ARTICLE

The MUMO (minimal under-restraining minimal over-restraining) method for the determination of native state ensembles of proteins

Barbara Richter · Joerg Gsponer · Péter Várnai · Xavier Salvatella · Michele Vendruscolo

Abstract While reliable procedures for determining the conformations of proteins are available, methods for generating ensembles of structures that also reflect their flexibility are much less well established. Here we present a systematic assessment of the ability of ensemble-averaged molecular dynamics simulations with ensemble-averaged NMR restraints to simultaneously reproduce the average structure of proteins and their associated dynamics. We discuss the effects that under-restraining (overfitting) and over-restraining (underfitting) have on the structures generated in ensemble-averaged molecular simulations. We then introduce the MUMO (minimal under-restraining minimal over-restraining) method, a procedure in which different observables are averaged over a different number of molecules. As both over-restraining and under-restraining are significantly reduced in the MUMO method, it is possible to generate ensembles of conformations that accurately characterize both the structure and the dynamics of native states of proteins. The application of the MUMO method to the protein ubiquitin yields a high-resolution structural ensemble with an RDC Q-factor of 0.19.

B. Richter · J. Gsponer · P. Várnai · X. Salvatella · M. Vendruscolo (✉)
Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK
e-mail: mv245@cam.ac.uk

Introduction

An accurate description of the inherent flexibility of proteins in solution is often indispensable in order to account for important biochemical processes such as enzymatic catalysis, signal transduction and molecular recognition (Karplus and Petsko 1990; Wand 2001; Palmer 2004; Carlson and McCammon 2000; Carlson 2002; Ma and Nussinov 2002; Teague 2003; Wong and McCammon 2003; Perryman et al. 2004; Grunberg et al. 2004; Karplus and Kuriyan 2005). The combination of experimental methods and molecular dynamics simulations, which has emerged as a powerful way to determine the average structures of proteins (Brünger et al. 1986; Scott et al. 1999; Brünger et al. 1998; Spronk et al. 2004; Schwieters et al. 2005; Rieping et al. 2005), also represents a promising route for generating ensembles of structures representing their dynamics (Torda et al. 1989; Bonvin et al. 1994; Hess and Scheek 2003; Best and Vendruscolo 2004; Clore and Schwieters 2004a; Lindorff-Larsen et al. 2005; Clore and Schwieters 2006). In standard structure determination by nuclear magnetic resonance (NMR) spectroscopy, a penalty is applied in a molecular simulation if an individual molecule does not satisfy the experimental restraints, most commonly distances derived from nuclear Overhauser effect (nOe) measurements (Wüthrich 1986), so that the minimization of the penalty energy yields a protein structure in good agreement with all experimental restraints simultaneously. Since the measured data are derived experimentally as averages over an ensemble of molecules over time, the structure obtained represents a model for the average structure rather than the conformational heterogeneity present in solution. When the

fluctuations take place around a well-defined average the structure produced by such a procedure tends to have the same conformation as the most probable conformer. This description, however, becomes less accurate when structural fluctuations lead to the occupation of multiple conformers (Bonvin and Brunger 1995; Clore and Schwieters 2006; Zagrovic and van Gunsteren 2006). In this case, the model does not account for the entire range of statistically significant conformations. This problem may be referred to as over-restraining (or underfitting).

In order to reproduce the structure and dynamics of proteins, the NMR observables used as restraints in the simulations can be imposed as averages over several copies of the protein molecule rather than on a single one. This idea was first implemented for the case of nOe distances, by applying a penalty if the time-average of an NMR observable calculated from a molecular dynamics trajectory differs from experiment (Torda et al. 1989, 1990; Bonvin et al. 1994; Kemmink and Scheek 1995; Bonvin and Brunger 1995). In an alternative approach, penalizing forces are applied if the calculated average distances at a given time across an ensemble of simulated molecules (the "replica ensemble") do not match the experimental ones. Since the early implementations of the ensemble-averaged nOe distance restraints (Bonvin et al. 1994), a variety of restraining algorithms, including simultaneous time and ensemble averaging (Fennen et al. 1995), have been developed for an array of experimental observables measured for native, transition, intermediate, and unfolded states (Vendruscolo and Paci 2003; Vendruscolo and Dobson 2005; Kuszewski et al. 1999; Clore and Schwieters 2004a; Clore and Schwieters 2004b; Clore and Schwieters, 2006).

While increasing the size of the replica ensemble alleviates the underfitting problem, it eventually introduces the well-documented problem of overfitting, or under-restraining. Overfitting arises because the addition of replicas in ensemble-averaging procedures (or the lengthening of the time in time-averaging schemes) increases the number of degrees of freedom and hence the number of free parameters, while the available experimental information provided by the restraints remains constant. A better fit between the experimental and the back-calculated data from the simulation can simply reflect the increased number of free parameters, rather than a better agreement with the true distribution of structures (the "Boltzmann ensemble") (Bonvin and Brunger 1995). In order to prevent overfitting it would in principle be necessary to increase the information content of the restraints (for example by increasing their number) to compensate

for the increase in the number of degrees of freedom of the ensemble. If no additional restraints are added, the force field plays an increasingly important role in determining the distribution of structures.

Overfitting is a concern in both time- and ensemble-averaged restrained simulations since the accurate description of complex protein structures requires many more restraints than those that can be determined routinely by NMR. In an important study, Bürgi et al. analytically demonstrated the insensitivity of NMR observables, particularly those that are not linearly averaged, such as nOe distances, to the underlying probability distribution of conformational states (Bürgi et al. 2001). An additional problem is that several conformations may yield the same measurable observables, as is the case for the ^3J-couplings due to the existence of multiple solutions to the Karplus equation (Scott et al. 1998). It is therefore extremely difficult to extract the Boltzmann ensemble given a set of NMR measurements (Bürgi et al. 2001).

Several studies have specifically addressed the overfitting problem in ensemble-averaged simulations. Bonvin et al. demonstrated that in the case of molecular dynamics simulations restrained only with nOe distances, overfitting occurs already with two replicas (Bonvin and Brunger 1995). Although the restraints are better satisfied when the number of molecules in the replica ensemble is increased, agreement with non-restrained distances diminishes. In a later article, the same authors concluded that, although determining the individual conformers present in the Boltzmann ensemble in solution is feasible, the exact enumeration of their Boltzmann probabilities is only possible when exact distance restraints, rather than distance bins that incorporate error bounds, are used (Bonvin and Brünger 1996). More recently, a suggestion has been made to surmount the overfitting problem by requiring individual molecules in the replica ensemble to remain close to an average conformation (Scott et al. 1998; Clore and Schwieters 2004a; Clore and Schwieters 2004b). A variety of restraints were considered, which restrict the movements of individual replicas to areas of conformational space close to the average conformation. The authors concluded that backbone movements can be explained with just two replicas when residual dipolar couplings (RDCs) are used as restraints (Clore and Schwieters 2004a; Clore and Schwieters 2004b). In a later study, the same authors found that four to eight copies of the molecule are adequate to obtain internally consistent results when residual dipolar couplings, backbone order parameters, and crystallographic B-factors are used as restraints (Clore and Schwieters 2006). In another approach, S^2

order parameters, which contain information about the dynamics of individual bond vectors, were restrained (Best and Vendruscolo 2004). By applying this idea, sixteen-replica simulations were found to reproduce the native state structure and dynamics of ubiquitin, as measured by a validation with scalar ^3J-couplings and RDCs and by the prediction of stability changes associated with mutations (Lindorff-Larsen et al. 2005).

Despite the increasing interest in the use of ensemble-averaged molecular dynamics simulations for the characterization of protein dynamics, a systematic assessment of ensemble-averaged simulations restrained by nOe distances and S^2 order parameters has not yet been performed. Here, we address this issue by adopting an approach in which an unrestrained molecular dynamics simulation of ubiquitin is used to generate a collection of structures (the "reference ensemble"), which serves as a model of the Boltzmann ensemble of the protein in solution. NMR data are then back-calculated from the reference ensemble and used as restraints in computational procedures aimed at reconstructing the reference distribution of structures. It is thus possible to perform a cross-validation analysis in which both the average structure and the structural heterogeneity obtained from the restrained simulations are compared to those of the known reference ensemble. This type of comparison avoids problems related to possible inaccuracies in the experimental data (since the reference ensemble grants access to the exact atomic coordinates of all its members) and in the translation of experimental NMR signals into structural restraints (since the same definitions are used to back-calculate the restraints from the reference ensemble and to enforce them in the structure determination procedure). An assumption in this type of approach is that the reference ensemble generated by the molecular dynamics simulation is a faithful representation of the protein structure and dynamics in solution. A significant body of evidence supporting this assumption has been accumulating over the years (see e.g., Karplus and McCammon 2002; Karplus and Kuriyan 2005). Moreover, possible inaccuracies in the simulations do not affect the outcome of the present "computer experiment," in which we assess how well different molecular dynamics strategies with structural restraints recover a known reference ensemble. The use of reference ensembles has been exploited in X-ray crystallography studies (Kuriyan et al. 1986; Ichiye and Karplus 1988) and has also been used to generate synthetic NMR data (Bonvin and Brunger 1995; Schneider et al. 1999).

In this work, we back-calculate nOe distances and S^2 order parameters from the reference ensemble and

enforce them as restraints in ensemble-averaged simulated annealing simulations while systematically varying the number of molecules in the replica ensemble, and the types, number, and error bounds of the restraints. The replica ensembles generated by multiple cycles of simulated annealing are then pooled together to form a larger ensemble (the "restrained ensemble") for each type of simulation. We then apply a variety of measures to assess the similarity between the reference ensemble and the various restrained ensembles. We find that the quality of the agreement with the reference ensemble depends on the type of observable considered and that the optimal number of replicas depends on the type of distribution that is analyzed. Taken together, our results indicate that different replica numbers are necessary for different types of restraints in ensemble-averaged molecular simulations. In fact, we find that two replicas are optimal for nOe restraints, while higher replica numbers yield the best cross-validation for the S^2 restraint. We thus introduce here a simulation protocol, named MUMO (Minimal-Under-restraining Minimal-Over-restraining) in which S^2 order parameters are averaged over eight or sixteen replicas and nOes are averaged in a linked pairwise manner (Fig. 1). The application of the MUMO method to the experimental data on ubiquitin yields an ensemble of structures (PDB code 2RN2) with an RDC Q-factor of 0.19, compared to 0.26 obtained using the dynamics ensemble refinement (DER) method (Lindorff-Larsen et al. 2005), which was also calculated with nOes and S^2 order parameters, thus confirming the effectiveness of MUMO method in determining both the structure and dynamics of proteins.
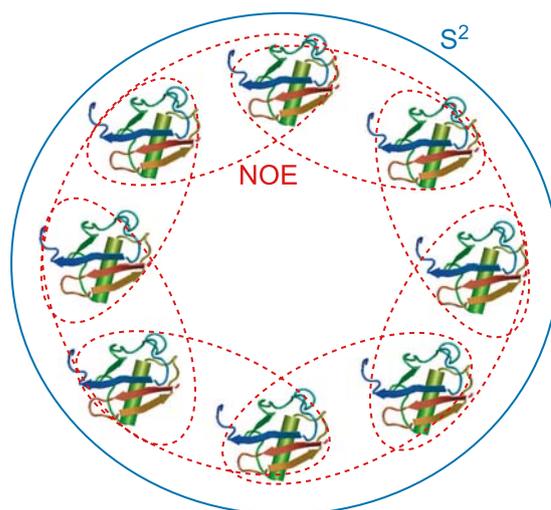
Fig. 1 Schematic illustration of the use of restraints in the MUMO procedure

Methods

Reference simulation

A 22 ns unrestrained molecular dynamics simulation was carried out for the 76-residue protein ubiquitin in the CHARMM22 force field (MacKerell et al. 1998) using the CHARMM molecular simulation package (Brooks et al. 1983). Ubiquitin, which is well characterized by NMR (Tjandra et al. 1995; Cornilescu et al. 1998; Lee et al. 1999; Chou et al. 2003) and used extensively as a model system in computational approaches (Clore and Schwieters 2004a; Lindorff-Larsen et al. 2005; Nederveen and Bonvin 2005), was solvated in a 4 Å explicit water shell containing 613 TIP3 water molecules (Jorgensen et al. 1983). The boundary potential of the MMFP module was used to prevent water molecules from escaping (Beglov and Roux 1994). The simulation was started from the minimized X-ray structure (Vijay-Kumar et al. 1987) and run at 300 K. The SHAKE algorithm was applied to all bonds to hydrogen atoms, eliminating high-frequency motions and thus allowing for an integration time step of 2 fs (Ryckaert et al. 1977). All calculations used an atom-based truncation scheme with a list cutoff of 14 Å, a non-bond cutoff of 12 Å, and the Lennard–Jones smoothing function initiated at 10 Å. Electrostatic and Lennard–Jones interactions were force switched. Atomic coordinates were saved every 500 integration steps (corresponding to every 1 ps) yielding 22,000 data points. The first 2,000 frames were discarded due to the potential existence of equilibration effects, leaving 20,000 frames for analysis.

Calculation of the nOe and S^2 restraints

Experimentally, nOe signals can be observed from protons less than about 6 Å apart (Neuhaus and Williamson 2000). For the reference simulation, all possible hydrogen–hydrogen interatomic distances were back-calculated using a $< r^{-3} >^{-1/3}$ average (Lindorff-Larsen et al. 2005; Neuhaus and Williamson 2000), resulting in 11,452 distances under 6 Å. These distances were split into short-range distances, where the two hydrogen atoms are on the same or adjacent residues, medium-range distances, where the two hydrogen atoms are between two and four residues apart, and long-range distances. Within each category, distances were randomly selected to match the proportions seen in experiment, yielding a total of 1,663 nOe restraints (Cornilescu et al. 1998); how the selected distances break down in terms of residue separation is shown in Table 1. In the *double* nOe

Table 1 Number of inter-proton distance restraints as a function of the distance along the polypeptide chain of residues i and j

Short-range	$|i–j| \leq 1$	825
Medium-range	$1 < |i–j| \leq 4$	234
Long-range	$|i–j| > 4$	604
Total restraints		1,663

simulations, approximately twice as many distances (3,446) were chosen, but the break-down between, short, medium, and long-range distances was left unchanged.

The nOe distances were binned using a lower bound of 0 and upper bounds of 3.5, 5, and 6 Å in order to reproduce the type of information available through the exploitation of nOe experiments. To incorporate an estimate of the experimental error, all the selected distances were increased by 15% and then placed in the appropriate bin based on the increased distance.

For the S^2 restraints, backbone amide and side-chain methyl S^2 order parameters were back-calculated from the reference trajectory using Eq. (6) (see below). Most backbone S^2 order parameters are well converged, as judged by investigating the tail of the autocorrelation function of the bond vectors, C_{tail} (Nederveen and Bonvin 2005; Chen et al. 2004). For all residues, except a few in the C-terminus of the protein, which is known to be highly flexible (Tjandra et al. 1995; Wang et al. 2003), we found $|C_{tail}–S^2| < 0.05$. Backbone amide order parameters were calculated for all residues except prolines and the N-terminal residue, resulting in 72 restraints. Fifty side-chain methyl order parameters were back-calculated for the following bonds: alanine $C_\alpha–C_\beta$, isoleucine $C_\gamma–C_\delta$ and $C_\beta–C\gamma$, leucine $C_\gamma–C_\delta$, methionine $S_\delta–C_\varepsilon$, threonine $C_\beta–C_\gamma$, and valine $C_\beta–C_\gamma$. Thus, a total of 122 S^2 order parameter restraints were applied in the ensemble-averaged simulations as described below.

Molecular dynamics simulations with ensemble-averaged restraints

The restraints were included in the force field by augmenting the CHARMM energy function

$$V_{TOTAL} = V_{CHARMM} + V_{nOe} + V_{S^2} \qquad (1)$$

The restraint terms, V_{nOe} and V_{S^2} are implemented as half-harmonic potentials that approach zero as the back-calculated observables approach the experimental one

$$V(\rho_X(t)) = \begin{cases} \frac{\alpha_X N_{rep}}{2}(\rho_X(t) - \rho_{X,0}(t))^2, & \rho_X(t) > \rho_{X,0}(t) \\ 0, & \rho_X(t) \leq \rho_{X,0}(t) \end{cases} \qquad (2)$$

where X may correspond to either the nOe or the S^2 term, N_{rep} is the number of replicas in the ensemble, α_X is the force constant, which relates to how strongly the restraint is applied, and

$$\rho_X(t) = \frac{1}{N_X} \sum_{k=1}^{N_X} (X_k^{exp} - X_k^{calc})^2 \qquad (3)$$

where N_X is the number of applied restraints and

$$\rho_{X,0}(t) = \min_{0 \le \tau \le t} \rho_X(\tau) \qquad (4)$$

In this implementation the back-calculated values are forced to approach the experimental ones by requiring them not to be worse than the best agreement previously achieved. Through the course of the simulation, thermal fluctuations decrease $\rho_{X,0}(t)$, eventually leading to agreement with the experimental values (Best and Vendruscolo 2004; Paci and Karplus, 1999).

The procedure for calculating X_k^{calc} varies with the type of restraint employed and includes a form of ensemble averaging appropriate to the particular observable. For nOes of proteins in the native state $< r^{-3} >^{-1/3}$ averaging was used since internal motions are assumed to be faster than the overall tumbling of the molecule (Lindorff-Larsen et al. 2005; Neuhaus and Williamson 2000). Thus, the interatomic distances are calculated according to

$$d_i^{calc} = \left(\frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} r_{k,i}^{-3} \right)^{-1/3} \qquad (5)$$

where $r_{k,i}$ refers to distance i of replica k.

For the S^2 restraint, X_k^{calc} is calculated using the Lipari–Szabo approximation (Lipari and Szabo 1982; Best and Vendruscolo 2004; Henry and Szabo 1985)

$$S_k^{2calc} = \frac{3}{2(r_k^{eff})^4} \left(\sum_{i=1}^{3} \sum_{j=1}^{3} \left[\frac{1}{N_{rep}} \sum_{l=1}^{N_{rep}} r_{i,k,l} r_{j,k,l} \right]^2 - 1 \right) \qquad (6)$$

where $r_{i,k}$ is the ith Cartesian component of bond vector k. The simplifying assumption is made that the bond vector does not change length during the molecular dynamics simulation, and an effective bond length of r_k^{eff} is used. This approximation is enforced through the use of the SHAKE algorithm (Ryckaert et al. 1977).

Using this procedure, the nOe distances and S^2 order parameters back-calculated from the reference simulation were employed as restraints in simulated annealing cycle simulations. These simulations were performed using the CHARMM22 force field, but were

carried out in vacuo, rather than in explicit solvent as was done for the reference simulation. The number of replicas used for the ensemble-averaging was systematically varied, as was the number of nOe restraints. In one set of simulations, exact distances rather than distance bins were used for the nOe restraints. All simulations contain 100 annealing cycles, except for the 16 replica simulations, which contain 50. This number of cycles proved sufficient for convergence, as measured by the observables considered for cross-validation (data not shown). The annealing cycles are composed of the following steps: (1) 20 ps at 298 K with $\alpha = \alpha_{high}$, (2) 10 ps for every 25 K up to 498 K with $\alpha = \alpha_{low}$, (3) 10 ps at 498 K with $\alpha = \alpha_{low}$, (4) 10 ps for every 25 K down to 348 K with $\alpha = \alpha_{low}$, (5) 10 ps for every 10 K down to 298 K with $\alpha = \alpha_{high}$, (6) 20 ps at 298 K with $\alpha = \alpha_{high}$, at the end of which the atomic coordinates are collected. Thus, the total simulation time for each replica is 230 ps per cycle or 23 ns for the 100 cycles. The force constant α that is applied to the restraints is lowered at higher temperatures in order to increase sampling; Table 2 lists the combinations of force constants that were applied to the different simulations.

Validation

To assess the similarity between the reference ensemble and the restrained ensembles, a variety of parameters were investigated.

Table 2 List of the force constants employed in the restrained simulations

Simulation	α_{nOe} (kcal/mol/Å⁴)		α_{S^2} (kcal/mol)	
	α_{high}	α_{low}	α_{high}	α_{low}
$N1$	$5 * 10^7$	$5 * 10^6$	–	–
$N2$	$8 * 10^7$	$8 * 10^6$	–	–
$N4$	$8 * 10^7$	$8 * 10^6$	–	–
$N8$	$8 * 10^7$	$8 * 10^6$	–	–
N_d1	$8 * 10^7$	$8 * 10^6$	–	–
N_d2	$8 * 10^7$	$8 * 10^6$	–	–
N_d4	$8 * 10^7$	$8 * 10^6$	–	–
N_d8	$8 * 10^7$	$8 * 10^6$	–	–
N_e1	$3 * 10^6$	$3 * 10^5$	–	–
N_e2	$8 * 10^6$	$8 * 10^5$	–	–
N_e4	$8 * 10^6$	$8 * 10^5$	–	–
N_e8	$4 * 10^6$	$4 * 10^5$	–	–
$NS2$	$1 * 10^7$	$1 * 10^6$	$1 * 10^7$	$1 * 10^6$
$NS4$	$1 * 10^7$	$1 * 10^6$	$1 * 10^7$	$1 * 10^6$
$NS8$	$1 * 10^7$	$1 * 10^6$	$3 * 10^7$	$3 * 10^6$
$NS16$	$1 * 10^7$	$1 * 10^6$	$1 * 10^7$	$1 * 10^7$
N_dS8	$8 * 10^6$	$8 * 10^5$	$1 * 10^7$	$1 * 10^6$
N_eS8	$5 * 10^6$	$5 * 10^5$	$1 * 10^7$	$1 * 10^6$
MUMO8	$3 * 10^7$	$3 * 10^6$	$3 * 10^7$	$3 * 10^6$
MUMO16	$3 * 10^7$	$3 * 10^6$	$3 * 10^7$	$3 * 10^7$

S^2 order parameters

To calculate the S^2 order parameters, the members of the restrained ensembles were first superimposed with a least-squares fit using backbone heavy atoms of regular secondary structures (residues 2–7, 12–16, 23–24, 41–45, 48–49, and 66–71) (Nederveen and Bonvin 2005). The S^2 order parameters were then obtained by the application of Eq. (6).

^3J-Couplings

Backbone and side-chain ^3J-couplings, which give information on dihedral angles, were calculated using the corresponding Karplus equation

$$^3J = A \cos^2(\theta + \delta) + B \cos(\theta + \delta) + C \tag{7}$$

where known values of A, B, C, and δ were used and where θ is the intervening dihedral angle (Karplus 1963). Five different backbone couplings ($^3J_{C_O-C_O}$, $^3J_{C_O-H_\alpha}$, $^3J_{H_N-C_\beta}$, $^3J_{H_N-C_O}$, $^3J_{H_N-H_\alpha}$) were considered for all applicable residues. $^3J_{N-C_\gamma}$ and $^3J_{C_O-C_\gamma}$ side-chain couplings were considered for threonine, valine, and isoleucine.

Hydrogen bond ^{h3}J-couplings

The same magnetization transfer mechanism that occurs across covalent bonds and is measured by ^3J-couplings can also be observed across hydrogen bonds in proteins (Cordier and Grzesiek 1999). The size of these $^{h3}J_{NC'}$-couplings depends crucially on the geometry of the hydrogen bond (Grzesiek et al. 2004). Using quantum mechanical methods, the hydrogen bond $^{h3}J_{NC'}$-couplings were parameterized as (Barfield 2002)

$$^{h3}J_{NC'} = -360 e^{-3.2r_{HO}} \cos^2 \theta + 0.04 \tag{8}$$

where r_{HO} is the length between the oxygen and the hydrogen involved in the hydrogen bond and where θ is the angle between the bond vector of the carbonyl group and the bond vector of the hydrogen bond.

Residual dipolar couplings

RDCs contain long-range information on the relative orientations of bond vectors with respect to an external magnetic field (Bax 2003; de Alba and Tjandra 2002). RDCs for the reference simulation for N–H_N, H_N–C_O, N–C_O, C_α–H_α, C_α–C_β, and C_α–C_O bond vectors were computed using PALES (Zweckstetter and Bax 2000)

with an approximation to the alignment tensor that considers only steric effects. RDCs for the restrained ensembles were back-calculated by fitting a single, effective alignment tensor by minimizing the Q-factor.

$$Q = \frac{\sqrt{\sum(\text{RDC}_{calc} - \text{RDC}_{exp})^2}}{\sqrt{\sum(\text{RDC}_{exp})^2}} \tag{9}$$

Interatomic distance distributions

Interatomic distances yielding nOe averages below 6 Å were grouped into a restrained set, which contains the atom pairs used as restraints, and an unrestrained set, which contains all other atom pairs. For every hydrogen–hydrogen pair, the distance calculated for each structure was binned into a histogram with a bin size of 0.4 Å and normalized by the number of structures in the restrained ensemble.

The similarity between the histograms of the reference and the restrained ensembles was computed for a given distance pair as

$$s_k = \sum_i \left| p_{i,k}^{ref} - p_{i,k}^{restrained} \right| \tag{10}$$

where the sum over i runs over all bins of the histogram, and where $p_{i,k}^{ref}$ and $p_{i,k}^{restrained}$ correspond to the normalized frequency of finding a particular distance in that particular bin for the reference and the restrained simulations, respectively. Thus, a low value for s_k represents very similar distributions while a high value for s_k represents dissimilar distributions. The similarity measures s_k were summed for all restrained and unrestrained distances separately and then averaged by the number of distances, N_{dist}, considered

$$S = \frac{1}{N_{dist}} \sum_{k=1}^{N_{dist}} s_k \tag{11}$$

Thus, we arrive at a value S for the overall distance distribution similarity of the restrained and unrestrained data sets.

Rotamer distributions

Distributions of rotamer states were examined for those side-chains for which S^2 order parameters were calculated. In total, 46 side-chain rotamers were investigated. The following dihedral angles were considered: isoleucine $\chi 1$ and $\chi 2$, leucine $\chi 1$ and $\chi 2$, methionine $\chi 1$, $\chi 2$, and $\chi 3$, threonine $\chi 1$, and valine $\chi 1$. Dihedral angles were calculated for each structure in

the reference and restrained ensembles, and then placed into a histogram with a bin width of 10°. The similarity between dihedral angle histograms was calculated using Eq. (10) and an overall distribution similarity, S, was derived using Eq. (11).

NH bond vector distributions

The orientations of the backbone amide bond vectors were considered by first aligning the structures and then binning the θ and φ angles in the spherical coordinate system into a two dimensional histogram using bin widths of 5°. The similarity to the reference was computed using Eq. (10) and an overall similarity measure, S, was calculated according to Eq. (11).

RMSD between average structures

CHARMM was used to calculate the geometric average structure of the reference and each restrained ensemble. Then the RMSD was calculated for regular secondary structure elements between the averaged structures of the reference and each restrained ensemble.

Mean pairwise RMSD between ensembles

The pairwise root mean square deviation (RMSD) was calculated within all frames in the reference and each restrained ensemble using CHARMM. The RMSDs were then averaged to obtain the mean pairwise RMSD for each simulation.

Per residue fluctuations

An average structure was calculated for the reference and for each restrained ensemble. Then the RMSD of the C_α at each residue was computed between each structure in the restrained ensemble and its average structure. In order to obtain the per residue fluctuations, F_i^{calc}, at residue i, the RMSDs were averaged across the restrained ensemble. The overall agreement of the per residue fluctuations, F, between the reference and restrained ensembles was calculated by averaging across all N_{C_α} residues

$$F = \frac{1}{N_{C_\alpha}} \sum_{i=1}^{N_{C_\alpha}} |F_i^{calc} - F_i^{ref}| \tag{12}$$

where F_i^{ref} is the per residue fluctuation at residue i in the reference ensemble.

Simulation with experimental NMR data

The starting structure, non-bonded interactions and solvation scheme for the simulation with experimental NMR data was identical to that of the ubiquitin reference simulation described above. Sixteen replicas were simulated with 2,663 nOe distances restraints and 141 backbone and sidechain S^2 order parameter restraints (Cornilescu et al. 1998; Chang and Tjandra 2005; Lee et al. 1999). The experimental restraints were ensemble-averaged according to the MUMO protocol with the S^2 order parameter restraints applied over all sixteen replicas and the nOe distance restraints applied to overlapping pairs of replicas.

A difference to the synthetic case is that many of the nOe distances are ambiguous. Thus $r_{k,i}$ in Eq. (5) is itself an average of all possible distance pairs contributing to the ambiguous nOe signal:

$$r_{k,i} = \left(\sum_{l=1}^{N_{ambig}} r_{k,i,l}^{-6} \right)^{-1/6} \tag{13}$$

where the sum runs over all N_{ambig} possible ambiguous distances for a particular member of the replica ensemble. The averaging between molecules and the calculation of the S^2 order parameters is identical to the synthetic case.

In the annealing cycles, which were similar to the synthetic case, the temperature was raised from 298 K to 598 K while decreasing the force constant for the nOes from 10^7 to 10^6 and for the S^2 order parameters from $5 * 10^7$ to $5 * 10^6$. A slightly higher maximum temperature was used than in the sythetic annealing cycles in order to speed up sampling, which is slower in explicit water. After an equilibration phase at 298 K, the final structure was recorded. The simulation time for each cycle was 260 ps per replica. Nine cycles were performed, yielding a simulation time of 2.34 ns per replica. After pooling, an ensemble containing 144 members resulted.

Validation with experimental backbone RDCs (Cornilescu et al. 1998) was performed by fitting a single alignment tensor to minimize the unnormalized Q-factors. Hydrogen bond J-couplings were calculated according to Eq. (8), and correlation coefficients to the experimental data were calculated (Cordier and Grzesiek 2002). Also the RMSD of the geometric average of all 144 structures to the average structure of the RDC-restrained NMR ensemble (Cornilescu et al. 1998) was calculated.

Results

Generation of the reference ensemble with unrestrained molecular dynamics simulations

We generated a reference native state ensemble of ubiquitin using an unrestrained molecular dynamics simulation. The RMSD from the starting structure, which is the minimized crystal structure (Vijay-Kumar et al. 1987), remains below 1.5 Å at all times during the simulation. The number and types of nOe distances and S^2 order parameters back-calculated from the reference ensemble are similar to those available experimentally for native state structure determination. This is necessarily the case for the nOe restraints since the proportion of short, medium, and long-range sequence distances was purposefully chosen to be similar to the experimental case. Figure 2 demonstrates that also the S^2 order parameters fall within the range of values typically observed for native state proteins. Backbone order parameters generally vary between 0.8 and 0.9, whereas the side-chain order parameters are lower, which is also observed for the reference simulation. The exception is the C-terminus, which has also been found experimentally to be very flexible (Tjandra et al. 1995; Wang et al. 2003). Nonetheless, any force field contains approximations (Karplus and McCammon 2002), and hence the reference ensemble may not fully represent the true structure and dynamics of native ubiquitin on the nanosecond timescale. It does, however, represent a model for a possible Boltzmann ensemble of a native state.

Reproduction of the reference ensemble using annealing simulations with ensemble-averaged restraints

We aim here at establishing a computational procedure that uses ensemble-averaged restraints to reproduce
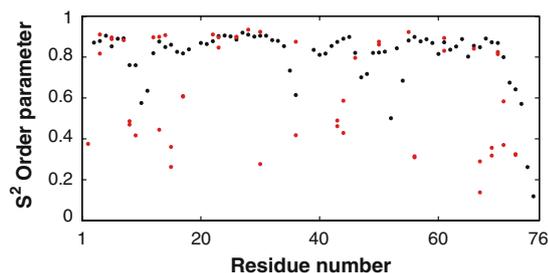
Fig. 2 S^2 order parameters back-calculated from the reference simulation. The backbone order parameters are shown in black and the side-chain order parameters are shown in red

the reference ensemble as accurately as possible. In one series of simulations, we tested the effect of imposing only nOe restraints on one (N1), two (N2), four (N4), and eight (N8) replicas of the molecule at each time step. In a second series, we applied both nOe and S^2 restraints to two (NS2), four (NS4), eight (NS8) and sixteen (NS16) replicas. Additionally, we also doubled the number of distance restraints and used exact distances rather than bins. The N_d1, N_d2, N_d4, and N_d8 simulations contain double the number of nOe restraints for one, two, four, and eight replicas, respectively. Similarly, the N_e1, N_e2, N_e4, and N_e8 simulations contain exact nOe restraints (the upper and lower bounds of the distance bins are identical) for one, two, four, and eight replicas, respectively. And, the N_dS8 ensemble was generated using eight replicas with the regular S^2 restraints, but with double the nOe restraints. The N_eS8 ensemble was created using the regular S^2 restraints in combination with exact nOe restraints. Finally, we introduced and applied a restraining algorithm (MUMO) in which nOes were restrained in a linked pairwise fashion, while the S^2 restraints were applied to either eight (MUMO8) or sixteen replicas (MUMO16).

The restraints back-calculated from the reference ensemble were imposed in ensemble-averaged annealing cycle simulations, where each annealing cycle yields one replica ensemble. The RMSD of the restrained NMR parameters from those in the reference simulation are summarized in Table 3. These values reflect how much, on average, individual restraints are violated for each replica ensemble. The RMSDs demonstrate that at each time step the restraints are closely reproduced, indicating the self-consistency of the procedure used to impose the restraints. In Table 3, the RMSDs for both the nOes and the S^2 order parameters tend to decrease as the number of replicas is increased. Since the number of degrees of freedom is proportional to the number of replicas (Bürgi et al. 2001), this result is expected. Subsequently, the quality of the determined ensembles is instead assessed by validation using unrestrained observables (Bonvin and Brunger 1995; Brünger 1992; Brünger et al. 1993).

Although the restraints are satisfied for each replica ensemble they are not necessarily satisfied for the overall restrained ensemble, in which all replica ensembles are pooled together. The restraining algorithm only enforces the restraints at each individual time step and not across time steps or annealing cycles. Nonetheless, in an N_{rep} replicas simulation with N_{cyc} cycles the nOe restraints must necessarily be satisfied when they are applied without a lower bound. At each

Table 3 Average deviations per cycle (or replica ensemble) of the nOe and the S^2 restraints

Simulation	nOe RMSD (Å)	S^2 RMSD
$N1$	0.080 ± 0.002	–
$N2$	0.041 ± 0.002	–
$N4$	0.031 ± 0.002	–
$N8$	0.029 ± 0.001	–
N_d1	0.103 ± 0.002	–
N_d2	0.063 ± 0.004	–
N_d4	0.047 ± 0.003	–
N_d8	0.040 ± 0.002	–
N_e1	0.427 ± 0.005	–
N_e2	0.205 ± 0.002	–
N_e4	0.174 ± 0.002	–
N_e8	0.213 ± 0.003	–
$NS2$	0.076 ± 0.003	0.058 ± 0.002
$NS4$	0.060 ± 0.002	0.046 ± 0.002
$NS8$	0.061 ± 0.004	0.031 ± 0.002
$NS16$	0.062 ± 0.002	0.049 ± 0.003
N_dS8	0.088 ± 0.003	0.045 ± 0.003
N_eS8	0.202 ± 0.002	0.047 ± 0.002
MUMO8	0.045 ± 0.002	0.032 ± 0.001
MUMO16	0.047 ± 0.002	0.032 ± 0.001

cycle, the back-calculated, average distances are forced to be below the upper bound, r_{upper}

$$\frac{1}{N_{\text{rep}}} \left(\sum_{j=1}^{N_{\text{rep}}} d_{i,j}^{-3} \right)^{-\frac{1}{3}} \leq r_{\text{upper}} \qquad (14)$$

where the cycle number i is in the range $1 \leq i \leq N_{\text{cyc}}$ and where $d_{i,j}$ is an interatomic distance for the jth replica in the ith cycle. Hence the overall distance measured, r, will also satisfy the upper bound, r_{upper}:

$$r = \frac{1}{N_{\text{rep}}N_{\text{cyc}}} \left(\sum_{i=1}^{N_{\text{cyc}}} \sum_{j=1}^{N_{\text{rep}}} d_{i,j}^{-3} \right)^{-\frac{1}{3}} \qquad (15)$$

and hence

$$r = \frac{1}{N_{\text{rep}}N_{\text{cyc}}} \left(\sum_{j=1}^{N_{\text{rep}}} d_{1,j}^{-3} + \cdots + \sum_{j=1}^{N_{\text{rep}}} d_{N_{\text{cyc}},j}^{-3} \right)^{-\frac{1}{3}} \qquad (16)$$

$$r \leq N_{cyc}^{-4/3} r_{\text{upper}} \qquad (17)$$

Since $0 < N_{\text{cyc}}^{-4/3} \leq 1$, the back-calculated distance across the whole restrained ensemble will always be smaller than the upper bound. Thus no distance restraint will be violated for the restrained ensemble if it is satisfied for each replica ensemble; note that this argument does not apply when there is a finite lower bound, as, for instance, when exact distances are used.

Any observable that is not a linear function of the geometries of the individual replicas may be satisfied for each of the replica ensembles, but not for the restrained ensemble. Since the S^2 order parameter does not depend linearly on the Cartesian bond vector components (see Eq. (6)), pooling effects may cause disagreements between the S^2 restraints and the values back-calculated from the restrained ensemble. Therefore we compare the backbone amide and side-chain methyl S^2 values back-calculated over the restrained ensembles to those of the reference ensemble in Fig. 3. In the $NS16$ case, several side-chain order parameters are lower in the restrained ensemble than in the reference ensemble (Fig. 3l). In fact, the correlation coefficient between the sidechain S^2 order parameters in the reference ensemble and in the $NS16$ ensemble drops to 0.89 from 0.97 in the $NS4$ case (Table 4).

It is also worth noting that the correlations for both the backbone and side-chain order parameters are lower for the $NS2$ than for the $NS4$ simulation (Fig. 3). Unlike in the $NS16$ case, the back-calculated order parameters are not exclusively too low, suggesting that issues other than the pooling of the replica ensembles to form the restrained ensemble are at work. A probable explanation is that two replicas are not sufficient to fully model bond vector dynamics. Since the entire range of the dynamical motions of a bond vector must be represented by the replicas at each cycle, there must be a minimum number of replicas necessary to enforce the S^2 order parameters (Best and Vendruscolo 2004). This means that the difficulties in satisfying the order parameters should be manifested also in the replica ensembles themselves. This is true as the $NS2$ simulations have the highest S^2 RMSD out of all the simulations with restrained order parameters (Table 3). Thus, we propose that attempts to satisfy low S^2 values imposed as restraints generate large forces at low replica numbers and hence an unrealistic degree of *frustration* (defined as "an inability to satisfy simultaneously all the inclinations of all the microscopic entities" (Mezard et al. 1987)) in the generated ensembles. As will be shown later, also validation with unrestrained observables indicates poor quality structures at low replica numbers. This effect, called here over-restraining (or underfitting), is due mainly to the use of too few replicas to fit the data.

Comparison of ensembles

In order to assess the different simulation procedures, we chose to compare the restrained ensembles, rather

Fig. 3 Comparison of the backbone (black) and side-chain (red) S^2 order parameters back-calculated for the reference and for the entire ensemble of restrained simulations. The panels refer to the following ensembles: $N1$ (**a**), $N2$ (**b**), $N4$ (**c**), $N8$ (**d**), N_d4 (**e**), N_d8 (**f**), N_e4 (**g**), N_e8 (**h**), $NS2$ (**i**), $NS4$ (**j**), $NS8$ (**k**), $NS16$ (**l**), N_dS8 (**m**), N_eS8 (**n**), MUMO8 (**o**), and MUMO16 (**p**)
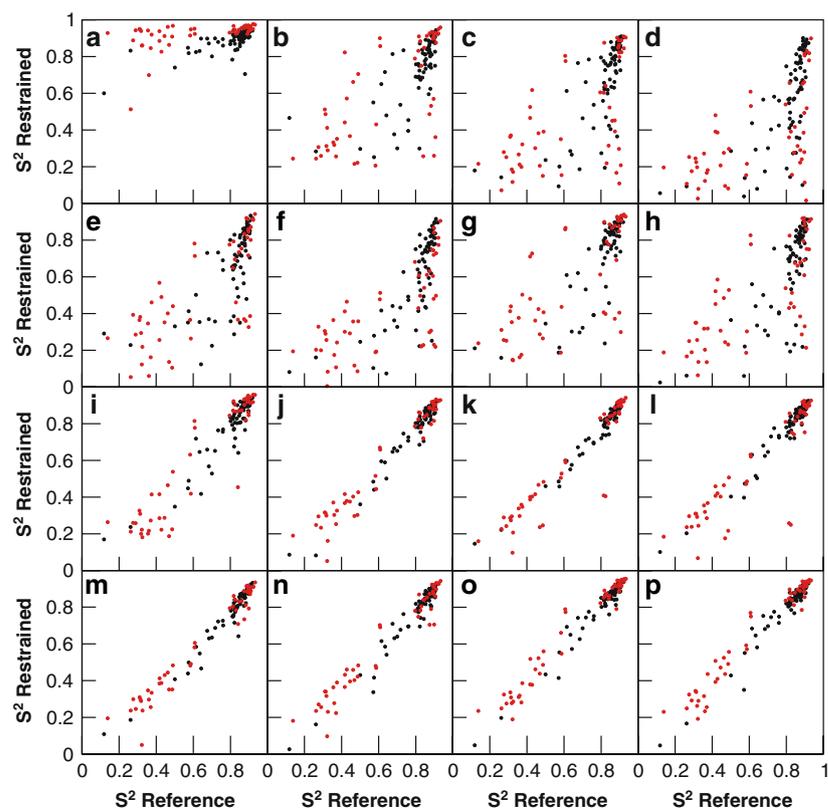
Table 4 Correlation coefficients between the S^2 order parameters of the reference ensemble and the various restrained ensembles analyzed in this work

Simulation	Amide S^2	Methyl S^2
$N1$	0.72	0.49
$N2$	0.69	0.64
$N4$	0.72	0.60
$N8$	0.67	0.56
N_d1	0.73	0.54
N_d2	0.74	0.63
N_d4	0.69	0.75
N_d8	0.75	0.68
N_e1	0.77	0.21
N_e2	0.82	0.70
N_e4	0.79	0.74
N_e8	0.73	0.70
$NS2$	0.95	0.94
$NS4$	0.98	0.97
$NS8$	0.98	0.94
$NS16$	0.97	0.89
N_dS8	0.97	0.98
N_eS8	0.97	0.97
MUMO8	0.98	0.98
MUMO16	0.97	0.98

than the replica ensembles, to the reference ensemble. Lindorff-Larsen et al. observed that cross-validation with the restrained ensembles yields better agreement than with any one of the individual replica ensembles

(Lindorff-Larsen et al. 2005). Ideally, we would simulate a large replica ensemble to obtain adequate sampling, but this is not feasible since we do not have enough experimental information to do so without overfitting. Therefore this study could also be considered a search for the optimal utilization of the limited experimental information available in reproducing the dynamics of proteins.

In order to determine the best way to reproduce the reference ensemble, we need to define a procedure for measuring the similarity between the reference and the restrained ensembles. Formally, two ensembles are identical if the same structures occur with the same relative probabilities. However, this is not a practical definition for our purposes.

A direct visual inspection of the ensembles (Fig. 4) already provides a quite accurate perspective on the quality of the reproduction of the structural heterogeneity of the reference ensemble. In order to obtain a more quantitative assessment of the reproduction, we back-calculate a variety of parameters, only some of which are experimentally accessible, from the restrained ensembles and compare them to the reference ensemble.

The accuracy in the determination of the average structure is measured by the average RMSD between the average structures of the restrained ensembles

Fig. 4 Fifty representative backbone traces for the following ensembles: (**a**) the reference, (**b**) *N1*, (**c**) *N2*, (**d**) *N4*, (**e**) *N8*, (**f**) *Nd4*, (**g**) *Nd8*, (**h**) *Ne4*, (**i**) *Ne8*, (**j**) *NS2*, (**k**) *NS4*, (**l**) *NS8*, (**m**) *NS16*, (**n**) *NdS8*, (**o**) *NeS8*, (**p**) MUMO8, and (**q**) MUMO16
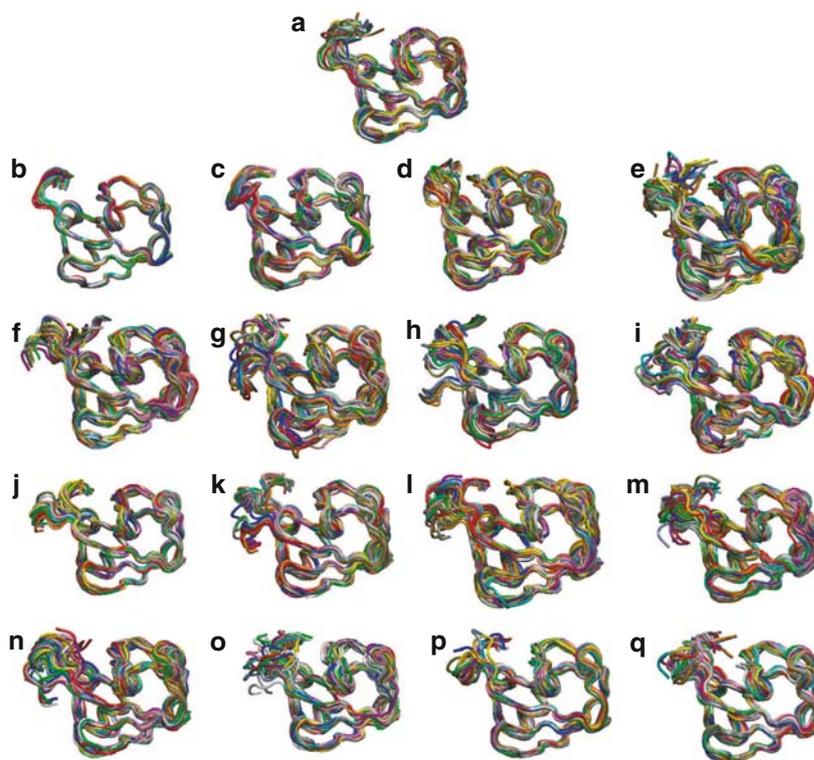
and the average structure of the reference ensemble (Table 5). The structural heterogeneity of the ensembles is assessed by measuring the RMSD between pairs of structures within the ensembles, shown in Table 5. Low intra-ensemble RMSDs indicate highly similar members in the structural ensembles. The structural fluctuations in different regions of the polypeptide chain are described by considering the per residue fluctuations (Fig. 5). This type of plot shows that the use of S^2 order parameters as restraints has a very significant impact on the quality of the reproduction of the structural fluctuations, by considerably reducing the problem of overfitting. This conclusion is supported by the analysis of the S^2 order parameters themselves (Fig. 3 and Table 4), which shows that excessive mobility is generated when nOe distances are the only restraints used.

NMR parameters not used as restraints can also be employed for validation. Table 6 lists the RDC Q-factors obtained with the different simulations. With the number and type of restraints held constant, the Q-factors tend to increase, signifying worse agreement with the reference RDCs, as the number of replicas is increased. Other NMR parameters analyzed include side-chain, backbone, and hydrogen bond J-couplings. The agreement with the reference simulation is quantified by the correlation coefficients listed in Table 6.

Table 5 RMSD of the average structures of the reference and the restrained ensembles; average difference of the per residue fluctuations; difference in average pairwise RMSD between the restrained and the reference ensembles

Simulation	RMSD of average structures (Å)	Average difference in per residue fluctuations (Å)	Difference in average pairwise RMSD (Å)
$N1$	0.46	0.39	−0.50 ± 0.05
$N2$	0.64	0.21	0.07 ± 0.33
$N4$	0.75	0.30	0.27 ± 0.32
$N8$	0.75	0.73	0.44 ± 0.33
N_d1	0.52	0.45	−0.50 ± 0.03
N_d2	0.59	0.31	0.10 ± 0.45
N_d4	0.56	0.31	0.16 ± 0.23
N_d8	0.69	0.50	0.38 ± 0.32
N_e1	0.43	0.47	−0.51 ± 0.03
N_e2	0.34	0.20	−0.16 ± 0.21
N_e4	0.33	0.32	0.22 ± 0.31
N_e8	0.46	0.42	0.35 ± 0.29
$NS2$	0.62	0.19	−0.13 ± 0.13
$NS4$	0.71	0.25	0.29 ± 0.30
$NS8$	0.80	0.20	0.10 ± 0.28
$NS16$	0.87	0.17	0.05 ± 0.27
N_dS8	0.76	0.17	0.11 ± 0.27
N_eS8	0.46	0.20	0.10 ± 0.23
MUMO8	0.50	0.17	−0.12 ± 0.20
MUMO16	0.56	0.13	−0.09 ± 0.18

Using a reference ensemble also allows us access to distributions of properties, such as hydrogen-hydrogen distances, NH bond vector orientations, and rotamer

Fig. 5 Per residue fluctuations for (**a**) the nOe-only simulations, for (**b**) the nOe plus S^2 simulations, for (**c**) the double and exact nOe-only simulations, and for (**d**) the double, exact, and pairwise nOe plus S^2 simulations
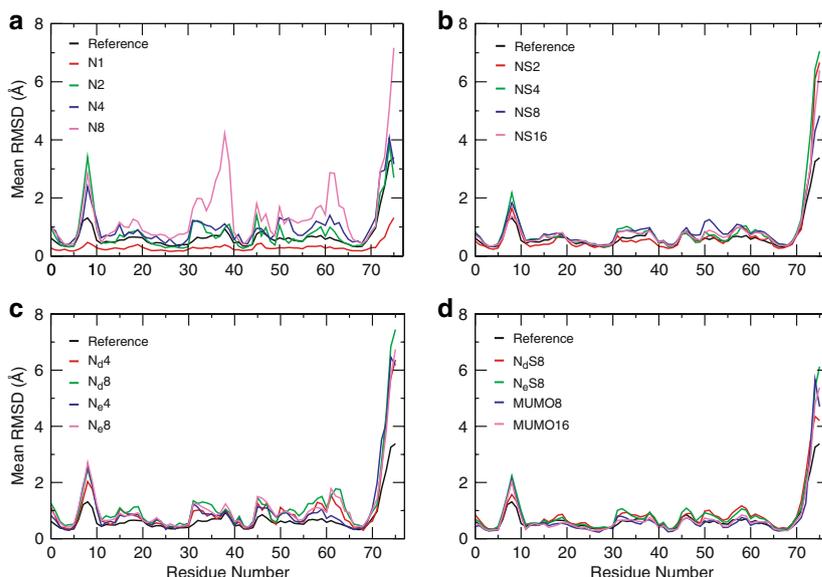
states. For different types of restrained simulations we show representative distance distributions (Fig. 6), rotamer distributions (Fig. 7), and angular NH bond vector distributions (Fig. 8). We observe that simulations with low replica numbers do not detect the range of substates present in the reference. By contrast, simulations with higher replica numbers generally identify the relevant substates, but do not always enumerate their correct relative probabilities. On some occasions, conformational substates that are not populated in the reference ensemble are occupied in the restrained ensembles, a result that indicates the presence of overfitting. The S-value, an overall measure for the goodness-of-fit of the distributions, is presented in Table 7.

Discussion

The onset of overfitting occurs beyond two replicas in the nOe-only restrained simulations

Since the use of a larger number of replicas implies increasing the number of degrees of freedom, the restraints are often better satisfied at higher replica numbers while agreement with non-restrained parameters deteriorates (overfitting). Previous studies of nOe-only simulations have indicated that overfitting occurs when using more than two replicas (Bonvin et al. 1994; Bonvin and Brunger 1995; Fennen et al. 1995). Our data confirm this early onset of overfitting when ensemble simulations are restrained only by nOes.

As shown in Table 5, the average structure of ubiquitin is best reproduced when only a single replica

Table 6 Correlation coefficients for various unrestrained NMR observables (the backbone, side-chain, and hydrogen bond J-couplings, and the RDC Q-factors)

Simulation	Backbone J-coupling	Side-chain J-coupling	Hydrogen Bond J-coupling	RDC Q-factor
$N1$	0.97	0.97	0.36	0.38
$N2$	0.99	0.96	0.59	0.35
$N4$	0.98	0.94	0.51	0.39
$N8$	0.98	0.92	0.39	0.44
N_d1	0.97	0.97	0.68	0.34
N_d2	0.98	0.98	0.73	0.35
N_d4	0.98	0.99	0.61	0.37
N_d8	0.98	0.98	0.55	0.37
N_e1	0.98	0.97	0.68	0.33
N_e2	0.99	0.98	0.90	0.23
N_e4	0.99	0.99	0.82	0.23
N_e8	0.98	0.98	0.58	0.33
$NS2$	0.99	0.98	0.73	0.36
$NS4$	0.99	0.98	0.78	0.33
$NS8$	0.98	0.98	0.71	0.41
$NS16$	0.98	0.94	0.71	0.40
N_dS8	0.99	0.98	0.78	0.38
N_eS8	0.99	0.99	0.82	0.24
MUMO8	0.99	0.98	0.74	0.29
MUMO16	0.99	0.98	0.75	0.27

In all cases the statistical errors on the correlation coefficients are estimated to be about 0.02

is used ($N1$). The agreement between the average structure of the nOe-restrained ensembles and the reference ensemble decreases as the number of replicas is increased. This result is not surprising as the restraints represent the average properties of the reference ensemble and must be satisfied for each member of the $N1$ ensemble. Thus each structure in $N1$ represents the average structure of the reference

Fig. 6 Example of a distance distribution from a restrained hydrogen–hydrogen pair. The reference distribution is shown in black, whereas the restrained distribution is shown in red for the ensembles (a) $N1$, (b) $N2$, (c) $N4$, (d) $N8$, (e) N_d4, (f) N_d8, (g) N_e4, (h) N_e8, (i) $NS2$, (j) $NS4$, (k) $NS8$, (l) $NS16$, (m) N_dS8, (n) N_eS8, (o) MUMO8, and (p) MUMO16
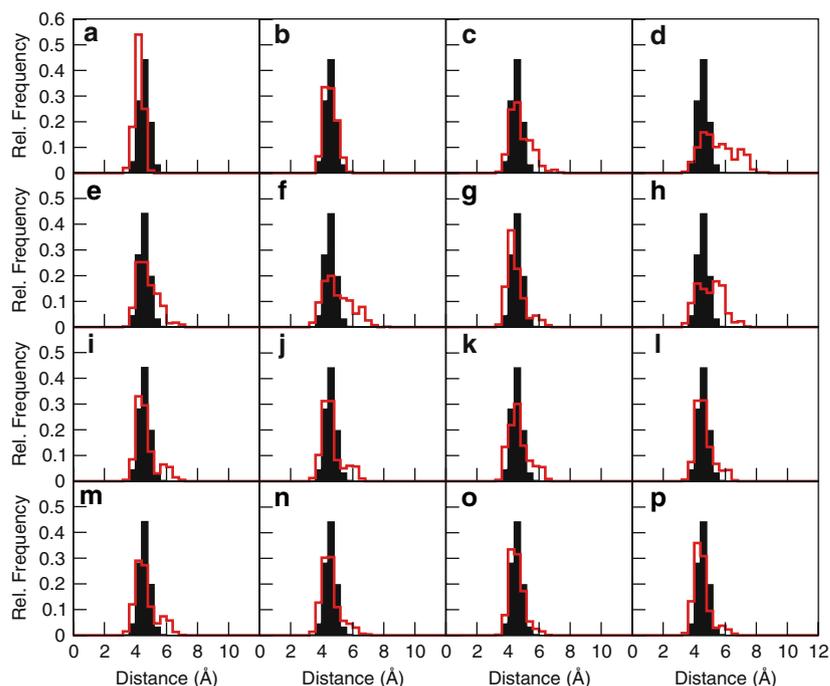
Fig. 7 Rotamer distributions for the $\chi2$ rotamer on residue 36 for the restrained simulations. The reference distribution is shown in black, whereas the restrained distribution is shown in red. The $N1$ simulation is shown in (a), $N2$ in (b), $N4$ in (c), $N8$ in (d), N_d4 in (e), N_d8 in (f), N_e4 in (g), N_e8 in (h), $NS2$ in (i), $NS4$ in (j), $NS8$ in (k), $NS16$ in (l), N_dS8 in (m), N_eS8 in (n), MUMO8 in (o), MUMO16 in (p)
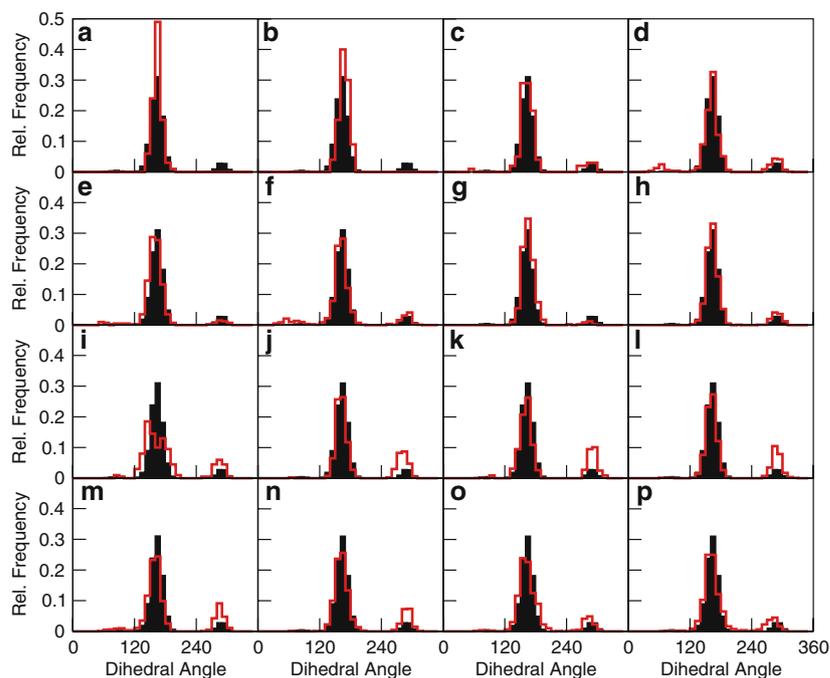
ensemble. By contrast, when two or more replicas are used, individual replicas may deviate significantly from the restraints and thus from the average structure.

This observation is supported by the analysis of the average pairwise RMSD, listed in Table 5. For the $N1$ simulation, the average pairwise RMSD is substantially smaller than for the reference simulation, implying that the N1 simulation does not reproduce the dynamic

fluctuations about the average structure that are present in the reference ensemble. For the $N2$ ensemble, the average pairwise RMSD (0.88 Å) is slightly larger than that of the reference (0.81 Å) but within the acceptable range given by the standard deviation. As the replica number is increased even further, the pairwise RMSD grows still larger. This result suggests that at higher replica numbers, the nOe-only ensembles are too

Fig. 8 NH bond vector distributions for residue Asp32 projected onto a unit sphere. The reference distribution is shown in red, while the restrained distribution is shown in blue. (**a**) the *NS*16 ensemble; (**b**) the MUMO16 ensemble
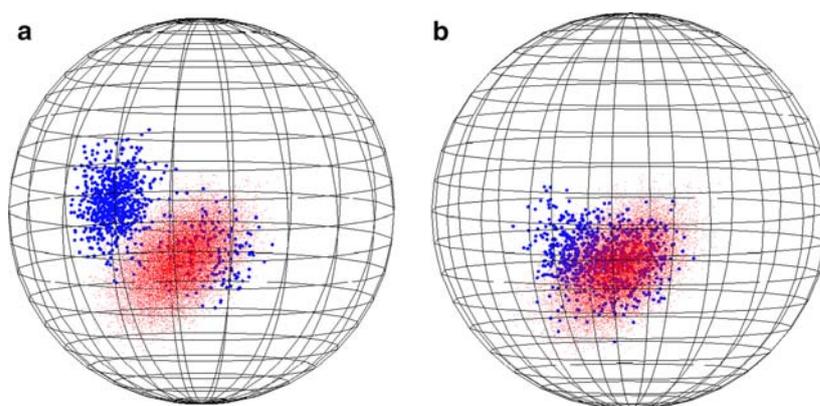
Table 7 Similarity scores *S* between the reference and the restrained ensembles; results are reported for distributions of the restrained and unrestrained distances, the sidechain rotamers, and the angular components of the *NH* bond vectors

Simulation	Restrained distance distributions	Unrestrained distance distributions	Sidechain rotamer distributions	Angular *NH* bond vector distributions
$N1$	0.87	0.89	0.73	1.17
$N2$	0.51	0.53	0.65	0.94
$N4$	0.44	0.47	0.53	1.05
$N8$	0.47	0.49	0.47	1.08
N_d1	0.91	0.94	0.83	1.14
N_d2	0.51	0.52	0.42	1.00
N_d4	0.39	0.41	0.33	0.97
N_d8	0.41	0.43	0.36	0.97
N_e1	1.00	1.02	0.98	1.15
N_e2	0.56	0.54	0.57	0.87
N_e4	0.34	0.34	0.30	0.76
N_e8	0.33	0.35	0.25	0.70
$NS2$	0.47	0.49	0.44	1.01
$NS4$	0.43	0.44	0.33	0.93
$NS8$	0.47	0.49	0.40	1.11
$NS16$	0.44	0.46	0.34	0.99
N_dS8	0.40	0.41	0.29	0.90
N_eS8	0.30	0.33	0.24	0.70
MUMO8	0.49	0.50	0.40	0.82
MUMO16	0.46	0.47	0.37	0.81

The statistical errors on the *S* scores are estimated to be less than 0.02 in all cases

heterogeneous. As the replica number is increased, the restraints can in principle still be satisfied, even if one replica adopts an extended, or potentially even unfolded, conformation (Zagrovic and van Gunsteren 2006). This problem is particularly severe because the $< r^{-3} >^{-1/3}$ averaging associated with nOes is especially insensitive to large distances.

The *N*2 ensemble provides the best fit to the reference ensemble in terms of the per residue fluctuations (Fig. 5a and Table 5). Also, the back-calculated S^2 order parameters show that the *N*1 ensemble is too rigid since especially for the sidechains but also for the backbone the order parameters in the restrained ensemble are higher than in the reference ensemble (Fig. 3a). As the replica number is increased (Fig. 3b–d), more and more data points fall below the diagonal, implying that the ensembles are too heterogeneous. Table 4 lists the correlation coefficients for the order parameters. The best correlation for the side-chain order parameters occurs at two replicas, while the best agreement for the backbone is found at one and four replicas. Nevertheless, the correlation for the order parameters for all nOe-only restrained simulations is fairly low, showing that nOe-only simulations are not very successful in capturing the fast (ps–ns) dynamics of native state proteins.

Validation with other unrestrained observables also illustrates that overfitting occurs already at low replica numbers. Table 6 demonstrates that backbone and hydrogen bond *J*-couplings are best reproduced for the *N*2 simulation, whereas the backbone *J*-couplings show optimal agreement for the *N*1 simulations. Increasing the replica number beyond two leads to worse agreement. Also the agreement with the RDCs tends to deteriorate as the replica number is increased; the best *Q*-factor is achieved in the *N*2 simulation.

One important aspect of using a reference ensemble and synthetic NMR data is that it becomes possible to compare directly distributions of distances and rotamer states. Representative distance and rotamer distributions shown in Fig. 6a–d and Fig. 7a–d respectively provide a visual illustration of the underfitting and overfitting effects. In the $N1$ simulations, the distributions are too narrow. This is especially true for the $\chi2$ distribution of Ile36 shown in Fig. 7a. The less populated rotamer state at 300° of the reference ensemble is not populated in the $N1$ ensemble due to underfitting. When eight replicas are used (Fig. 7d), this rotamer state is populated, but so is an additional state at 60°, which does not exist in the reference ensemble; this overfitting effect can also be seen in the distance distributions (Fig. 6a–d). As the replica number is increased, the distributions become broader. In particular, long-distance tails tend to develop, a situation particularly prevalent for the $N8$ ensemble in Fig. 6d.

Results of other properties that we investigated, such as the per residue fluctuations and S^2 order parameters, support the conclusion that the $N1$ ensemble does not display enough structural heterogeneity, while the opposite is found for the $N4$ and $N8$ ensembles. Therefore the $N1$ ensemble is over-restrained (underfitted), whereas the $N4$ and $N8$ ensembles are under-restrained (overfitted). Hence, consistent with the results by Bonvin and Brünger (Bonvin and Brunger 1995), the $N2$ ensemble seems to display the best possible balance between over- and underfitting in the case of nOe-only restrained simulations.

Doubling the number of nOe restraints or using exact distances delays overfitting

The extent of overfitting and underfitting depends on the balance between the number of degrees of freedom and the amount of experimental information available. As more information is added, increasingly many replicas are necessary to avoid underfitting, while it should be possible to delay the onset of overfitting at high replica numbers. We tested the influence of the information content of the restraints by doubling the number of nOe restraints in one series of simulations and using exact bins for the distance restraints in another series.

The double and exact restraint simulations were carried out with one, two, four, and eight replicas. Virtually all validation measures considered in this study improve when augmenting the available experimental data. Lower Q-factors are observed (Table 6), particularly for the simulations with the exact nOe restraints. Also the order parameters (Table 4) and

J-couplings (Table 6) improve. The distance and rotamer distributions (Table 7) become more similar to the reference ensemble. Not only do these parameters improve when compared to the nOe-only (N) simulations, better agreement is now more often found at higher replica numbers.

The improved agreement with non-restrained parameters and distributions of observables confirms that overfitting is in fact due to a lack of information about the system. The number of replicas at which overfitting occurs increases as the amount of information about the system is increased. Furthermore, using exact bins tends to yield better results than doubling the number of restraints. This result indicates that significant information is lost when placing distances into large bins. Also, doubling the number of restraints is not equivalent to doubling the information as many restraints are redundant, especially when large distance bins are used. These results suggest that bins should be tightened as far as the experimental error allows (Schneider et al. 1999).

Addition of the S^2 restraints to the nOe restraints delays overfitting

Although doubling the number of restraints or using exact bins delays the overfitting problem, we do not suggest this as a viable solution since it would require a currently infeasible increase in the amount and precision of experimental nOe data. Encouragingly, however, the addition of the S^2 restraints, which are measurable experimentally, also delays the onset of overfitting.

Ensembles in the nOe-only (N) simulations are structurally too heterogeneous when using two or more replicas. Ensembles from the nOe plus S^2 (NS) simulations are only slightly too heterogeneous when four or more replicas are used. Also the per residue fluctuations (Fig. 5) are well reproduced at high replica numbers. Since the S^2 restraint causes a delay of the ensemble broadening effect at high replica number seen in N simulations, we also expect cross-validation with unrestrained NMR observables to yield higher optimal replica numbers. The agreement with side-chain, backbone, and hydrogen bond J-couplings improves at higher replica numbers compared to the N simulations (Table 6). Interestingly, when considering distributions of properties, such as distances, rotamer states, and NH bond vector orientations, the best agreement is found for the $NS4$ simulation. The $NS8$ simulation is significantly worse, but agreement improves again for the $NS16$ ensemble. However, the $NS4$ simulation has a higher average pairwise RMSD

than both the NS8 and NS16 simulations. These results indicate that both underfitting and overfitting effects for the nOe and the S^2 restraints are at play.

Finally, it is worth pointing out that although many observables require higher replica numbers, some measures like the RMSD of the average structures and the RDC Q-factors remain optimal when two or four replicas are used. As these are observables that depend strongly on the average structure (although dynamics also play a role in the RDCs), this result suggests that although the higher replica NS simulations reproduce the dynamics very well, they do so at the expense of the average structure. Increasing the replica number to a point where the S^2 restraints can be satisfied with high precision has the effect that the nOe restraints are no longer effective in maintaining the average structure.

Tradeoff between overfitting and underfitting

The dependence of the optimal replica number on the observable under consideration complicates our goal of determining the appropriate types of restraints and the optimal number of replicas to use in order to produce ensembles that simultaneously model the structure and dynamics of native states. In nOe-only (N) simulations low replica numbers generate the best possible cross-validation, whereas higher numbers (eight or sixteen) create optimal agreement for many but not all of the observables when using both nOe and S^2 restraints (NS). Now we address which types of restraints yield the best overall agreement with the reference ensemble.

Rotamer and distance distributions, which are sensitive to the variability of structures within the restrained ensemble, are best reproduced in the NS simulations, as shown in Table 7. Also hydrogen bond J-couplings are much better reproduced in the NS than in the N simulations (see Table 6). Not surprisingly, observables specifically designed to measure dynamics, such as the per residue fluctuations, are generally better reproduced in the NS simulations (see Fig. 5). However, the RMSD between the average structures of the reference and the restrained ensembles at the optimal replica numbers deteriorates when S^2 order parameters are added as restraints in addition to nOes. In other words, observables reporting on dynamics are better reproduced in the NS simulations, whereas observables sensitive to the average structure are optimally reproduced in the N simulations with low replica numbers.

In the NS simulations, two types of restraints are used and they are susceptible to overfitting and underfitting at different replica numbers. As the N

simulations show, nOe restraints are prone to overfitting, even at relatively low replica numbers. The S^2 restraints, on the other hand, are much more susceptible to underfitting, since at each time step, all possible NH bond vector or side chain bond vector orientations should be represented. Thus a fairly large number of replicas is necessary for the S^2 restraint to work with accuracy. As a result, average properties are sacrificed at high replica numbers when the dynamics are best characterized.

The MUMO approach reproduces native state structure and dynamics

The considerations so far suggest that when restraining two or more different NMR observables simultaneously methods should be devised such that each observable is restrained with its optimal replica number. Such an approach should in principle alleviate the overfitting and underfitting problems that are observed in the NS simulations. The MUMO (minimal under-restraining minimal over-restraining) procedure implements these ideas and enables us to reproduce the average structure as well as in the case of the $N1$ and $N2$ simulations and the dynamics as well as in the NS8 and NS16 simulations.

In the MUMO method, nOe distances are restrained for pairs of replicas, whereas the S^2 restraints are applied to eight or sixteen replicas. In order to prevent structures that share nOe restraints from becoming too dissimilar, the pairs were overlapped as shown in Fig. 1. This procedure prevents overfitting the nOe distances no matter how large an ensemble is used for the S^2 order parameters, while ensuring that the pairs do not diverge too far from each other.

Validation for the MUMO restrained ensembles demonstrates the success of this method in producing structure and dynamics simultaneously; the RMSD between the average structures remains almost as low as in the $N1$ simulation (Table 5). The best Q-factor (0.27) for the RDCs (N–H_N, H_N–C_O, N–C_O, C_α–H_α, C_α–C_β, and C_α–C_O, see Methods), is observed for the MUMO16 ensemble (Table 6), which, remarkably, is close to that (0.24) obtained with the use of exact distances as restraints (N_eS8). Cross-validation with backbone, side-chain, and hydrogen bond J-couplings is as good or better as for the NS and N simulations for both the MUMO8 and MUMO16 ensembles (Table 6). Similarly, the per residue fluctuations and the average pairwise RMSD closely reproduce those of the reference ensemble (Table 5 and Fig. 5).

Agreement for the distance and rotamer distributions is comparable to that achieved by NS8 for

MUMO8 and to that achieved by NS16 for MUMO16 (Table 7). The NH bond vector distributions, on the other hand, are significantly better reproduced in the MUMO simulations. Figure 8 shows the NH bond vector distributions for residue Asp32, which has a backbone order parameter of 0.88 in the reference simulation, for both the NS16 and the MUMO16 case. The MUMO simulation is more successful in replicating the behavior of this bond vector. Additionally, the correlation for the side-chain S^2 order parameters for the pooled ensembles is higher for the MUMO than for the NS simulations (Table 4 and Fig. 3), although the same algorithm was used to enforce the S^2 restraints at each cycle. This result demonstrates that the added rigidity provided by enforcing the nOe restraints over pairs of overlapping replicas in the MUMO algorithm aids in the effective enforcement of the order parameter restraint.

Determination ubiquitin ensemble using MUMO with experimental data

The use of a synthetic reference ensemble proved very effective in understanding the trends of overfitting and underfitting and in developing the MUMO method. However, there are many assumptions and simplifications inherent in the use of synthetic data. Thus we applied the MUMO16 method to the experimental nOe distances and S^2 order parameters available for native state ubiquitin.

We assess the generated ensemble (PDB code 2RN2) by comparing it to the DER ensemble of 128 structures (PDB code 1XQQ), which was produced with the same nOe and S^2 data and a similar solvation and annealing set-up (Lindorff-Larsen et al. 2005). Validation with unrestrained parameters demonstrates the success of the introduction of the pairwise nOe restraint in the MUMO16 algorithm (Table 8). The MUMO16 method yields an RDC Q-factor of 0.19 compared to 0.26 for the DER ensemble. The correlation coefficient for the hydrogen bond J-couplings increases to 0.84 from 0.70 observed for the DER ensemble. The RMSD of the average structure of the MUMO16 ensemble to the average of the RDC-restrained NMR ensemble (1D3Z), which represents the average geometry of the native state ubiquitin extremely well), is 0.31 Å, whereas the RMSD of the average DER structure is 0.40 Å. Taken together, these results demonstrate that the MUMO method introduced in this study presents a highly accurate protocol for simultaneously determining the structure and dynamics of native state proteins.

Table 8 Validation of experimental MUMO16 and DER ensembles with backbone RDC Q-factors, hydrogen bond ^{3h}J-coupling correlation coefficients, and the RMSD of the geometric average structures to the average structure of the RDC-restrained NMR ensemble (1D3Z) (Cornilescu et al. 1998)

Observable	MUMO16 Ensemble	DER Ensemble
RDC Q-factor	0.19	0.26 (Lindorff-Larsen et al. 2005)
^{h3}J-Coupling correlation	0.84	0.70
RMSD of average structures	0.31Å	0.40 Å

Conclusions

We have analysed the effects of the underfitting and the overfitting problems on procedures for simultaneously determining the structure and the dynamics of the native states of proteins. An extensive comparison of different types of simulated annealing simulations with ensemble-averaged restraints has shown that nOe restraints are extremely sensitive to overfitting whereas S^2 restraints are more susceptible to underfitting. As a solution to this problem, we proposed the MUMO procedure, in which different observables are ensemble-averaged over a different number of molecules. The best results were obtained when nOe distances were averaged over pairs of molecules in a sixteen-member ensemble while S^2 restraints were enforced on all members. Application to the native state of ubiquitin using experimentally measured nOe distances and S^2 order parameters shows that the MUMO method is capable of providing ensembles at high resolution. Furthermore, as the MUMO approach can be readily extended to include other NMR observables that contain information about the dynamics of proteins, in particular RDCs and J couplings, it should serve as a general procedure for performing restrained molecular dynamics simulations.

References

Barfield M (2002) Structural dependencies of interresidue scalar coupling (h3)J(NC), and donor H-1 chemical shifts in the hydrogen bonding regions of proteins. J Am Chem Soc 124:4158–4168

Bax A (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. Prot Sci 12:1–16

Beglov D, Roux B (1994) Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. J Chem Phys 100:9050–9063

Best RB, Vendruscolo M (2004) Determination of protein structures consistent with NMR order parameters. J Am Chem Soc 126:8090–8091

Bonvin AMJJ, Boelens R, Kaptein R (1994) Time- and ensemble-averaged direct NOE restraints. J Biomol NMR 4:143–149

Bonvin AMJJ, Brunger AT (1995) Conformational variability of solution nuclearmagnetic-resonance structures. J Mol Biol 250:80–93

Bonvin AMJJ, Brünger AT (1996) Do NOE distances contain enough information to assess the relative populations of multi-conformer structures? J Biomol NMR 7:72–76

Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: A program for macromolecular energy, minimization and dynamics calculations. J Comp Chem 4:187–217

Brünger A (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355:472–475

Brünger A, Adams P, Clore G, DeLano W, Gros P, Grosse-Kunstleve R, Jiang J, Kuszewski J, Nilges M, Pannu N, Read R, Rice L, Simonson T, Warren G (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Cryst D 54:905–921

Brünger A, Clore GM, Gronenborn A, Saffrich R, Nilges M (1993) Assessing the quality of solution nuclear magnetic resonance structures by complete cross-validation. Science 261:328–331

Brünger AT, Clore GM, Gronenborn AM, Karplus M (1986) Three-dimensional structure of proteins determined by molecular dynamics with interproton distance restraints: Application to crambin. Proc Natl Acad Sci USA 83:3801–3805

Bürgi R, Pitera J, van Gunsteren WF (2001) Assessing the effect of conformational averaging on the measured values of observables. J Biomol NMR 19:305–320

Carlson HA (2002) Protein flexibility and drug design: How to hit a moving target. Curr Opin Cell Biol 6:447–452

Carlson HA, McCammon JA (2000) Accomodating protein flexibility in computational drug design. Mol Pharmacol 57:213–218

Chang S, Tjandra N (2005) Temperature dependence of protein backbone motion from carbonyl ^{13}C and amide ^{15}N NMR relaxation. J Magn Res 174:43–53

Chen J, Brooks CL, Wright PE (2004) Model-free analysis of protein dynamics: Assessment of accuracy and model selection protocols based on molecular dynamics simulation. J Biomol NMR 29:243–257

Chou J, Case D, Bax A (2003) Insights into the mobility of methyl-bearing side chains in proteins from $^3J_{CC}$ and $^3J_{CN}$ couplings. J Am Chem Soc 125:8959–8966

Clore GM, Schwieters CD (2004) How much backbone motion in ubiquitin is required to account for dipolar coupling data measured in multiple alignment media as assessed by independent cross-validation? J Am Chem Soc 126:2923–2938

Clore GM, Schwieters CD (2004) Amplitudes of protein backbone dynamics and correlated motions in a small alpha/beta protein: Correspondence of dipolar coupling and heteronuclear relaxation measurements. Biochemistry 43:10678–10691

Clore GM, Schwieters CD (2006) Concordance of residual dipolar couplings, backbone order parameters and crystallographic B-factors for a small alpha/beta protein: A unified picture of high probability, fast atomic motions in proteins. J Mol Biol 355:879–886

Cordier F, Grzesiek S (1999) Direct observation of hydrogen bonds in proteins by interresidue 3hJNC' scalar couplings. J Am Chem Soc 117:5179–5197

Cordier F, Grzesiek S (2002) Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR. J Mol Biol 317:739–752

Cornilescu G, Marquardt JL, Ottiger M, Bax A (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. J Am Chem Soc 120:6836–6837

de Alba E, Tjandra N (2002) NMR dipolar couplings for the structure determination of biopolymers in solution. Prog Nucl Mag Res Spec 40:175–197

Fennen J, Torda AE, van Gunsteren WF (1995) Structure refinement with molecular-dynamics and a boltzmann-weighted ensemble. J Biomol NMR 6:163–170

Grunberg R, Leckner J, Nilges M (2004) Complementarity of structure ensembles in protein–protein binding. Science 12:2125–2136

Grzesiek S, Cordier F, Jaravine V, Barfield M (2004) Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings. Prog Nucl Mag Res Spec 45:275–300

Henry ER, Szabo A (1985) Influence of vibrational motion on solid state line shapes and NMR relaxation. J Chem Phys 82:4753–4761

Hess B, Scheek RM (2003) Orientation restraints in molecular dynamics simulations using time and ensemble averaging. J Magn Res 164:19–27

Ichiye T, Karplus M (1988) Anisotropy and anharmonicity of atomic fluctuations in proteins: Implications for x-ray analysis. Biochemistry 27:3487–3497

Jorgensen WJ, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935

Karplus M (1963) Vicinal proton coupling in nuclear magnetic resonance. J Am Chem Soc 85:2870–2871

Karplus M, Kuriyan J (2005) Molecular dynamics and protein function. Proc Natl Acad Sci USA 102:6679–6685

Karplus M, McCammon JA (2002) Molecular dynamics simulations of biomolecules. Nat Struct Biol 9:646–652

Karplus M, Petsko G (1990) Molecular dynamics simulations in biology. Nature 347:631–639

Kemmink J, Scheek RM (1995) Dynamic modeling of a helical peptide in solution using NMR data—multiple conformations and multi-spin effects. J Biomol NMR 6:33–40

Kuriyan J, Petsko G, Levy RM, Karplus M (1986) Effect of anisotropy and anharmonicity on protein crystallographic refinement: An evaluation by molecular dynamics. J Mol Biol 190:227–254

Kuszewski J, Gronenborn AM, Clore GM (1999) Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. J Am Chem Soc 121:2337–2338

Lee AL, Flynn PF, Wand AJ (1999) Comparison of ^2h and ^{13}c NMR: relaxation techniques for the study of protein methyl group dynamics in solution. J Am Chem Soc 121:2891–2902

Lindorff-Larsen K, Best RB, DePristo MA, Dobson CM, Vendruscolo M (2005) Simultaneous determination of protein structure and dynamics. Nature 433:128–132

Lipari G, Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. J Am Chem Soc 104:4546–4559

Ma B, Nussinov R (2002) Stabilities and conformations of Alzheimer's β-amyloid peptide oligomers (A$\beta_{16–22}$, A$\beta_{16–35}$,

Aβ_{10-35}): Sequence effects. Proc Natl Acad Sci USA 99:14126–14131

MacKerell Jr AD, Bashford D, Bellot M, Dunbrack RLJ, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Ha S, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich B, Smith JC, Stote RH, Straub J, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586–3616

Mezard M, Parisi G, Virasoro M (1987) Spin glass theory and beyond. Singapore, World Scientific Publishing

Nederveen A, Bonvin AMJJ (2005) NMR relaxation and internal dynamics of ubiquitin from a 0.2 microsecond MD simulation. J Chem Theor Comp 1:363–374

Neuhaus D, Williamson MP (2000) The nuclear Overhauser effect in structural and conformational analysis. New York, Wiley

Paci E, Karplus M (1999) Forced unfolding of fibronectin type 3 modules: An analysis by biased molecular dynamics simulations. J Mol Biol 288:441–459

Palmer AG (2004) NMR characterization of the dynamics of biomacromolecules. Chem Rev 104:3623–3640

Perryman AL, Lin J, McCammon JA (2004) HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: Possible contributions to drug resistance and a potential new target site for drugs. Prot Sci 13:1108–1123

Rieping W, Habeck M, Nilges M (2005) Inferential structure determination. Science 309:303–306

Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion if a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341

Schneider TR, Brünger AT, Nilges M (1999) Influence of internal dynamics on accuracy of protein NMR structures: Derivation of realistic model distance data from a long molecular dynamics trajectory. J Mol Biol 285:727–740

Schwieters CD, Kuszewski JJ, Clore GM (2005) Using Xplor-NIH for NMR molecular structure determination. Prog Nucl Mag Res Spec 48:47–62

Scott WRP, Hunenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Kruger P, van Gunsteren WF (1999) The GROMOS biomolecular simulation program package. J Phys Chem A 103:3596–3607

Scott WRP, Mark AE, van Gunsteren WF (1998) On using time-averaging restraints in molecular dynamics simulation. J Biomol NMR 12:501–508

Spronk CAEM, Nabuurs SB, Krieger E, Vriend G, Vuister GW (2004) Validation of protein structures derived by NMR spectroscopy. Prog Nucl Mag Res Spec 45:315–337

Teague SJ (2003) Implications of protein flexibility for drug discovery. Nat Rev Drug Discov 2:527–541

Tjandra N, Feller SE, Pastor RW, Bax A (1995) Rotational diffusion anisotropy of human ubiquitin from n-15 NMR relaxation. J Am Chem Soc 117:12562–12566

Torda AE, Scheek RM, van Gunsteren WF (1989) Time averaged distance restraints in molecular dynamics simulations. Chem Phys Lett 157:289–294

Torda AE, Scheek RM, van Gunsteren WF (1990) Time-averaged nuclear Overhauser effect distance restraints applied to tendamistat. J Mol Biol 214:223–235

Vendruscolo M, Dobson CM (2005) Towards complete descriptions of the free-energy landscapes of proteins. Phil Trans R Soc A 363:433–450

Vendruscolo M, Paci E (2003) Protein folding: Bringing theory and experiment closer together. Curr Opin Struct Biol 13:82–87

Vijay-Kumar S, Bugg CE, Cook WJ (1987) Structure of ubiquitin refined at 1.8 A resolution. J Mol Biol 194:531–544

Wand AJ (2001) Dynamic activation of protein function: A view emerging from NMR spectroscopy. Nat. Struct Biol 8:926–931

Wang T, Cai S, Zuiderweg ERP (2003) Temperature dependence of anisotropic protein backbone dynamics. J Am Chem Soc 125:8639–8643

Wong CF, McCammon JA (2003) Protein flexibility and computer-aided drug design. Annu Rev Pharmacol Toxicol. 43:31–45

Wüthrich K (1986) NMR of proteins and nucleic acids. New York, Wiley

Zagrovic B, van Gunsteren WF (2006) Comparing atomistic simulation data with the NMR experiment: How much can NOEs actually tell us?. Proteins 63:210–218

Zweckstetter M, Bax A (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. J Am Chem Soc 122:3791–3792