH3) domain	α/β	57	1
35	1d3z	ubiquitin	α/β	76	0
36	1bta	barstar	α/β	89	−6

[a] PDBID denotes the PDB entry name. [b] SSE refers to the classification of secondary structure content according to the definitions given in the corresponding PDB files. In the cases of 2ci2 and 3ci2, 1pgb and 2gb1, 1shg and 1aey, 1ubi and 1d3z, and 1a19 and 1bta, which correspond to pairs, structures were solved by using X-ray and NMR, respectively. The overlap of the definitions of secondary structures given in the PDB files was used. For 2gb1, the definitions given in the file 1pgb were used as no definition of secondary structure was given in the PDB file of 2gb1. For the same reason the definition for 1aey was used for 1shg. For the structures solved by using NMR the PDB files for 1sap, 1a1z, 1bw6, 1coo, 1lea, 2af8, 2ezh, 2gb1, and 1bta contain only a single averaged structure. Where multiple structures had been deposited in the PDB, the first structure in each set was chosen to represent the molecule. Two C-terminal residues of 2ci2 were removed for compatibility with 3ci2. The C-terminal residue of 1aey was removed for compatibility with 1shg. [c] N_{cha} is the expected net charge of each structure in water at neutral pH.

assuming standard pK_a values. Histidines were protonated on the δ1 nitrogen atom. No counterions were added to neutralize the systems.

Each structure was placed in a periodic truncated octahedral box filled with Simple Point Charge (SPC) water molecules.[70] The minimum distance between the solute and the wall of the unit cell was 1.0 nm. During the simulations, bond lengths within the proteins were constrained by using the LINCS algorithm.[71] The SETTLE algorithm was used to constrain the geometry of water molecules.[72] A time step of 2 fs was used to integrate the equations of motion. Nonbonded interactions were evaluated with the use of a twin range cutoff. Interactions within the shorter range cutoff (0.9 nm) were calculated every step whereas interactions within the longer cutoff (1.4 nm) were updated every 10 steps, together with the pair list. A reaction field correction was applied to the electrostatic interactions beyond 1.4 nm,[73] using a water dielectric constant of 78. To maintain constant temperature the protein and solvent were independently coupled to an external heat bath (300 K) with a Berendsen thermostat.[74] The pressure was weakly coupled to an isotropic pressure bath (1 bar).[74] The isothermal compressibility was 4.5×10^{-5} bar^{-1}. The protein−water system was first minimized by using a steepest descent method and then equilibrated by a 10 ps MD simulation with positional restraints on the protein atoms. Two independent simulations, 5 and 10 ns in length, were performed for each structure in combination with each of the three sets of parameters giving 216 simulations in total. Note the parameters used for the simulations (cutoffs, reaction field, etc.) were chosen so as to be consistent with the original parametrization of the force fields.

Analysis of System Properties. A range of instantaneous properties of the system were analyzed with the use of each simulation. The value of the property in question was averaged over the period from 4 to 5 ns in both simulations. For the 10 ns simulations, the properties were also averaged over the period 4 to 10 ns. The properties included (1) the root-mean-square deviation (rmsd) from the corresponding experimentally determined structure, (2) the radius of gyration (RG), (3) the number of backbone hydrogen bonds within secondary structure elements defined in the experimental structure (HB), (4) the hydrophobic solvent accessible surface (HSA), and (5) the hydrophilic solvent accessible surface (PSA). The average structure during a certain period of a given simulation was used for rmsd calculation while the other properties were calculated as the mean of the instantaneous observations during the same period. All properties were analyzed by using routines within the GROMACS package.[69] The rmsd was determined for backbone atoms in the whole structure as well as in elements

of secondary structure (α-helices or/and β-sheet) defined in the experimental structure after fitting to the reference structure. The solvent-accessible surface was computed numerically by using atomic radii of 0.16 nm for carbon, 0.13 nm for oxygen, 0.14 nm for nitrogen, 0.20 nm for sulfur, and 0.10 nm for hydrogen atoms.[75] The atomic radius of water was 0.14 nm. An atom was recognized as hydrophobic if the absolute value of its partial charge was less than 0.2 e. A hydrogen bond was considered to exist when the distance between the hydrogen acceptor was less than 0.25 nm and the donor−hydrogen−acceptor angle was greater than 60°. Only H-bonds between backbone atoms involved in elements of secondary structures were considered.

For the statistical analysis of the data the results were first normalized to eliminate the contribution from the different proteins. In this way, a modified linear model could be applied, excluding the protein term. This allowed a direct comparison of the effects of the force field and of the effects due to specific interactions between the protein and the force field as opposed to the variation due to the proteins themselves. An assessment was then made of the similarity or otherwise of results obtained with the three different parameter sets.

Statistical Analysis. A multivariate statistical analysis[27] was performed to determine if the differences between the simulations performed with the different versions of the force field exceeded the level of natural variation in these simulations. To perform the statistical analysis the set of structural properties were first expressed as observation vectors for all simulations. The general linear model that was assumed to describe the original observations is given by

$$x_{ijkl} = \mu_{il} + \tau_{jl} + \gamma_{ijl} + \epsilon_{ijkl}$$

Here x_{ijkl} is the kth observation of property l for protein i, under condition (parameter set) j. μ_{il} is a general location parameter for property l specific to protein i and is comparable to the protein mean for that property. τ_{jl} is a deviation from μ_{il} due to the force field used, and γ_{ijl} is a deviation specific to the combination of the force field used and the protein. This is called an interaction term. ϵ_{ijkl} is a random error, which is assumed to be independent of the force field. The error has a mean of zero and a variance specific to the protein.

The primary objective of the statistical analysis was to test for differences between simulations that could be attributed to differences between the force fields used, not differences between proteins within a given force field. The incorporation of protein-specific error terms means that it is not possible to make a direct comparison between the sets of simulations based on a simple analysis of variance. This is because the analysis of variance (ANOVA) requires that the residual variances be equal (homoscedastic criterium). For this reason, the data were transformed (normalized) such that the variation due to the different proteins in the test set was removed from the model. The normalization was performed by scaling the results according to

$$x'_{ijkl} = \frac{(x_{ijkl} - a_{il})}{b_{il}}$$

$$a_{il} = \bar{x}_{il} = \frac{1}{cr}\sum_{j=0}^{c}\sum_{k=0}^{r} x_{ijkl} \quad \text{and}$$

$$b_{il} = \frac{1}{\sqrt{\mathrm{var}(\epsilon_{il})}} = \left[\frac{1}{cr-1}\sum_{j=0}^{c}\sum_{k=0}^{r}(x_{ijkl} - \bar{x}_{ijl})^2\right]^{-1/2}$$

where c is the number of force fields (3) and r denotes the number of simulations per protein−force field combination (2). The linear model describing the transformed results is then given by

$$x'_{ijkl} = \tau'_{jl} + \gamma'_{ijkl} + \epsilon'_{ijkl}$$

This model was then used as the basis for a fixed-effects multivariate analysis of variance (MANOVA) with provision for force field effects and interactions between the protein and the force field. The purpose of MANOVA is to test whether the effects due to conditions or interactions between conditions exceed the residual variance on a given significance level. In other words, we determined the probability that the different samples were obtained from a common underlying distribution. For this purpose the total variance was decomposed into a set of components determined by the external conditions and an unexplained residual variance. The ratio of the condition-determined and residual variance in a univariate case has the variance ratio (F) distribution with the appropriate degrees of freedom. This ratio provides the probability that the condition-determined and residual variances are equal and that the given condition consequently has no effect.

For the multivariate case several test statistics are available which are based on the distribution of the nonzero eigenvectors of the ratio of covariance matrices equivalent to the ratio of variances in the univariate analysis of variance (ANOVA). In this study Wilk's λ was used.[76] If the MANOVA results suggested a significant difference between at least two sets of simulations, multiple comparisons were made to investigate the source(s) of these differences. This post-hoc analysis was done by using the Roy union−intersection approach.[77,78] Finally, for each pair of parameter sets showing statistically significant differences in the simulation outcomes, regression analysis was performed to understand the difference in terms of specific properties of the protein.

All statistical analysis was performed with the program R (R Development Core Team, 2005), a language and environment for statistical computing (http://www.r-project.org/).

Results

Root-Mean-Square Deviation to Experimental Structures. The positional root-mean-square deviation (rmsd) of backbone atoms was calculated for the mean structure (averaged over the period from 4 to 5 ns in each simulation) with respect to the corresponding experimental structure (see the Supporting Information: Table S1). Since the fluctuations in the rmsd can be dominated by the motions of flexible regions in proteins such as loops and termini, the backbone rmsd of only those residues involved in secondary structural elements (SSERMSD) is also reported (see the Supporting Information: Table S2). With the 43a1 parameter set (Table S1) the lowest rmsd value was 0.06 nm. This was obtained for the X-ray structure of α/β-fold protein G B1 domain (1pgb). The highest value was 0.75 nm for the NMR structure of α-fold protein actinorhodin polyketide synthase acyl carrier (2af8). With the 53a5 and 53a6 parameter sets, the NMR and X-ray structures, respectively, of the sh3

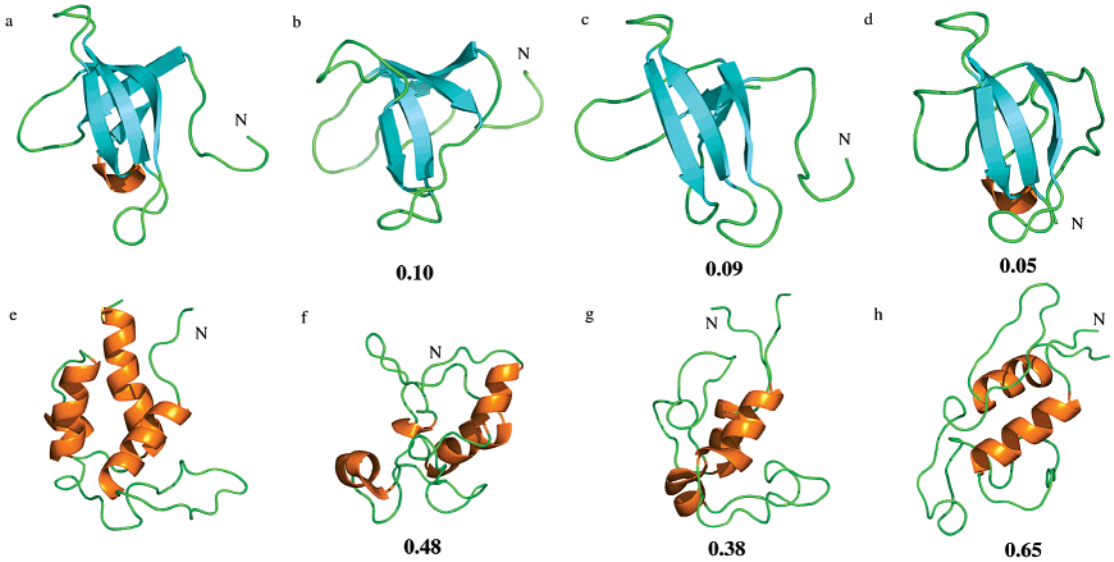

Figure 1. Top row: From left to right are shown the experimental structure and the final structures after 10 ns of MD simulations performed with the 43a1, 53a5, and 53a6 GROMOS force field for the protein 1vif. Bottom row: The corresponding structures for the protein 2af8. Where appropriate the root-mean-square (rms) positional deviation (in nm) of the backbone atoms of residues in elements of secondary structure with respect to the starting experimental structure is given below each figure.

domain of α-spectrin (1aey and 1shg, α/*β*-fold) gave the lowest rmsd values (0.10 and 0.06 nm) while the highest rmsd values, 0.91 and 1.01 nm, respectively, were obtained using the NMR structure of a DNA-binding protein (1bb8, α/*β*-fold). Compared to the rmsd values, the SSERMSD values fall within a much smaller range (0.05 to 0.53, 0.79, and 0.75 nm for 43a1, 53a5, and 53a6 parameter sets, respectively). However, using all of the parameter sets there were marked differences between the different proteins with respect to their deviations from the corresponding experimental structures during the simulations. In some cases this is likely to be due to problems in the experimental structure. As illustrated in Figure 1, the X-ray structure of dihydrofolate reductase (1vif) deviates much less (SSERMSD < 0.10 nm) from the experimental structure under all three parameter sets than does the NMR structure 2af8 (SSERMSD > 0.30 nm). Compared with the difference between proteins, the differences between the three different sets of parameters are relatively small. The distribution of the rmsd and SSERMSD values is shown in Figure 2. With the 43a1 parameter set, the SSERMSD values were <0.2 nm in 50 of the 72 simulations (69%). In 18 cases (25%) the values of the SSERMSD were in the range from 0.20 to 0.40 nm with only 4 (6%) above 0.40 nm. For 53a5, 43 simulations (60%) gave SSERMSD values <0.20 nm, 24 (33%) were in the range from 0.20 to 0.40 nm, and 5 (7%) were >0.40 nm. The corresponding values for 53a6 were 44 (61%), 19 (26%), and 9 (13%) (Figure 2). Averaged over the 72 simulations performed using each parameter set, the mean value for the rmsd was 0.28, 0.31, and 0.32 nm for 43a1, 53a5, and 53a6, respectively. The tendency for a slightly lower value to be obtained by using the 43a1 parameter set compared with the 53a5 and 53a6 parameter sets was also observed for the average values of the SSERMSD which are 0.17, 0.20, and 0.22 nm for the three sets, respectively. The extent to which two otherwise identical simulations might deviate on a 5 ns time scale due to differences in the starting velocities was also examined. Comparing the rmsd between the two structures averaged from 4 to 5 ns from the replicate simulations it was found that the deviation between the two simulations was similar to the deviation from the corresponding experimental structures. The average values of the rmsd (all the backbone atoms) between the replicate runs were

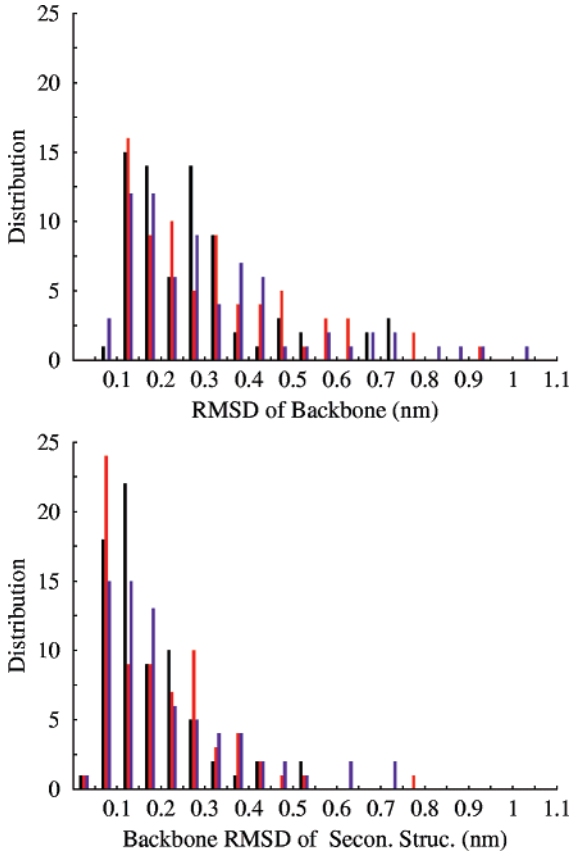

Figure 2. The distribution of values of the rms positional deviation (nm) with respect to the corresponding starting structures for all backbone atoms (top) and for backbone atoms involved in elements of secondary structure (bottom) obtained after 5 ns. The figure shows the combined results from the duplicate simulations of the 36 starting structures performed with the different GROMOS force field parameter sets: 43a1 (black), 53a5 (red), and 53a6 (blue).

0.25 (43a1), 0.31 (53a5), and 0.32 nm (53a6). Considering only secondary structure elements, the results were 0.15 nm for 43a1 and 0.20 nm for both 53a5 and 53a6. Although convergence has clearly not been achieved during these simulations the 43a1

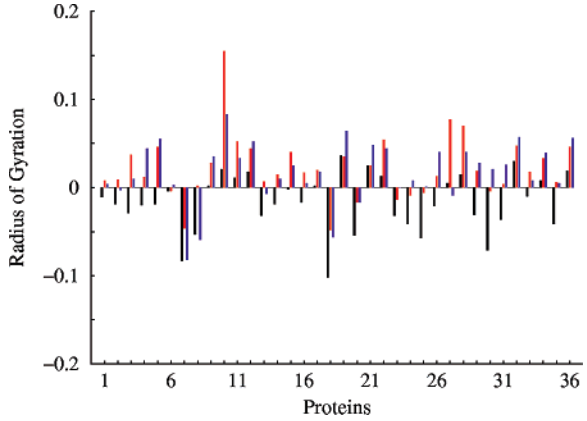

Figure 3. The relative deviation of the radii of gyration (RG) with respect to the value obtained from the corresponding experimental structure for each protein. The values are averaged over 5 ns for the duplicate simulations performed with each parameter set and plotted as a function of the protein identifier number (ID) (see Table 1): 43a1 (black), 53a5 (red), and 53a6 (blue).

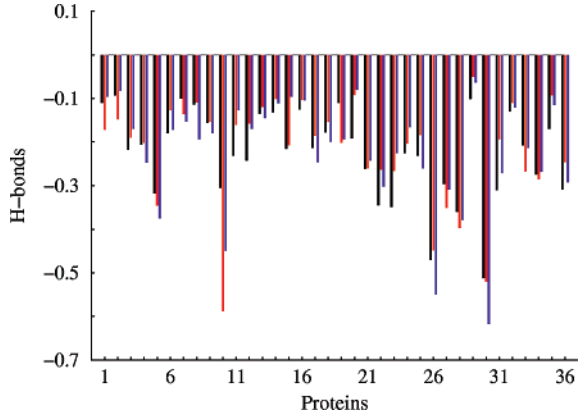

Figure 4. The relative deviation in the number of H-bonds involved in elements of secondary structure with respect to that found in the starting experimental structure. The values are an average over 5 ns for the duplicate simulations and plotted as a function of the protein ID (see Table 1): 43a1 (black), 53a5 (red), and 53a6 (blue).

parameter set appears to maintain the initial structures slightly better than other sets on a 5 ns time scale.

Radius of Gyration. The radius of gyration (RG) of protein structures was monitored during the simulations. The mean values of the RG averaged over the period from 4 to 5 ns were determined and are given as Supporting Information (Table S3). The average of the deviation from the replicate runs for each of the proteins in combination with each of the parameters sets from the initial structure are shown in Figure 3. With the 43a1 parameter set, the NMR structure of the DNA-binding domain of centromere binding protein B (1bw6, α-fold) showed the largest relative increase of RG (4.6%). The largest relative decrease in RG (−11.7%) was for the NMR structure of the DNA-binding domain of the arginine repressor (1aoy, α/β-fold). Among the 72 simulations, 24 (33%) showed an increase in RG (the proteins expanded) and 47 (65%) showed a decrease in RG (the proteins contracted). In the case of the parameter set 53a5, the maximal relative increase in RG was 19.8% for the X-ray structure of ferredoxin (2fxb, α/β-fold). Note, the experimental structures contained a number of ions. These were removed before the simulations were performed. This structure also had the highest overall net charge. Two structures, 1aoy and the X-ray structure of the B2 domain of protein G (1pgx) both of which have long flexible regions at the chain termini, showed the largest decrease in RG (−4.9%). In total, the RG increased in 53 simulations (74%) and decreased in 18 simulations (25%). The behavior of the 53a6 parameter set, in terms of the relative change in RG, is similar to that of 53a5. The largest increase was for the structure 2fxb (10.5%) while the largest decrease in RG occurred in the case of 1pgx (−11.2%). The number of simulations in which the RG increased or decreased were 55 (76%) and 16 (22%), respectively. On average, the relative deviations are −1.7%, 2.2%, and 1.8% for the 43a1, 53a5, and 53a6 parameter sets, respectively. The results suggest that when using the 43a1 parameter set the proteins tend to become slightly more compact while when using 53a5/6 the proteins tend to expand slightly.

Number of Native Hydrogen Bonds. Like the RG, the number of mainchain−mainchain hydrogen bonds within elements of secondary structure defined in the experimental structures (HB) was evaluated and averaged from 4 to 5 ns. The raw data are given as Supporting Information (Table S4). The average deviation from the experimental value for each protein is presented in Figure 4. The largest relative decrease in HBs was found for the NMR structure 2af8 (61.3%) for both the 43a1 and 53a6 parameter sets. For the 53a5 parameter set the X-ray structure 2fxb showed the greatest decrease in HBs (60.2%). With the 43a1 parameter set, in 7 of the 72 simulations (10%), <10% of the HBs were lost. In 23 simulations (32%), between 10% and 20% of the HBs were lost. In 26 simulations (36%), the loss was between 20% and 30% and in 16 (22%) the loss was >30%. In the case of the parameter set 53a5, the corresponding numbers are 7 (<10%), 35 (>10% and <20%), 16 (>20% and <30%), and 14 (>30%). With the 53a6 parameter set, these numbers become 12, 26, 20, and 14, respectively. The average loss of HBs was 22.5%, 21.6%, and 22.2% for the 43a1, 53a5, and 53a6 parameter sets, respectively. All of the parameter sets showed a similar ability to preserve the HBs between elements of secondary structure.

It should be noted in regard to the HB analysis that (1) only HBs present within elements of secondary structure in the experimental structure were considered in the analysis and (2) that the values listed above relate to whether the specific donor and acceptor pairs involved satisfied the geometric criteria used to define a HB in each frame of the trajectory. As even small fluctuations in the structure during the simulations can result in a transient loss of a given HB, all structures using all force fields show an apparent net loss. This HB definition was used to enable a direct comparison between the force fields. Although expressed as an apparent net "loss" this value does not directly reflect the quality of the force field nor should it be used to infer a systematic problem with any of the force fields. A 20% loss in the number of HBs, for example, primarily reflects the fact that there are fluctuations within the elements of secondary structure. It does not translate into a 20% loss of secondary structure.

Solvent Accessible Surface Area. In addition to the rmsd, RG, and HB, the hydrophobic and hydrophilic solvent accessible surface areas (HSA and PSA) were also analyzed between 4 and 5 ns. The raw data are available as Supporting Information (Tables S5 and S6). The deviation from the value obtained from the experimental structure is shown in Figure 5. On average, the HSA decreased with respect to the experimental value by 4.3% for the 43a1 parameter set but increased by 9.2% for both the 53a5 and 53a6 parameter sets. In the case of PSA, the sample mean increased by 0.2% for 43a1 and by 6.5% and 5.3% for 53a5 and 53a6, respectively. In general, these results are consistent with the RG values and show that the proteins tend to interact more strongly with water when using the 53a5/6

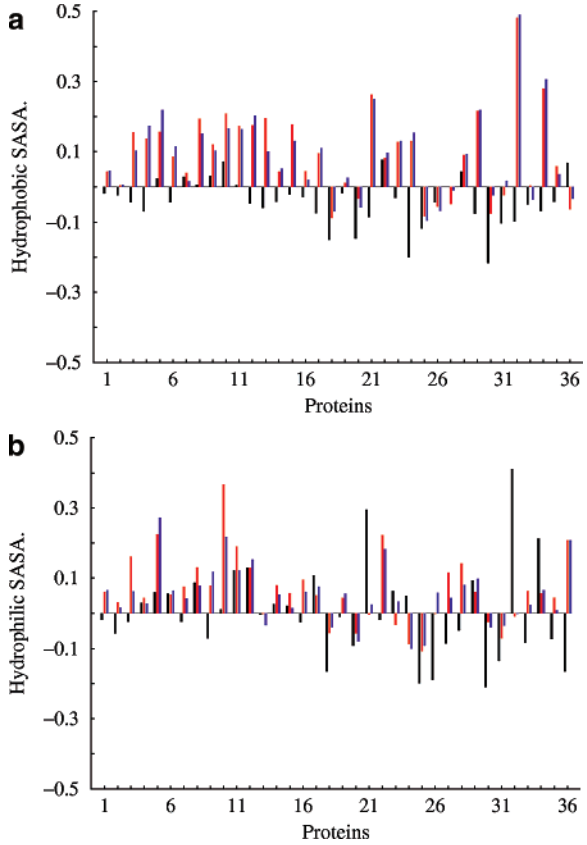

Figure 5. The relative deviation in the hydrophobic (a) and hydrophilic (b) solvent accessible surface area with respect to that found in the starting experimental structure. The values are an average over 5 ns for the duplicate simulations and plotted as a function of the protein ID (see Table 1): 43a1 (black), 53a5 (red), and 53a6 (blue).

parameter sets than with the 43a1 parameter set. The increased HSA might be partly due to the higher repulsion between aliphatic groups in the 53a5/6 parameter sets than 43a1.

Simulation Time. To examine the sensitivity of the results to the length of the simulations, one of the two 5 ns simulations was extended to 10 ns for each parameter set and protein combination. The various structural properties discussed above—rmsd, SSERMSD, RG, HB, HSA, and PSA—were again monitored during the simulations and averaged over the time period from 4 to 10 ns. The raw results are presented as Supporting Information (Tables S1−S6). In terms of the deviation from the starting experimental structure, doubling the length of the simulations to 10 ns did not significantly alter the overall results. The SSERMSD in particular are very similar to those discussed above. With the 43a1 parameter set, 25 out of the 36 simulations (70%) had a SSERMSD of <0.20 nm, 8 (22%) had SSERMSD values in the range 0.20 to 0.40 nm, and only 3 (8%) had SSERMSD values >0.40 nm. The corresponding numbers for the 53a5 and 53a6 parameter sets respectively were 20 and 21 (<0.20 nm), 13 and 10 (>0.20 and <0.40 nm), and 3 and 5 (>0.40 nm). The mean values of SSERMSD averaged over all of the extended simulations were 0.18, 0.20, and 0.22 nm for 43a1, 53a5, and 53a6. Similar to SSERMSD, the differences in the results for the RG and HSA/PSA were relatively small. In contrast, the results of the HB analysis differed markedly between the two periods in the cases of 43a1 and 53a5 parameter sets. In the period from 4 to 10 ns, no structure maintained >90% of the HBs with the 43a1 parameter set. In 11 simulations (31%) the loss was between 10% and 20%, in 17 (47%) the loss was between 20% and 30%, and in 8 (22%) the loss was more than 30%. With the 53a5 parameter set, the corresponding numbers are 1 (<10%), 19 (>10% and <20%), 8 (>20% and <30%), and 8 (>30%). Only with the 53a6 parameter set did the number of HBs remain roughly the same during the two periods: 5 out of 36 in the range from 0 to 10%, 14 in the range from 10% to 20%, 11 in the range from 20% to 30%, and 6 in the range >30%. Averaged over the different parameter sets the values were −25.4%, −23.4%, and −22.8% for 43a1, 53a5, and 53a6, respectively. Again it should be noted that these values include transient fluctuations in secondary structure and do not necessarily involve the actual loss of elements of secondary structure.

TABLE 2: Averages and Contrasts (Differences) of the Relative Deviations of Descriptive Properties for the Different Force Fields

property	average	43a1−53a5	43a1−53a6	53a5−53a6
SSERMSD[a]	0.194	−0.023	−0.041	−0.018
RG[b]	0.008	−0.038	−0.034	0.004
HB[c]	−0.221	−0.009	−0.004	0.005
HSA[d]	0.040	−0.063	−0.051	0.012
PSA[e]	0.047	−0.136	−0.135	0.000
P-value[f]		4.47×10^{-34}	1.14×10^{-35}	0.55

[a] Positional root-mean-square deviation of secondary structure elements with respect to the experimental structures. [b] Radius of gyration. [c] The number of intramolecular hydrogen bonds associated with secondary structure elements as defined in the experimental structure. [d] Hydrophobic solvent accessible surface. [e] Hydrophilic solvent accessible surface. [f] Probability that the difference vector is equal to the null vector determined by using the Roy union−intersection approach.

Statistical Analysis. As is very evident from the distributions shown in Figures 2−5 and in the raw data provide as Supporting Information there is a large degree of variation not only between the different proteins but also between the replicate simulations performed with the different parameter sets. The question is, given such variation, to what extent is it possible to attribute the differences in the simulations to the effect of a specific parameter set? Or, more critically, are any trends observed throughout this set of 216 simulations statistically significant? To address these questions the set of simulations performed with each parameter set was characterized in terms of an observation vector that included the following five instantaneous properties: SSERMSD, RG, HB, HSA, and PSA. This observation vector was assessed for each protein under each of the three alternative parameter sets. Then by using a MANOVA approach (see methods and ref 27) it was tested whether the results obtained with the three parameter sets were equivalent given the intrinsic variation in the data. It was found that the probability that these sets are equal with regard to the results is negligible ($p < 2.2 \times 10^{-16}$) and that there was a statistically significant contribution due to interactions between the parameter set and the specific protein. That is, the differences observed between the parameter sets were clearly significant considering the three parameter sets and all of the simulations simultaneously.

To identify if a specific parameter set (or sets) was responsible for the rejection of the hypothesis of equal effects (null hypothesis), contrasts were tested by using the Roy union−intersection method.[27,77,78] Differences between the mean vectors of each pair of sets are given in Table 2, together with the *p*-values denoting the probability that the vector is equal to the null vector (no differences). The table shows that the difference between the two parameter sets, 53a5 and 53a6, is not statistically significant at the 95% confidence level ($p = 0.55$). However, the 43a1 set is found to be different from that of both 53a5 ($p < 4.5 \times 10^{-34}$) and 53a6 ($p < 1.1 \times 10^{-35}$).

TABLE 3: Results for Multilinear Regressions of Differences of Selected Properties between Force Fields 53a5 and 43a1 (a) and 53a6 and 43a1 (b) against a Number of Properties Characterizing the Proteins[a]

	SSERMSD			RG			HB			HSA			PSA			
regressor[b]	β[c]	p[d]	R^2_{adj}[e]	β[c]	p[d]	R^2_{adj}[e]	β[c]	p[d]	R^2_{adj}[e]	β[c]	p[d]	R^2_{adj}[e]	β[c]	p[d]	R^2_{adj}[e]	MCC[f]
(a) 53a5 and 43a1 parameter sets																
N_{res}	0.000	0.924	−0.029	0.000	0.336	−0.001	0.001	0.377	−0.022	**0.004**	**0.050**	**0.083**	−0.002	0.613	−0.006	0.249
$N_{cha.res}$	0.000	0.979	−0.029	0.000	0.639	−0.023	0.001	0.395	−0.024	−0.003	0.662	−0.024	0.004	0.665	−0.007	0.049
N_{cha}	**−0.006**	**0.015**	**0.138**	**−0.003**	**0.001**	**0.259**	0.006	0.245	0.176	**−0.014**	**0.004**	**0.192**	**0.005**	**0.006**	**0.011**	0.410
$N_{pos.cha}$	−0.007	0.059	0.075	**−0.003**	**0.033**	**0.102**	0.009	0.131	0.136	**−0.020**	**0.012**	**0.147**	**0.011**	**0.015**	**0.038**	0.274
$N_{neg.cha}$	0.006	0.089	0.056	**0.003**	**0.008**	**0.163**	−0.006	0.772	0.048	0.013	0.091	0.055	−0.002	0.105	−0.027	0.233
$N_{cha.res}$ %	−0.077	0.755	−0.026	−0.032	0.693	−0.025	0.037	0.050	−0.029	**−1.222**	**0.013**	**0.144**	0.840	0.874	0.082	0.259
$N_{pos.res}$ %	−0.530	0.052	0.080	**−0.209**	**0.018**	**0.129**	0.476	0.051	0.069	**−1.880**	**0.000**	**0.289**	0.950	0.066	0.081	0.407
$N_{neg.res}$ %	**0.267**	**0.020**	**0.124**	**0.107**	**0.004**	**0.201**	−0.274	0.262	0.152	**0.600**	**0.012**	**0.149**	**−0.237**	**0.011**	**0.009**	0.338
α-helix	0.001	0.566	−0.019	0.000	0.539	−0.018	0.001	0.360	−0.016	0.002	0.203	0.019	−0.001	0.504	−0.004	0.116
β-sheet	0.000	0.803	−0.028	−0.001	0.053	0.079	0.000	0.687	−0.029	−0.002	0.353	−0.003	0.001	0.911	−0.024	0.163
loop	−0.001	0.620	−0.022	**0.001**	**0.019**	**0.125**	0.000	0.979	−0.026	0.003	0.216	0.017	0.000	0.735	−0.029	0.210
α-helix %	0.028	0.639	−0.023	0.010	0.600	−0.021	0.038	0.421	−0.016	0.114	0.366	−0.005	−0.087	0.504	−0.010	0.083
β-sheet%	−0.001	0.985	−0.029	**−0.048**	**0.040**	**0.092**	−0.028	0.530	−0.024	−0.198	0.186	0.023	0.081	0.684	−0.017	0.192
loop %	−0.081	0.435	−0.011	**0.067**	**0.044**	**0.088**	−0.056	0.630	−0.020	0.071	0.744	−0.026	0.090	0.573	−0.022	0.198
sulfur %	0.341	0.650	−0.023	0.452	0.059	0.074	−1.877	0.323	0.186	2.071	0.179	0.025	**−1.314**	**0.005**	**0.000**	0.293
aromatic %	0.373	0.239	0.012	−0.021	0.842	−0.028	0.068	0.696	−0.028	−0.089	0.894	−0.029	0.223	0.823	−0.025	0.086
amide %	−0.186	0.644	−0.023	−0.009	0.945	−0.029	0.257	0.969	−0.015	0.143	0.865	−0.029	−0.028	0.498	−0.029	0.025
(b) 53a6 and 43a1 parameter sets																
N_{res}	0.002	0.134	0.037	**0.001**	**0.038**	**0.095**	−0.001	0.160	0.029	−0.001	0.512	−0.016	**0.005**	**0.011**	**0.151**	0.371
$N_{cha.res}$	**0.008**	**0.018**	**0.130**	0.001	0.188	0.023	−0.003	0.282	0.006	0.005	0.304	0.003	0.000	0.947	−0.029	0.228
N_{cha}	**−0.008**	**0.007**	**0.175**	**−0.002**	**0.004**	**0.195**	**0.007**	**0.001**	**0.272**	0.005	0.257	0.009	**−0.012**	**0.009**	**0.162**	0.484
$N_{pos.cha}$	−0.002	0.660	−0.024	−0.002	0.220	0.016	0.006	0.080	0.060	0.011	0.107	0.048	−0.015	0.051	0.082	0.198
$N_{neg.cha}$	**0.016**	**0.000**	**0.307**	**0.003**	**0.003**	**0.207**	**−0.009**	**0.002**	**0.236**	−0.001	0.874	−0.029	0.013	0.059	0.075	0.520
$N_{cha.res}$ %	0.398	0.205	0.019	−0.028	0.710	−0.025	0.012	0.956	−0.029	**0.848**	**0.050**	**0.082**	**−1.056**	**0.020**	**0.124**	0.269
$N_{pos.res}$ %	−0.452	0.206	0.019	**−0.179**	**0.028**	**0.109**	**0.507**	**0.025**	**0.115**	**0.968**	**0.049**	**0.083**	**−1.622**	**0.001**	**0.251**	0.374
$N_{neg.res}$ %	**0.415**	**0.004**	**0.195**	**0.087**	**0.012**	**0.149**	**−0.301**	**0.001**	**0.252**	−0.251	0.239	0.012	**0.516**	**0.019**	**0.127**	0.439
α-helix	0.002	0.062	0.072	**0.001**	**0.020**	**0.125**	**−0.001**	**0.041**	**0.091**	0.000	0.873	−0.029	0.003	0.062	0.073	0.324
β-sheet	−0.002	0.224	0.015	**−0.001**	**0.008**	**0.168**	0.001	0.249	0.011	−0.001	0.766	−0.027	−0.002	0.254	0.010	0.229
loop	0.000	0.845	−0.028	0.001	0.093	0.054	0.000	0.824	−0.028	0.000	0.887	−0.029	0.003	0.242	0.012	0.145
α-helix %	0.110	0.155	0.031	0.031	0.084	0.058	−0.082	0.096	0.052	−0.007	0.947	−0.029	0.150	0.192	0.022	0.184
β-sheet %	−0.122	0.187	0.023	**−0.057**	**0.006**	**0.180**	0.074	0.218	0.016	−0.009	0.944	−0.029	−0.222	0.104	0.049	0.288
loop%	−0.073	0.586	−0.020	0.026	0.400	−0.008	0.092	0.287	0.005	0.041	0.830	−0.028	0.014	0.946	−0.029	0.081
sulfur %	1.261	0.186	0.023	0.386	0.081	0.060	**−1.650**	**0.005**	**0.184**	−1.360	0.311	0.002	1.675	0.239	0.012	0.347
aromatic %	0.569	0.161	0.029	0.123	0.196	0.021	0.298	0.260	0.009	0.239	0.678	−0.024	0.094	0.878	−0.029	0.171
amide %	**−0.991**	**0.049**	**0.083**	−0.173	0.149	0.033	0.436	0.189	0.022	−0.102	0.888	−0.029	−0.030	0.969	−0.029	0.190

[a] Each line in the table summarizes the results of a multiple regression of the five measured properties against the regressor on the left. Results which were found to be statistically significant are shown in bold type. The table shows how the differences between the relative deviations from the references for force fileds 53a6 and 43a1 (part a) and 53a5 and 43a1 (part b) are related to properties characterizing the proteins. A positive slope (β) indicates that the differences become larger with increasing values of the regressor. A negative value for β indicates that the difference becomes smaller with increasing values of the regressor. [b] The regressors include N_{res} (the number of residues in a structure), $N_{cha.res}$ (the number of charged residues at neutral pH), N_{cha} (the net charge), $N_{pos.cha}$ (the number of positively charged residues), $N_{neg.cha}$ (the number of negatively charged residues), $N_{cha.res}$ % (the percentage of charged residues in a structure), α-helix (the number of residues belonging to α-helical structures), sulfur % (the percentage of Met and Cys residues in a structure), aromatic % (the percentage of residues containing aromatic groups), amide % (the percentage of Asn and Gln residues in a structure). [c] The slope of the regression line. [d] The probability that the slope equals 0 (no regression relation). [e] R^2_{adj} denotes the adjusted correlation coefficient. [f] MCC is the multilinear correlation coefficient and indicates the extent to which the joint observed values are correlated to a regressor.

To gain further insight into the source(s) of the differences the data were subjected to regression analysis with a number of characteristic properties of the proteins as independent variables. These properties were related to the charge on the protein, the secondary structure content, and the presence of specific residue types, such as aromatic and sulfur containing residues. The results for the regression analysis are shown in Table 3a,b. In this table each line summarizes the results of performing a multiple regression analysis on the five descriptive properties obtained from the simulations against the single independent variable indicated on the left. The multiple correlation coefficient (MCC) indicates how well the five properties together correlate to the characteristic property of the protein. Note that the dependent variable is the difference in the specific structural properties (SSERMSD, RG, HB, HSA, PSA) between using the 53a5 and 43a1 parameter sets or between using the 53a6 and 43a1 parameter sets. Thus, a positive number indicates that the measured value for a property is increasing for the 53a5/6 parameter sets relative to the 43a1 parameter set as the values for that property of the protein increase. From Table 3 it can be seen that the highest MCCs are obtained for properties related to the charge on the protein. In fact, the differences between the 53a5/53a6 parameter sets and the 43a1 parameter set correlate most consistently with the net charge in the protein. In the case of 53a5 it is the highest MCC value (0.410) whereas for 53a6 the net charge corresponds to the second highest MCC value (0.484). The highest MCC value in this case was obtained for the number of negative charges (0.520). However, although the effect of charge shows the strongest correlation, in all cases the correlations are relatively weak and it would be inappropriate to draw to specific conclusions in regard to the significance of these results.

Discussion

The three different versions of the GROMOS96 force field compared in this study, the original 43a1 parameter set and two most recent versions of the force field the 53a5 and 53a6 parameter sets, differ significantly in terms of both the non-

bonded parameters themselves and the experimental data against which they were parametrized. In particular, the three parameter sets differ markedly in their ability to reproduce the liquid properties of certain small organic molecules and also in their ability to reproduce the partitioning behavior of analogues of the 18 common amino acids between a polar (water) and an apolar (cyclohexane) environment. The greatest differences between the parameter sets are in the description of the aromatic amino acids (histidine, phenylalanine, tryptophan, tyrosine), the description of the neutral polar amino acids (asparagine, glutamine, serine, threonine), and the description of the amino acids that contain sulfur (methionine, cysteine).

In terms of the effect on the structural properties analyzed in this work the most important difference between the parameter sets relates to their ability to reproduce the correct partitioning behavior of the amino acids as this will determine whether specific amino acids will remained buried within the matrix of the protein or will interact strongly with water. Indeed the development of the 53a5 and 53a6 parameter sets was prompted by the fact that the relative free energy of solvation between cyclohexane and water for certain compounds such as methyl ethyl sulfide, an analogue of the side chain of methionine, not only differed from the experimental value by >15 kJ/mol but also had the incorrect sign. The average deviation from the experimental solvation free energies for the analogues of the amino acids listed above with use of the 43a1 parameter set was 10.0 kJ/mol in water and 3.0 kJ/mol in cyclohexane. In contrast, with the 53a6 parameter set the solvation free enthalpies in water and cyclohexane are reproduced to within 0.8 and 2.2 kJ/mol, respectively.

In this study, we attempted to determine whether the changes within the GROMOS96 force field parameter sets led to systematic, measurable differences in the simulations on a 5−10 ns time scale. The results show that given a sufficiently large sample (in this case 36 structures and two replicate simulations) systematic differences attributable to changes in the force field could be detected between the parameter sets 53a5/6 and the parameter set 43a1. No statistically significant differences could, however, be detected between the 53a5 and the 53a6 parameter sets. This does not mean that the force fields are equivalent, only that the differences are not significant given the expected variation within a set of simulations performed with a given force field on the time scale investigated (5−10 ns).

Regression analysis showed that the differences between the simulations performed with the 43a1 and the 53a5/6 parameter sets were primarily correlated with the presence of charged residues in the proteins. This is despite the fact that the parameters of the charge residues themselves were largely unchanged between the parameter sets. However, the charges on the polar residues are larger in the 53a5/6 parameter sets compared with the 43a1 parameter set. With the 53a5/6 parameter sets the structures tended to be more hydrated and this was associated with a slight increase in the structure deviation (rmsd), radii of gyration (RG), and solvent accessible surface area (HSA and PSA). Other properties investigated, such as the secondary structure content of a protein, were poor determinants of the differences observed between the parameter sets. In particular there was no obvious correlation between the presence of specific amino acids that had been extensively revised during the development of the 53a5/6 parameter sets (aromatic, amide, and sulfur-containing residues) and the differences between the simulations. This may seem surprising but it simply highlights the difficulty in assessing the quality of a force field by comparing the properties of whole proteins given that there may be compensating changes within a given force field.

TABLE 4: The Absolute Differences between Two Replicate Simulations Averaged over the Three Parameter Sets[a]

PDBID	SSERMSD	RG	HB	HSA	PSA
1vif	0.02	0.01	0.52	0.44	0.54
1tuc	0.02	0.01	0.66	0.64	0.47
1vcc	0.10	0.02	0.88	0.32	0.92
1ail	0.07	0.02	6.37	1.29	0.88
1cei	0.21	0.02	2.88	1.38	1.67
1ctf	0.08	0.01	2.31	0.64	0.93
1pgx	0.04	0.02	0.92	1.00	0.21
1tif	0.16	0.04	2.22	0.91	0.40
2acy	0.01	0.00	0.60	0.95	0.21
2fxb	0.13	0.07	2.62	1.06	0.41
1r69	0.13	0.04	0.52	0.42	1.78
1bm8	0.05	0.01	1.49	0.78	0.74
2ci2	0.02	0.02	0.25	1.27	1.62
1pgb	0.10	0.02	1.08	0.13	0.40
1shg	0.01	0.01	0.34	0.33	0.72
1ubi	0.02	0.02	0.63	1.02	0.90
1a19	0.03	0.01	1.17	0.84	0.46
1aoy	0.05	0.03	1.29	0.59	0.58
1stu	0.03	0.01	1.33	0.75	1.16
1sro	0.02	0.02	0.78	0.85	1.68
1sap	0.05	0.03	1.06	0.49	0.60
1afi	0.03	0.01	1.12	0.82	0.30
1bb8	0.03	0.05	1.66	2.04	1.98
2bby	0.03	0.01	1.09	0.97	1.17
2fmr	0.08	0.01	1.05	1.57	1.34
1a1z	0.15	0.03	4.50	2.08	1.54
1bw6	0.11	0.12	0.85	1.14	1.37
1coo	0.04	0.07	0.61	1.27	2.28
1lea	0.08	0.01	0.27	0.88	0.93
2af8	0.14	0.03	3.28	1.98	1.93
2ezh	0.14	0.01	1.42	0.22	0.57
3ci2	0.02	0.01	0.38	1.07	0.10
2gb1	0.02	0.02	0.16	0.39	0.41
1aey	0.00	0.01	0.54	0.76	0.45
1d3z	0.01	0.01	1.44	0.77	1.32
1bta	0.02	0.01	1.40	0.37	0.52
average	0.06	0.02	1.38	0.90	0.93

[a] The differences correspond to the values obtained when the results for each individual simulation (see the Supporting Information) are compared to the values obtained for the experimental structure. They do not correspond to the differences between the replicate simulations themselves. Units: SSERMSD (nm), RG (nm), HSA (nm^2), and PSA (nm^2).

This work also provides an indication of the magnitude of the variation in a range of properties that might reasonably be expected when comparing the results of replicate simulations of a given system or simulations of two closely related systems on a multinanosecond time scale to a particular reference structure. In Table 4 the absolute difference between the two replicate simulations for the five properties monitored, averaged over the three parameter sets is reported. For each property, the differences averaged over the whole test set are also given. The results in Table 4 reflect the chaotic nature of atomistic simulations of proteins, which are highly dependent on the precise initial conditions. For example, in the case of the protein 1cei there is an average difference of almost 3 backbone hydrogen bonds between replicate simulations after 5 ns. Note, this corresponds to the difference between the replicate simulations and the initial structure. Clearly, the differences between simulations themselves would in general be larger. The average difference in rmsd for residues in elements of secondary structure comparing the final structures after 5 ns performed with the same parameter sets was 0.06 nm. However, for certain proteins this value was much larger. Whereas in one simulation

of the protein 1tuc using the 53a5 parameter set the SSERMSD from the initial structure was only 0.08 nm in the replicate simulation it was 0.26 nm. In more extreme examples the values of the SSERMSD from the initial structure varied from 0.17 to 0.75 nm. This underlines the need for extreme caution when attempting to use a single or small number of nanosecond simulations to draw conclusions either in regard to the validity of a given force field or in regard to the functional significance of structural changes associated with changes in simulation conditions or amino acid substitutions.

Conclusions

In this study the effect of different parameter sets on the outcome of simulations of a number of proteins has been assessed. A series of 216 simulations consisting of 2 simulations starting from 36 different initial structures corresponding to 31 different proteins in combination with three different GROMOS parameter sets (43a1, 53a5, and 53a6) were performed giving a total sampling time of 1.62 μs. Although there are significant differences between the parameter sets in terms of the description of the individual amino acids, no major differences in a range of structural properties including the root-mean-square deviation, radius of gyration, solvent accessible surface area, and number of H-bonds in secondary structures were apparent on a multinanosecond time scale. In fact, it was only by combining all data and using a detailed statistical analysis that it could be demonstrated that there were systematic differences between 53a5/6 and 43a1 sets and that these were significant.

The study also provides information on the variation in various structural properties that can be expected in replicate simulations performed with any given force field. We have shown that there can be large variations in the outcome of even short simulations performed under almost identical conditions and that the expected variation in the properties investigated must be considered when analyzing the results from any simulation. In particular, the current study underlines the need to move toward the use of large data sets in conjunction with statistically rigorous methods if we are to draw meaningful conclusions based on simulation studies of proteins and other biomolecules especially in regard to determining the validity of any given force field.

Supporting Information Available: Tables showing the raw data on which the analysis presented in the main body of the paper is based. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Hünenberger, P. H.; van Gunsteren, W. F. In *Computer simulation of biomolecular systems, theoretical and experimental applications;* van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1997; pp 3−82.

(2) MacKerell, A. D., Jr. *J. Comput. Chem.* **2004**, *25*, 1584.

(3) Jorgensen, W. L.; Tirado-Rives, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6665.

(4) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Mertz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(5) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. A. *J. Comput. Chem.* **2003**, *24*, 1999.

(6) MacKerell, A. D., Jr.; Bashford, D.; Bellot, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(7) Foloppe, N.; MacKerell, A. D. *J. Comput. Chem.* **2000**, *21*, 86.

(8) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation*: *GROMOS96 Manual and User Guide*; Hochschulverlag AG an der ETH Zürich: Zürich, Switzerland, 1996.

(9) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656.

(10) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(11) van Gunsteren, W. F.; Mark, A. E. *J. Chem. Phys.* **1998**, *108*, 6109.

(12) Villa, A.; Mark, A. E. *J. Comput. Chem.* **2002**, *23*, 548.

(13) MacCallum, J. L.; Tieleman, D. P. *J. Comput. Chem.* **2003**, *24*, 1930.

(14) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740.

(15) Roccatano, D.; Nau, W. M.; Zacharias, M. *J. Phys. Chem. B* **2004**, *108*, 18734.

(16) Mu, Y.; Kosov, D. S.; Stock, G. *J. Phys. Chem. B* **2003**, *107*, 5064.

(17) Gnanakaran, S.; García, A. E. *Proteins* **2005**, *59*, 773.

(18) Hess, B.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2006**, *110*, 17616.

(19) Jorgensen, W. L.; Ulmschneider, J. P.; Tirado-Rives, J. *J. Phys. Chem. B* **2004**, *108*, 16264.

(20) Price, D. J.; Brooks, C. L., III *J. Comput. Chem.* **2002**, *23*, 1045.

(21) van der Spoel, D.; Lindahl, E. *J. Phys. Chem. B* **2003**, *107*, 11178.

(22) Sorin, E. J.; Pande, V. S. *Biophys. J.* **2005**, *88*, 2472.

(23) Patel, S.; Mackerell, A. D.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1504.

(24) Soares, T. A.; Daura, X.; Oostenbrink, C.; Smith, L. J.; van Gunsteren, W. F. *J. Biomol. NMR* **2004**, *30*, 407.

(25) Oostenbrink, C.; Soares, T. A.; van der Vegt, N. F. A.; van Gunsteren, W. F. *Eur. Biophy. J. Biophy. Lett.* **2005**, *34*, 273.

(26) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. *J. Comput. Chem.* **2001**, *22*, 1205.

(27) Wassenaar, T. A.; Mark, A. E. *J. Comput. Chem.* **2006**, *27*, 316.

(28) Hermans, J.; Berendsen, H. J. C.; van Gunsteren, W. F.; Postma, J. P. M. *Biopolymers* **1984**, *23*, 1513.

(29) Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **1998**, *19*, 535.

(30) Narayana, N.; Matthews, D. A.; Howell, E. E.; Nguyen-huu, X. *Nat. Struct. Biol.* **1995**, *2*, 1018.

(31) Viguera, A. R.; Blanco, F. J.; Serrano, L. *J. Mol. Biol.* **1995**, *247*, 670.

(32) Sharma, A.; Hanai, R.; Mondragon, A. *Structure* **1994**, *2*, 767.

(33) Liu, J.; Lynch, P. A.; Chien, C. Y.; Montelione, G. T.; Krug, R. M.; Berman, H. M. *Nat. Struct. Biol.* **1997**, *4*, 896.

(34) Chak, K. F.; Safo, M. K.; Ku, W. Y.; Hsieh, S. Y.; Yuan, H. S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6437.

(35) Leijonmarck, M.; Liljas, A. *J. Mol. Biol.* **1987**, *195*, 555.

(36) Achari, A.; Hale, S. P.; Howard, A. J.; Clore, G. M.; Gronenborn, A. M.; Hardman, K. D.; Whitlow, M. *Biochemistry* **1992**, *31*, 10449.

(37) Biou, V.; Shu, F.; Ramakrishnan, V. *EMBO J.* **1995**, *14*, 4056.

(38) Thunnissen, M. M.; Taddei, N.; Liguri, G.; Ramponi, G.; Nordlund, P. *Structure* **1997**, *5*, 69.

(39) Fukuyama, K.; Okada, T.; Kakuta, Y.; Takahashi, Y. *J. Mol. Biol.* **2002**, *315*, 1155.

(40) Mondragon, A.; Subbiah, S.; Almo, S. C.; Drottar, M.; Harrison, S. C. *J. Mol. Biol.* **1989**, *205*, 189.

(41) Xu, R. M.; Koch, C.; Liu, Y.; Horton, J. R.; Knapp, D.; Nasmyth, K.; Cheng, X. *Structure* **1997**, *5*, 349.

(42) McPhalen, C. A.; James, M. N. *Biochemistry* **1987**, *26*, 261.

(43) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 4721.

(44) Musacchio, A.; Noble, M.; Pauptit, R.; Wierenga, R.; Saraste, M. *Nature* **1992**, *359*, 851.

(45) Ramage, R.; Green, J.; Muir, T. W.; Ogunjobi, O. M.; Love, S.; Shaw, K. *Biochem. J.* **1994**, *299*, 151.

(46) Ratnaparkhi, G. S.; Ramachandran, S.; Udgaonkar, J. B.; Varadarajan, R. *Biochemistry* **1998**, *37*, 6958.

(47) Sunnerhagen, M.; Nilges, M.; Otting, G.; Carey, J. *Nat. Struct. Biol.* **1997**, *4*, 819.

(48) Bycroft, M.; Grunert, S.; Murzin, A. G.; Proctor, M.; St Johnston, D. *EMBO J.* **1995**, *14*, 3563.

(49) Bycroft, M.; Hubbard, T. J.; Proctor, M.; Freund, S. M.; Murzin, A. G. *Cell* **1997**, *88*, 235.

(50) Edmondson, S. P.; Qiu, L.; Shriver, J. W. *Biochemistry* **1995**, *34*, 13289.

(51) Steele, R. A.; Opella, S. J. *Biochemistry* **1997**, *36*, 6885.

(52) Connolly, K. M.; Wojciak, J. M.; Clubb, R. T. *Nat. Struct. Biol.* **1998**, *5*, 546.

(53) Groft, C. M.; Uljon, S. N.; Wang, R.; Werner, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9117.
(54) Musco, G.; Kharrat, A.; Stier, G.; Fraternali, F.; Gibson, T. J.; Nilges, M.; Pastore, A. *Nat. Struct. Biol.* **1997**, *4*, 712.
(55) Eberstadt, M.; Huang, B.; Chen, Z.; Meadows, R. P.; Ng, S. C.; Zheng, L.; Lenardo, M. J.; Fesik, S. W. *Nature* **1998**, *392*, 941.
(56) Iwahara, J.; Kigawa, T.; Kitagawa, K.; Masumoto, H.; Okazaki, T.; Yokoyama, S. *EMBO J.* **1998**, *17*, 827.
(57) Jeon, Y. H.; Negishi, T.; Shirakawa, M.; Yamazaki, T.; Fujita, N.; Ishihama, A.; Kyogoku, Y. *Science* **1995**, *270*, 1495.
(58) Fogh, R. H.; Ottleben, G.; Ruterjans, H.; Schnarr, M.; Boelens, R.; Kaptein, R. *EMBO J.* **1994**, *13*, 3936.
(59) Crump, M. P.; Crosby, J.; Dempsey, C. E.; Parkinson, J. A.; Murray, M.; Hopwood, D. A.; Simpson, T. J. *Biochemistry* **1997**, *36*, 6000.
(60) Clubb, R. T.; Schumacher, S.; Mizuuchi, K.; Gronenborn, A. M.; Clore, G. M. *J. Mol. Biol.* **1997**, *273*, 19.
(61) Ludvigsen, S.; Shen, H. Y.; Kjaer, M.; Madsen, J. C.; Poulsen, F. M. *J. Mol. Biol.* **1991**, *222*, 621.
(62) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657.
(63) Blanco, F. J.; Ortiz, A. R.; Serrano, L. *J. Biomol. NMR* **1997**, *9*, 347.
(64) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836.
(65) Lubienski, M. J.; Bycroft, M.; Freund, S. M.; Fersht, A. R. *Biochemistry* **1994**, *33*, 8866.
(66) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
(67) Fan, H.; Mark, A. E. *Proteins* **2003**, *53*, 111.
(68) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *95*, 43.
(69) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306.
(70) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. In *Intermolecular Forces*; Pullman B., Ed.; Reidel D. Publishing Company: Dordrecht, The Netherlands, 1981; p 331.
(71) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463.
(72) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952.
(73) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451.
(74) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(75) Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M. *J. Comput. Chem.* **1995**, *16*, 273.
(76) Wilks, S. S. *Annals Mathe. Stati.* **1932**, *3*, 163.
(77) Roy, S. N.; Bose, R. C. *Annals Mathe. Stati.* **1953**, *24*, 513.
(78) Morrisson, D. F. *Multivariate Statistical Methods*; McGraw-Hill series in probability and statistics; Blackwell, D., Solomon, H., Eds.; McGraw-Hill Kogakusha: Tokyo, Japan, 1976.