Biophys Rev (2012) 4:189–203
DOI 10.1007/s12551-012-0087-6

REVIEW

Assessing and refining molecular dynamics simulations of proteins with nuclear magnetic resonance data

Jane R. Allison

Received: 28 April 2012 / Accepted: 12 June 2012 / Published online: 2 August 2012

Abstract The sophistication of the force fields, algorithms and hardware used for molecular dynamics (MD) simulations of proteins is continuously increasing. No matter how advanced the methodology, however, it is essential to evaluate the appropriateness of the structures sampled in a simulation by comparison with quantitative experimental data. Solution nuclear magnetic resonance (NMR) data are particularly useful for checking the quality of protein simulations, as they provide both structural and dynamic information on a variety of temporal and spatial scales. Here, various features and implications of using NMR data to validate and bias MD simulations are outlined, including an overview of the different types of NMR data that report directly on structural properties and of relevant simulation techniques. The focus throughout is on how to properly account for conformational averaging, particularly within the context of the assumptions inherent in the relationships that link NMR data to structural properties.

Keywords Molecular dynamics · Nuclear magnetic resonance · Protein · Biomolecular simulation

Introduction

In order to understand the function and malfunction of proteins, it is essential to characterise not only their structure but also their dynamics. Biomolecular simulation is an extremely valuable tool for probing these at a detailed molecular level, uniquely allowing direct visualisation. It can suffer, however, from limited force field accuracy and insufficient sampling. Because of this, it is important to check, and, if necessary, improve the relevance of, the structures sampled during a simulation by comparing back-calculated experimentally measurable quantities with experimental data. Nuclear magnetic resonance (NMR) data is particularly useful for such comparisons because it reports on both structure and dynamics, although this information can be encoded in a complex manner.

The focus here is on the use of solution NMR data to validate the ensembles of structures produced by molecular dynamics (MD) simulations of proteins using atomic-level force fields, and to bias such simulations towards structures that are in keeping with the data. Whilst solid state NMR data have recently been utilised for protein structure determination and refinement (Castellani et al. 2002; Andronesi et al. 2005; Siemer et al. 2005; Manolikas et al. 2008; Van Melckebeke et al. 2010), a field in which many of the principles and techniques discussed here are also relevant, these topics are outside the scope of this article.

This review begins by considering the average nature of NMR data, followed by an outline of the relationships that link NMR observables to structural properties of proteins. Subsequently, the use of NMR data to validate MD simulations of proteins is discussed, including factors that affect the agreement between back-calculated and experimental observables. Different means of restraining or biasing simulations to fit NMR data are then described and evaluated. The article ends with a summary of the state of the art, and some thoughts on future developments.

Averaging of NMR data

An important property of NMR data is that they are ensemble-averages over the potentially different conformations of the

Special issue: Computational Biophysics

J. R. Allison (✉)
Centre for Theoretical Chemistry and Physics, Institute of Natural Sciences, Massey University Albany,
Albany Highway,
Auckland 0632, New Zealand
e-mail: j.allison@massey.ac.nz

many molecules present in the experiment, and time-averages over motions that occur faster than the chemical shift difference or the inverse of the coupling constant, depending on the type of NMR experiment (Bryant 1983). This averaging has important consequences when NMR data are interpreted in terms of structural properties, particularly for mobile or disordered states.

The ensemble of different structures sampled by the many molecules present over the duration of the experiment can be thought of in terms of distributions of values of the structural properties, such as dihedral angles or inter-nuclear distances, that are related to NMR observables as outlined in Section 3. These distributions may or may not be correlated to one another. Each value of the structural property in the distribution may give rise to a different value of the related NMR observable; it is the average value of the NMR observable that is measured in an experiment. Because of this, the most naive structural interpretation of an NMR observable in terms of a single value of a structural property back-calculated from the averaged observable is only appropriate if the molecule is rigid (Fig. 1). Reconstruction of the distributions of structural properties, and therefore the ensemble of structures, from the averaged NMR observables is a non-trivial problem, as the relationships linking NMR observables to structural properties are often multiple-valued and non-linear (see Section 3). An advantage of using MD simulations to interpret NMR data is that the shape of the underlying distribution of structural properties described by the data does not have to be assigned a priori as it does when using analytical models. Properly accounting for conformational averaging remains difficult, however, and is therefore discussed throughout this review.

Relating NMR data to protein structure

In order to compare MD simulations with NMR data, the value of the NMR observables Q must be calculated from the Cartesian coordinates of the protein $\mathbf{r}$ or vice versa. In principle, Q can be calculated from $\mathbf{r}$ using quantum-chemical methods (Oldfield 2002; Bühl and van Mourik 2011), but the accuracy that can currently be reached is rather low because a number of approximations must be made due to the expense of the calculations (Mulder and Filatov 2010; Frank et al. 2012). This also renders such calculations too expensive to carry out "on the fly" during an MD simulation. Therefore, protein structure is generally related to observable quantities according to semi-empirical, approximate functions $Q(\mathbf{r})$ and their inverses. These relationships are outlined and discussed below for commonly used NMR observables that provide three main types of structural information: distance, angular, and orientational (Fig. 2).

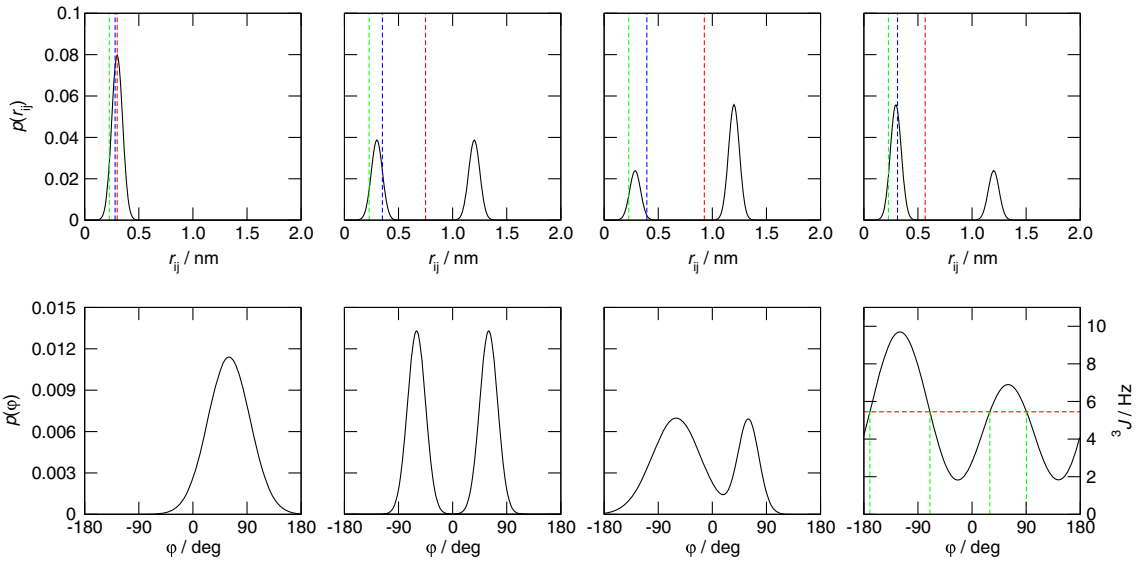

Fig. 1 The combined effect of conformational averaging and the non-linear and multiple-valued relationships linking NMR observables to structural properties. *Upper row* four different distributions of inter-nuclear distances r_{ij} and (*dashed lines*) the corresponding (*red*) linearly, (*blue*) r^{-3} and (*green*) r^{-6} averaged distances. All four distributions have the same r^{-6} average (0.227 nm) upon which PREs and many NOEs depend, and are therefore indistinguishable at the level of the NMR observable. *Lower row* three different distributions of values of a dihedral angle φ and (*rightmost graph*) the Karplus relation for the φ angle of the protein backbone (Pardi et al. 1984). All three distributions give rise to the same average ^3J-value (5.45 Hz, *red dashed line in rightmost graph*) and thus are indistinguishable at the level of the NMR observable. (*Rightmost graph, green dashed line*) the four dihedral angle values that would be predicted from a ^3J-value of 5.45 Hz without allowing for conformational averaging

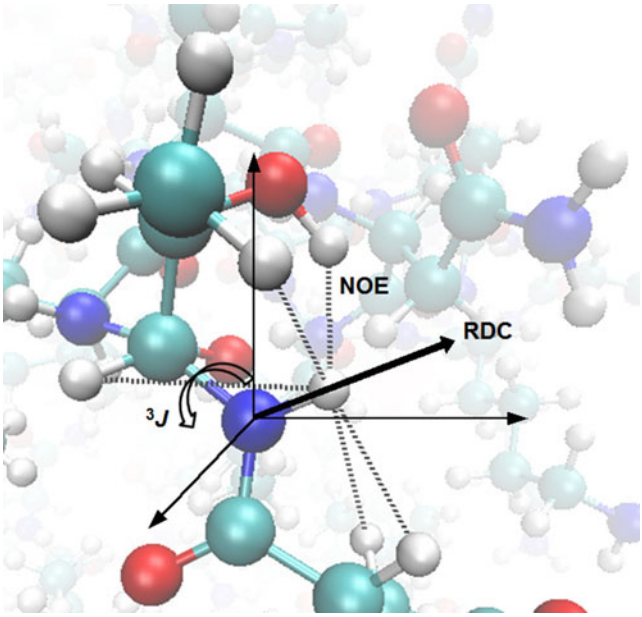

Fig. 2 The types of structural information obtained from the NMR observables introduced in Section 3: *dashed lines* short-range NOE-derived distances, *curved arrow* dihedral angle φ reported on by $^3J_{\mathrm{HNC}\alpha\mathrm{H}\alpha}$-couplings and *heavy line* the orientation of the N–H bond vector relative to the molecular frame specified by an alignment tensor **A** attainable from RDCs. The protein atoms are coloured according to type: *cyan* carbon, *blue* nitrogen, *red* oxygen and *white* hydrogen

Distance information

Short-range distances up to approximately 0.6 nm are obtained from nuclear Overhauser effects (NOEs), the transfer of spin polarisation between nuclei through space via cross-relaxation (Wüthrich 1986). Generally, only semi-quantitative distance information is extracted from NOESY spectra using the isolated spin pair approximation (ISPA)

$$r_{ij} = r_{\mathrm{ref}} \left(\frac{a_{\mathrm{ref}}}{a_{ij}} \right)^{-1/x}, \qquad (1)$$

where r_{ij} is the distance between two nuclei i and j, a_{ij} is the cross-peak intensity or volume and x=−3 if internal structural fluctuations are significantly faster than the tumbling time of the molecule, or −6 otherwise (Tropp 1980). a_{ref} is a reference intensity corresponding to an inter-nuclear distance r_{ref} that is known due to internal geometry; these serve to calibrate the conversion. This approach assumes that the mixing time τ_{m} is short, so that the NOE buildup is approximately linear, and that internal motion is negligible. In practise, these expectations are not always met. A short mixing time also limits the effects of spin diffusion (Kalk and Berendsen 1976), one of the major causes of inaccuracy in NOE-derived distances (Keepers and James 1984). In most cases, spin diffusion leads to enlarged NOE intensities and, therefore, under-estimated inter-nuclear distances. A complete description of spin diffusion can only be obtained with a full relaxation matrix formalism, which is computationally very intensive. Nevertheless, several approaches are available for correcting NOE-derived distance restraints for spin diffusion effects during iterative structure refinement (Marion et al. 1987; Boelens et al. 1988; Yip and Case 1989; Borgias and James 1990; Borgias et al. 1990; Post et al. 1990; Edmondson 1992; Leeflang and Kroon-Batenburg 1992; Linge et al. 2004), although these are typically not used in the analysis or biasing of MD simulations.

Distances obtained using Eq. 1 are applied with large tolerances to account for the many sources of uncertainty mentioned above. Often, only an upper bound on the inter-nuclear distance is used, with the sum of the van der Waals radii of the two nuclei involved acting as a proxy for the lower bound. The use of conservative bounds compensates for alteration of cross-peaks by spin diffusion or partial overlap, as well as other possible sources of error. Alternatively, the NOE intensities may simply be sorted into a few groups based on relative peak intensities, with weak, medium and strong intensities corresponding to short, medium and long distances. The upper bounds on the distance for each group are chosen based on cross-peaks stemming from pairs of protons for which the distance can reasonably be estimated, such as those identified as being in regular secondary structure elements according to chemical shift data or a priori structural knowledge. Upper bounds can also be estimated from the maximum distance for which an NOE cross-peak is expected to be observed (0.5−0.6 nm). Whilst a single NOE-derived distance is relatively imprecise, this is compensated for by the large numbers of NOEs typically measured.

In certain cases, such as large proteins or disordered states, long-range distance information may be desirable. Distances in the range 1.2−2.0 nm can be obtained from paramagnetic relaxation enhancement (PRE) experiments, which exploit the change in the relaxation rate of a nuclear spin induced by the presence of a distant paramagnetic group (Solomon 1955; Gillespie and Shortle 1997; Battiste and Wagner 2000; Donaldson et al. 2001; Clore and Iwahara 2009). The paramagnetic relaxation enhancement is quantified by the ratio of the intensities of the cross-peaks with the paramagnetic group present and absent (diamagnetic state), $I_{\mathrm{para}}/I_{\mathrm{dia}}$. From this, the paramagnetic relaxation rate R_2^{sp} can be determined by fitting (Gillespie and Shortle 1997; Battiste and Wagner 2000)

$$\frac{I_{\mathrm{para}}}{I_{\mathrm{dia}}} = \frac{R_2 \exp\left(-R_2^{\mathrm{sp}} t\right)}{\left(R_2 + R_2^{\mathrm{sp}}\right)}, \qquad (2)$$

where t is the total INEPT delay time and R_2 is the intrinsic relaxation rate, usually estimated from the half-height linewidth in the diamagnetic state. The electron–proton distance is then calculated according to

$$r = \left[\frac{K}{R_2^{sp}}\left(4\tau_c + \frac{3\tau_c}{1+\omega_H^2\tau_c^2}\right)\right]^{-1/6}, \quad (3)$$

where ω_H is the Larmor frequency of the proton, τ_c is the correlation time of the electron–proton vector and K is a combination of physical constants.

The number of PRE-derived distances obtained is limited by the need to include a paramagnetic centre, which typically requires the creation of many artificial constructs with paramagnetic labels attached. Moreover, the presence of the label must be accounted for when inferring and utilising the inter-nuclear distances.

Angular information

Information regarding dihedral angles spanned by three covalent bonds or the geometry of hydrogen bonds is provided by through-bond ^3J-couplings. These are usually related to the molecular geometry using the Karplus relation: (Karplus 1959)

$$^3J(\theta) = A\cos^2\theta + B\cos\theta + C, \quad (4)$$

where θ is the torsion or hydrogen bond angle and A, B and C are empirical parameters. More complex relationships have also been proposed (Haasnoot et al. 1979; Haasnoot et al. 1981; Imai and Osawa 1990; Suardıaz et al. 2007; Schmidt 2007). The constants A, B and C are generally estimated by fitting ^3J-values measured for molecules whose dihedral angle values are presumed to be known. More sophisticated fitting methods such as Bayesian inference have also been used (Habeck et al. 2005). Efforts have been made to incorporate the effects of dynamics and averaging into the parameters by self-consistent fitting (Schmidt et al. 1999; Pérez et al. 2001) and by deriving the parameters from ensembles of structures obtained from MD simulations (Brüschweiler and Case 1994; Lindorff-Larsen et al. 2005b; Vögeli et al. 2007), but such parameters are not necessarily transferable. When the Karplus relation is used, its approximate form and parameters mean that ^3J-couplings should be interpreted with an uncertainty of at least ±1 Hz (Allison and van Gunsteren 2009; Steiner et al. 2012). The multiple-valued nature of this relationship further complicates the extraction of angular information, as a single ^3J-value can be compatible with up to four dihedral angle values (see Fig. 1, lower right graph). Whilst this degeneracy can, in some cases, be resolved by the measurement of ^3J-couplings between different sets of atoms spanning a given dihedral angle (Smith et al. 1991; Schwalbe et al. 2001; Allison and van Gunsteren 2009), for dihedral angles in mobile parts of a protein such as loops or side-chains, the analytical derivation of angular information remains extremely difficult.

Dihedral angle information can also be extracted from the chemical shifts of the protein nuclei, albeit indirectly. Although chemical shifts are influenced by a wide array of structural properties (Wagner et al. 1983; Osapay and Case 1991; Williamson and Asakura 1993; Case 1995; Wishart et al. 1991, 1995), the dependence of Cα and Cβ chemical shifts on the backbone ϕ and ψ dihedral angles (Pastore and Saudek 1990; Spera and Bax 1991; de Dios et al. 1993) is routinely used to infer protein secondary structure (Spera and Bax 1991; Wishart et al. 1991, 1992; Cornilescu et al. 1999; Peti et al. 2001; Yao et al. 2001; Wang and Jardetzky 2002; Eghbalnia et al. 2005; Marsh et al. 2006; Shen et al. 2009). There is no simple analytical relationship linking chemical shifts to protein structure. Methods for calculating chemical shifts using quantum mechanics are improving, but remain too slow to be applicable during simulations (Mulder and Filatov 2010; Frank et al. 2012). A number of fast empirical methods for calculating chemical shifts approximately have been developed (Xu and Case 2001; Neal et al. 2003; Meiler 2003; Shen and Bax 2007; Kohlhoff et al. 2009; Shen and Bax 2010; Han et al. 2011; Nielsen et al. 2012), although even these state-of-the-art programs often result in errors much larger than the deviation between back-calculated and experimental chemical shifts (Kjaergaard and Poulsen 2012), and so should be used with caution.

Orientational information

One source of orientational information is residual dipolar couplings (RDCs) (Tjandra and Bax 1997; Blackledge 2005) between bonded or non-bonded pairs of nuclei. Measurement of RDCs requires the protein to be partially aligned with respect to the magnetic field direction (Bax 2003). Various alignment media have been developed for this task (Sanders et al. 1994; Tjandra and Bax 1997; Clore et al. 1998; Hansen et al. 1998; Sass et al. 1999. 2000; Koenig et al. 1999; Tycko et al. 2000). Partial alignment prevents complete isotropic rotation of the protein, which would average the couplings to zero, but allows sufficient rotation to reduce the dipolar couplings to manageable magnitudes.

RDCs are related to the angle θ between an inter-nuclear vector r_{ij} and the direction of the magnetic field, as well as the length of the inter-nuclear vector r_{ij} , according to:

$$RDC = -\frac{\gamma_i\gamma_j\mu_0 h}{8\pi^3}\left\langle\frac{(3\cos^2\theta - 1)}{2r_{ij}^3}\right\rangle_{mol,time}, \quad (5)$$

where γ_i and γ_j are the gyromagnetic ratios of the two

nuclei, μ_0 is the magnetic permeability of vacuum and h is Plancks constant. The inter-nuclear distance r_{ij} is usually assumed to be constant when i and j are bonded, consistent with the constraint of bond lengths in classical atomic-level force fields.

In order to simplify the interpretation of RDCs, the alignment of the protein with respect to the magnetic field is often separated from its internal structure. If the protein is assumed to be rigid, Eq. 5 can be factorised into structural and orientational components by using an alignment tensor, **A** (Saupe 1968), to describe the non-isotropic alignment of the protein, represented by the molecular frame, and defining the orientation of the inter-nuclear vector with respect to the molecular frame (Blackledge 2005):

$$RDC = -\frac{\gamma_i \gamma_j \mu_0 h}{8\pi^3} \sum_{\alpha \in x,y,z} \sum_{\beta \in x,y,z} A_{\alpha\beta} \frac{\cos(\zeta_{ij,\alpha}) \cos(\zeta_{ij,\beta})}{r_{ij}^3} \quad (6)$$

where $\zeta_{ij,\alpha}$ is the angle between r_{ij} and the α-axis of the molecular frame, and the nine components, five of which are independent, of the alignment tensor are given by $A_{\alpha\beta} = 3/2\langle \cos\xi_\alpha \cos\xi_\beta \rangle$, with ξ_α the angle between the α-axis of the molecular frame and the magnetic field direction. The alignment tensor encompasses averaging over different orientations of the protein. Whilst some have averaged the angles $\zeta_{ij,\alpha}$ over the trajectory prior to fitting **A** (Markwick et al. 2009), this is not strictly correct, as the derivation of Eq. 6 depends on the assumption that the molecule is rigid. Averaging of $\zeta_{ij,\alpha}$ also requires superposition of the structures, which cannot be done sensibly if the protein undergoes large-scale structural changes during the simulation. Thus, although the use of an alignment tensor greatly simplifies the extraction of structural information from RDCs, the assumptions inherent in its use are not always appropriate. This is particularly pertinent for proteins that are partially or entirely disordered, but even small fluctuations may be sufficient to perturb the alignment (Louhivuori et al. 2007), although such fluctuations appear to be largely uncorrelated with the changes in alignment (Louhivuori et al. 2006). Even when it is appropriate to use an alignment tensor, the extraction of structural information from RDCs is not trivial, due to the degeneracy and non-linearity of Eqs. 5 and 6.

Validation of MD simulations with NMR data

Comparison of the values of NMR observables back-calculated from and averaged over an ensemble of structures generated by an MD simulation to experimental data allows assessment of both the quality of the structures and the extent of sampling. There are two important points to consider in making such a comparison, however.

Firstly, because there are almost always fewer NMR observables N^{obs} than degrees of freedom of a protein N^{df}, conformational averaging must be accounted for, and the relationships linking observables to structural properties are non-linear and/or multiple-valued, it is likely that many different ensembles of structures can fit the experimental data equally well on average. There are many documented examples of this (Glättli and van Gunsteren 2004; Dolenc et al. 2010; Missimer et al. 2010; Allison et al. 2011, 2012; Eichenberger et al. 2012; Niggli et al. 2012), including several in which the majority of the NMR data are equally compatible with the folded or unfolded states (Daura et al. 1999, 2001; Peter et al. 2003; Zagrovic and van Gunsteren 2006). This insensitivity of NMR observables to the shape of the underlying probability distribution of structural properties has been demonstrated analytically (Bürgi et al. 2001) and is illustrated in the upper panel of Fig. 1. If it is possible to use multiple experimental techniques to measure different observables that report on the same or related structural features, but are subject to different types of averaging, then these can be combined to give more information about the shape of the underlying distribution (Choy et al. 2002; Allison et al. 2009). In general, combining many different datasets to maximise the amount of experimental information is a useful tactic, but is only strictly appropriate if the data were measured at the same or very similar thermodynamic state points.

A further concern is the extent of sampling of different conformations. In the limits of an infinitely long simulation, all possible conformations will occur with a relative probability given by the Boltzmann distribution. In practice, however, the length of a simulation is finite and often rather short in comparison to the time-scale encompassed by NMR data. The latter depends on the experiment and on the field strength (Fig. 3). Whilst the length of MD simulations of proteins has increased by several orders of magnitude from the 9.2 ps of the first MD simulation of a protein in 1976 (McCammon et al. 1977), with ns and even μs time-scales now routinely accessible, and ms simulations possible with specialised hardware (Shaw et al. 2009), still longer simulations may be needed for comparison with RDCs and chemical shifts, as well as to obtain statistically good sampling of faster motions. Indeed, earlier studies utilising shorter simulation times mostly, although not always (Philippopoulos et al. 1997), found reasonable agreement with S^2 order parameters, which report on ps–ns motions of bond vectors (Fig. 3), but sometimes struggled to reproduce NMR data averaged over longer time-scales (Brüschweiler et al. 1992; Chandrasekhar et al. 1992; Koerdel and Teleman 1992; Palmer and Case 1992; Eriksson et al. 1993; Smith

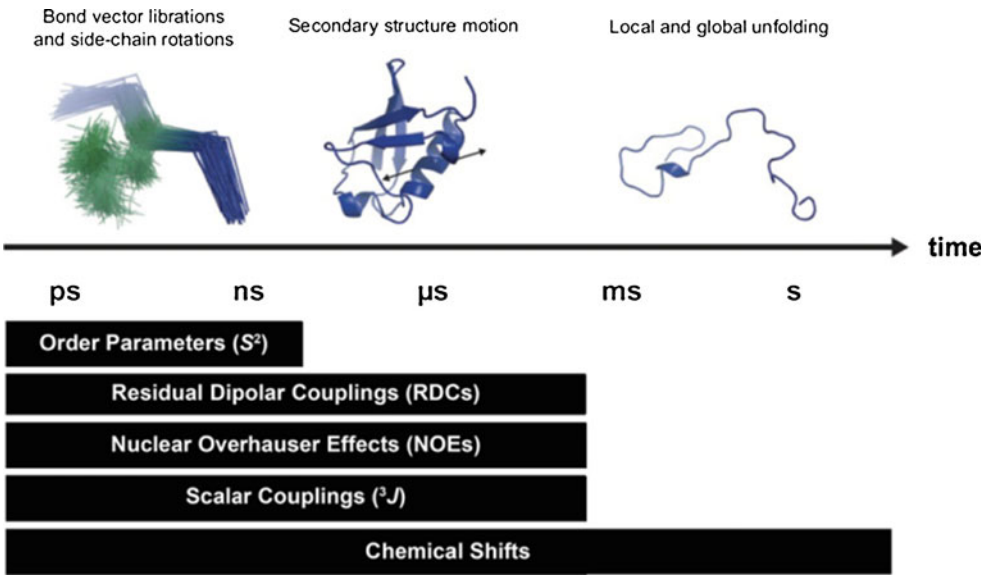

Fig. 3 The degree of protein conformational motion associated with the time-scales encompassed by the different types of NMR observables discussed here. Note that the scale on the *time axis* is logarithmic

et al. 1995a, b; Nanzer et al. 1997). Even in more recent studies, the agreement with experimental data is often worst for residues in loops or disordered regions (Wrabl et al. 2000; Tong et al. 2009); in some cases, these regions are simply excluded from the analysis (Showalter and Brüschweiler 2007).

Further evidence that longer simulations are required is the typically better agreement with experimental data obtained for small systems such as peptides, for which longer simulations have been possible for some time (Trzesniak et al. 2005). Only with the recent construction of ANTON, a massively parallel supercomputer purpose-built for carrying out MD simulations, has it become tractable to run ms-length MD simulations of proteins (Shaw et al. 2009). As well, extensive parallelisation of MD codes (Plimpton 1995; Phillips et al. 2005) and use of GPUs (Schmid et al. 2010; Stone et al. 2010; Brown et al. 2011) have opened up the possibility of conducting hundreds of ns or even μs simulations routinely even without access to supercomputers. Early results from a 1.2 μs simulation of ubiquitin showed that simultaneous consideration of all, rather than solely internal, motions reduced the discrepancies between calculated and experimental S^2 values (Maragakis et al. 2008). More recently, a 200 μs high-temperature simulation aimed at sampling the disordered state of Acyl-CoA-binding protein (ACBP) gave reasonable agreement with experimental NMR relaxation rate data (Lindorff-Larsen et al. 2012b), although the radius of gyration was smaller than the experimental value despite the elevated temperature, and comparison with the PRE-derived distances available for this protein (Teilum et al. 2002) was not made.

Despite being two orders of magnitude longer than previous simulations, the aforementioned simulation of ACBP was not converged, giving some indication of the lengths of simulations that will be required to characterise disordered states and other systems in which large-scale conformational motions are expected. Such long simulation times are required because of the low probability of crossing the high energy barriers that can exist between different conformations that co-exist or interconvert during an NMR experiment. One possible means of counteracting this is to simultaneously run multiple independent simulations of the same protein starting from randomly selected different conformations. So long as the starting conformations are sufficiently different to allow access to all barrier-separated regions of the potential energy surface, averaging over the resulting pooled ensemble of structures should give better agreement with experimental data. This was shown to be the case for NMR data including RDCs, although not ^3J-couplings across hydrogen bonds (Lange et al. 2010). In a similar way, enhanced sampling methods such as replica exchange (Fawzi et al. 2008) also result in improved agreement with experimental data.

Another means of visiting more regions of conformational phase space and crossing barriers that looks to be particularly promising is accelerated molecular dynamics (AMD) (Hamelberg et al. 2004). The potential energy landscape is smoothed by raising the low-energy regions, resulting in accelerated exchange between low-energy conformations whilst retaining the essential details of the landscape. The acceleration level is chosen according to reproduction of experimental data, and the resulting ensemble of structures is reweighted such that the correct canonical Boltzmann distribution is obtained. Due to the acceleration procedure, the time-scale of the motions sampled using AMD is not known explicitly, and can only be inferred indirectly by comparison with the time-scale of the NMR data to which the ensembles of structures are compared. Somewhat

remarkably, given the range of time-scales associated with different types of NMR data (Fig. 3), ensembles of structures obtained at the RDC-optimal acceleration level also reproduce experimental ^3J-couplings and order parameters (Markwick et al. 2009) or chemical shifts (Markwick et al. 2010).

A further potential source of discrepancy between back-calculated and experimental NMR data is the quality of the force field. Newly accessible long simulations of proteins are allowing detection of shortcomings in force fields that were not visible in shorter simulations (Klepeis et al. 2009) or in the simulations of small molecules usually used for force field parametrisation (Lindorff-Larsen et al. 2012a). Eight different force fields (Mackerell et al. 1998, 2004; Kaminski et al. 2001; Duan et al. 2003; Mackerell 2004; Hornak et al. 2006; Best and Hummer 2009; Lindorff-Larsen et al. 2010; Piana et al. 2011) were recently used to run 10-μs simulations of ubiquitin and GB3 (Lindorff-Larsen et al. 2012a). Four of the most recent variants produced ensembles of structures in very good agreement with experimental NMR data including ^3J-couplings, RDCs and S^2 order parameters. The key change responsible for the improved agreement with experimental data appears to be modification of the backbone potential, as many other studies using the same and other force fields refined in this way have also shown good agreement with experimental NMR data including ^3J-couplings, RDCs and chemical shifts (Hornak et al. 2006; Buck et al. 2006; Showalter and Brüschweiler 2007; Showalter et al. 2007; Markwick et al. 2007; Trbovic et al. 2008; Li and Brüschweiler 2009; Schmid et al. 2011; Beauchamp et al. 2012; Robustelli et al. 2012).

It cannot be excluded that uncertainty introduced by the use of approximate relationships $Q(\mathbf{r})$ could be causing some of the disagreement between back-calculated and experimental data. For instance, ^3J-couplings calculated using density functional theory rather than the Karplus relation (Eq. 4) from 500 ps simulations agree better with experimental data than those calculated from the X-ray or NMR model structures, despite the short length of the simulations (Markwick et al. 2002). It is likely, however, that limited sampling and force field inaccuracy play the most significant roles. A means of overcoming both of these problems simultaneously is to restrain or bias simulations to fit experimental data, as outlined in the next section. Effectively, the force field is adjusted to favour structures that agree with the experimental data and the sampling is directed towards these structures.

Biasing MD simulations to fit NMR data

To drive simulations to sample areas of conformational space that are in agreement with experimental data, a restraining function $V^{\mathrm{restr}}(\mathbf{r}(t))$ that penalises deviation from the experimental data is added to the physical interaction function term $V^{\mathrm{phys}}(\mathbf{r}(t)) : V^{\mathrm{pot}}(\mathbf{r}(t)) = V^{\mathrm{restr}}(\mathbf{r}(t)) + V^{\mathrm{phys}}(\mathbf{r}(t))$. Conventional restraining terms are harmonic:

$$V_i^{\mathrm{restr}}(\mathbf{r}(t)) = \frac{1}{2}K_i^{\mathrm{restr}} = \left(f_i^{\mathrm{obs}} - f_i^{\mathrm{calc}}(\mathbf{r}(t))\right)^2, \tag{7}$$

where K_i^{restr} is a user-defined force constant, f_i^{obs} is the value of the i^{th} observable measured experimentally or the corresponding structural property and $f_i^{\mathrm{calc}}(\mathbf{r}(t))$ is the value of the observable or structural property back-calculated from the simulation. The more $f_i^{\mathrm{calc}}(\mathbf{r}(t))$ deviates from f_i^{obs} the larger the energy penalty and the resulting forces on the protein. The force constant K_i^{restr} allows tuning of the strength of the restraint, providing a means of accounting for experimental or other uncertainty.

Various modifications of Eq. 7 are possible (Fig. 4). In many applications, the penalty function becomes linear when $\Delta f_i = \left|f_i^{\mathrm{obs}} - f_i^{\mathrm{calc}}(\mathbf{r}(t))\right|$ is large to prevent excessive forces. The restraint potential may be made "flat-bottomed" by only enforcing the restraint potential if Δf_i is greater than some threshold value. This is another useful means of accounting for uncertainty in the experimental data or in the relationship that converts between observables and structural properties. For NOE-derived distances, a half-harmonic potential is typically used, so that only distances larger than f_i^{obs} are penalised.

Either the experimental observable or the related structural property may be restrained. In either case, relationships such as those outlined in Section 3 are used to convert between structural properties and experimental observables, thus any inaccuracies or uncertainties in the nature or parametrisation of these relationships cannot be escaped. If the relationship is highly non-linear, a small change in the protein coordinates can have a large impact on the value of the back-calculated observable. For multiple-valued relationships, such as the Karplus relation (Eq. 4), restraining the structural property is only appropriate in cases where additional information exists indicating that only one of the multiple values is possible and specifying which value it is. For both non-linear and multiple-valued relationships, the restraining potential energy landscape is likely to be rugged, and specialised sampling techniques may be required.

The issues discussed at the start of Section 4 mean that restrained simulations tread a fine line between over- and under-fitting. Over-fitting occurs when the restraints are too stringently enforced (see Section 5.1), whereas under-fitting occurs when $N^{\mathrm{df}} \gg N^{\mathrm{obs}}$ (Bonvin and Brünger 1995), such that many different ensembles of structures fit the data equally well. A means of checking for under-fitting is cross-validation (Brünger 1992, 1993; Brünger et al. 1993), in which the reproduction of "free" data not used as restraints, which may comprise completely independent data

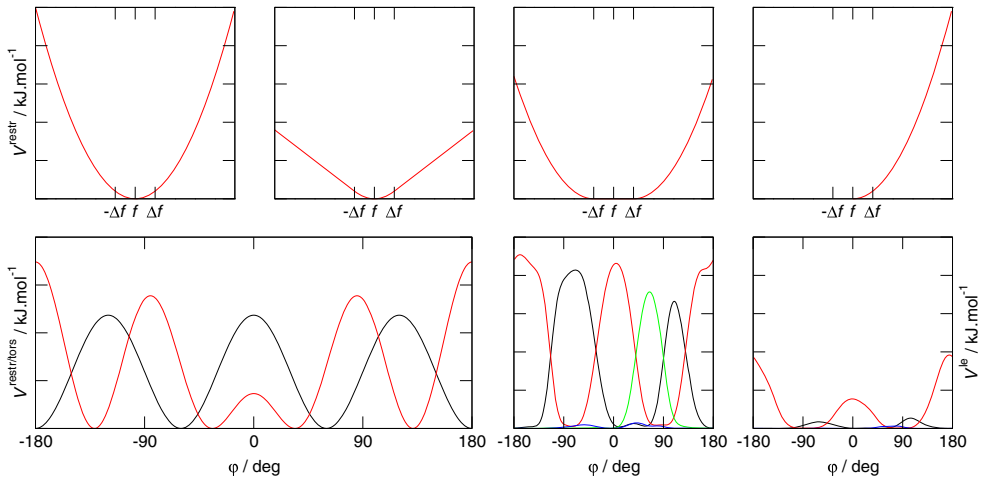

Fig. 4 Examples of the types of potential functions used to restrain or bias MD simulations to fit experimental data. *Upper row, left to right* a harmonic potential (f_i^{obs} is shown as f for clarity), a harmonic potential that becomes linear when $\Delta f_i = \left|f_i^{\text{obs}} - f_i^{\text{calc}}(\mathbf{r}(t))\right| > \Delta f$, a "flat-bottomed" harmonic potential that does not penalise deviations of less than Δf from f_i^{obs} and a half-harmonic potential as used for NOE-derived distances. *Lower row, left-hand side* potentials from (*black*) the torsional term of a force field and (*red*) a harmonic ^3J-value restraint acting on a dihedral angle φ. *Lower row, right-hand side* two sets of local elevation end-point potentials from simulations biased to fit multiple types (shown in *black*, *red*, *green* and *blue*) of ^3J-values reporting on the same dihedral angle. *Left* a case in which even allowing for conformational averaging, it is not possible to find a set of dihedral angle values in keeping with all four ^3J-values, so that the local elevation potentials clash with one another and continue to build throughout the simulation. *Right* a case where after an initial build-up period in which the dihedral angle is biased towards values that on average satisfy the data, the potentials remain static for the remainder of the simulation (Allison and van Gunsteren 2009)

or a randomly excluded subset of the restraint data, as well as of the restraint data, is assessed. Satisfaction of the restraints generally improves with more degrees of freedom, whereas reproduction of the free data becomes worse.

Instantaneous restraining

The most basic restraining method is instantaneous restraining, in which a single copy of the protein is required to fit all of the data simultaneously at each point in time. Whilst this method is widely used, particularly in protein structure determination, it ignores the average nature of NMR data. Even if a protein only samples two different conformations, the structure produced by enforcing the restraints on a single copy will not only be irrelevant, but may well be physically impossible (Kessler et al. 1988). The rarity of examples in which instantaneous restraining fails simply highlights the lack of NMR data relative to the number of degrees of freedom ($N^{\text{obs}} \ll N^{\text{df}}$). Additionally, data pertaining to mobile regions are often omitted. Given that a special feature of NMR data is that they can be measured for dynamic states, it is desirable to account for conformational averaging in the restraining methodology. There are two possible means of doing so: ensemble-averaging (Scheek et al. 1991; Kemmink et al. 1993; Bonvin et al. 1994; Mierke et al. 1994) and time-averaging (Torda et al. 1989, 1990).

Ensemble averaging

In ensemble-averaged restraining, multiple independent replicas are simulated in parallel. At each integration step, the quantity to be restrained is back-calculated from each replica and then averaged over the N^{rep} replicas before comparison with the experimental value:

$$f_i^{\text{calc}}(\mathbf{r}(t)) = \frac{1}{N^{\text{rep}}} \sum_{k=1}^{N^{\text{rep}}} \left(f_{i,k}^{\text{calc}}(\mathbf{r}(t))^x \right)^{1/x}. \quad (8)$$

The value of x depends on the type of data and whether the observable or the related structural property is restrained. For ^3J-couplings, the non-linearity is already accounted for by use of the Karplus relation to back-calculate the couplings from the protein coordinates, so $x=1$, whereas for NOE- and PRE-derived distances, which are structural properties, $x=-3$ or -6.

Ensemble-averaging accounts, to some extent, for the average nature of NMR data, although the number of replicas can never be as large as the number of different conformations over which the measured NMR data are averaged, because of the limited amount of experimental data. Averaging over too many replicas results in underfitting, where the ensemble is primarily determined by the force field. A more elaborate version of ensemble-averaging that attempts to correct for this is the MUMO method (Richter et al. 2007), in which each type of data can be restrained across a different number of replicas. This

provides a means by which to represent the different time- and spatial-scales over which each type of data are averaged, but thus far MUMO has only been used in simulated annealing-based structure refinement (Richter et al. 2007; Fenwick et al. 2011).

A further problem with ensemble-averaging is that, although the structures sampled in each individual trajectory will be correctly Boltzmann-weighted, the ensemble of structures drawn from the different trajectories over which the average is calculated at each point in time is not necessarily Boltzmann-weighted. A possible solution to this in which the probability of each replica was treated as a parameter in the simulation was explored, but found to suffer from a number of drawbacks of its own (Fennen et al. 1995).

Despite the aforementioned issues, ensemble-averaging has been used to generate ensembles of structures representative of the dynamics of folded proteins (Hess and Scheek 2003; Best and Vendruscolo 2004; Clore and Schwieters 2004; Lindorff-Larsen et al. 2005a; Clore and Schwieters 2006; De Simone et al. 2009; Robustelli et al. 2010), the variability of the transition-state for folding (Vendruscolo et al. 2001; Paci et al. 2002, 2004; Lindorff-Larsen et al. 2004), molten-globule (Paci et al. 2005) and disordered states (Dedmon et al. 2005; Kristjansdottir et al. 2005; Francis et al. 2006; Allison et al. 2009; Esteban-Martín et al. 2010) of proteins, the conformational equilibrium between major substates (Camilloni et al. 2012) and transiently populated excited states (Robustelli et al. 2010) using NMR and other types of experimental data.

Time-averaging

Another means of accounting for the average nature of NMR data is time-averaging (Torda et al. 1989, 1990, 1993; Bonvin et al. 1994). A single copy of the protein is simulated, and an exponentially-weighted time-average is compared to the experimental observable:

$$\overline{f_i^{\mathrm{calc}}(\mathbf{r}(t))} = \frac{1}{\tau}\frac{1}{1-\exp(t/\tau)} \times \int_0^t \exp\left(-\frac{t-t'}{\tau}\right) f_i^{\mathrm{calc}}(\mathbf{r}(t'))dt', \quad (9)$$

where τ is the memory relaxation time. In this way, more recent structures contribute most to the average, avoiding $\overline{f_i^{\mathrm{calc}}(\mathbf{r}(t))}$ becoming less sensitive to instantaneous fluctuations for increasing values of t (Scott et al. 1998). Time-averaging is computationally more efficient than ensemble-averaging, as only a single copy of the protein needs to be simulated. However the time-scale encompassed by the average is generally much shorter than the time-scale over which the experimental data were averaged, with τ typically on the order of 5–10 ps. This means that the structures contributing to the time-average are likely to be highly correlated, in comparison to ensemble-averaging, where each replica is independent. Thus, whilst time-averaging is efficient at capturing fluctuations up to the time-scales of the averaging time (Hess and Scheek 2003), ensemble-averaging may be a better approach when multiple rather different conformations contribute to the average, as it allows them to co-exist simultaneously (Fennen et al. 1995).

One problem with time-averaging is that $\overline{f_i^{\mathrm{calc}}(\mathbf{r}(t))}$ lags behind $f_i^{\mathrm{calc}}(\mathbf{r}(t))$, thus even if the current conformation of the protein is in keeping with the experimental data, the time-average will not be and the protein will continue to experience a restraining force. This issue is particularly important in the case of ^3J-values. Alternative functional forms for the restraint term have been proposed to overcome this problem. One possibility is to use a biquadratic restraining function (Scott et al. 1998; Christen et al. 2007):

$$V_i^{\mathrm{restr}}(\mathbf{r}(t)) = \frac{1}{2}K_i^{\mathrm{restr}}\left(f_i^{\mathrm{obs}} - \overline{f_i^{\mathrm{calc}}(\mathbf{r}(t))}\right)^2\left(f_i^{\mathrm{obs}} - f_i^{\mathrm{calc}}(\mathbf{r}(t))\right)^2. \quad (10)$$

In this way, the energy penalty will be zero if either $f_i^{\mathrm{calc}}(\mathbf{r}(t))$ or $\overline{f_i^{\mathrm{calc}}(\mathbf{r}(t))}$ agrees with f_i^{obs}.

Rather than weight the time-average and instantaneous values equally, an elliptic restraining function, incorporating a linear combination $\overline{f_i^{\mathrm{calc}}(\mathbf{r}(t))}$ and $f_i^{\mathrm{calc}}(\mathbf{r}(t))$ with relative weights specified by the mixing parameter $A_{\mathrm{ell}} \in [0 \ldots 1]$, may be used (Scott et al. 1998):

$$V_i^{\mathrm{restr}}(\mathbf{r}(t)) = \frac{1}{2}K_i^{\mathrm{restr}}\left[A_{\mathrm{ell}}\overline{f_i^{\mathrm{calc}}(\mathbf{r}(t))} + (1-A_{\mathrm{ell}})f_i^{\mathrm{calc}}(\mathbf{r}(t)) - f_i^{\mathrm{obs}}\right]^2 \quad (11)$$

This method allows the relative weights of $\overline{f_i^{\mathrm{calc}}(\mathbf{r}(t))}$ and $f_i^{\mathrm{calc}}(\mathbf{r}(t))$ to be adjusted by the user.

The particular characteristics of inversion of the Karplus relation inspired further adaptation of the restraining function specifically for ^3J-couplings by introduction of an oscillating factor $\cos^2(\omega^{\mathrm{restr}}t)$ that scales the torsion angle potential function V_i^{tors} of the dihedral angle associated with the i^{th} ^3J-coupling as well as the restraining function V_i^{restr} for a period $\tau^{\mathrm{restr}} = \frac{\pi}{\omega^{\mathrm{restr}}}$ (Keller et al. 2007). The oscillating factor is switched on when the average ^3J-value deviates more than a certain threshold ΔJ from the experimental value, temporarily scaling down the potential energy for the dihedral angle, allowing it to escape from local minima.

A related method is local elevation biasing (Christen et al. 2007), based on the local elevation enhanced searching method (Huber et al. 1994). The simulation is driven away from structures that do not fit the experimental data by

adding Gaussian functions to a biasing coordinate, usually the structural property related to the NMR observable. For ^3J-couplings, for example, the biasing coordinate is the dihedral angle θ upon which the couplings report. The range of values of θ is divided into N^{le} bins with midpoints $\theta^0_{i,k}$. At each step, Gaussian functions centered on $\theta^0_{i,k}$ are added if $\overline{f_i^{\mathrm{calc}}(\mathbf{r}(t))}$ and $f_i^{\mathrm{calc}}(\mathbf{r}(t))$ do not match f_i^{obs}:

$$V^{\mathrm{le}}_{i,k}(\mathbf{r}(t)) = K^{\mathrm{le}}_i w_{\theta_{i,k}}(\mathbf{r}(t)) \times \exp\left(-\left(\theta_i(\mathbf{r}(t)) - \theta^0_{i,k}\right)^2 / 2\left(\Delta\theta^0\right)^2\right). \quad (12)$$

K^{le}_i is typically orders of magnitude lower than the force constants used in Eqs. 7, 9, 10 and 11. The weight of the k^{th} penalty term $w_{\theta_{i,k}}(r(t))$ is calculated according to

$$w_{\theta_{i,k}}(\mathbf{r}(t)) = t^{-1}\int_0^t \delta_{\theta_i}(\mathbf{r}(t'))\theta^0_{i,k} V_i^{\mathrm{biq}}(\mathbf{r}(t'))dt', \quad (13)$$

where $V^{\mathrm{biq}}(\mathbf{r}(t'))$ is calculated in the same manner as $V_i^{\mathrm{restr}}(\mathbf{r}(t))$ in Eq. 10 and δ is

$$\delta_{\theta_i}(\mathbf{r}(t))\theta^0_{i,k} = \begin{cases} 1 \; if \; \theta^0_{i,k} - \Delta\theta^0/2 \le \theta_i(\mathbf{r}(t)) < \theta^0_{i,k} + \Delta\theta^0/2 \\ 0 \; otherwise. \end{cases} \quad (14)$$

Local elevation biasing greatly enhances sampling and has been shown to produce superior agreement with experimental data comprising side-chain ^3J-couplings compared to conventional instantaneous and time-averaged biquadratic restraining methods (Christen et al. 2007; Allison and van Gunsteren 2009; Dolenc et al. 2010; Missimer et al. 2010).

Conclusions

The speed, accuracy and length of MD simulations of proteins have increased markedly over the years, but it remains essential to check the quality of the structures produced and the extent of sampling by comparison with quantitative experimental data. Solution NMR provides a number of observables that can be directly linked to structural properties of proteins and that encompass dynamics on different time-scales. Modern force fields produce ensembles of structures that, on the whole, agree reasonably well with NMR data, particularly when enhanced sampling methods are used. The recent advent of specialised hardware and highly parallelised software capable of simulations of similar lengths to the time-scales encompassed by NMR data has revitalised the field whilst revealing that there is still work to be done. The alternative, biasing or restraining MD simulations to fit experimental data, has also undergone much development over the years, with methods that include enhanced sampling methods showing particular promise. Such methods will always be limited, however, by the need to balance over- and under-fitting.

In the future, the deficiencies in atomic-level protein force fields that are beginning to be exposed by the longer simulations now possible need to be addressed. One possible means of doing so is to utilise the memory of the alterations made to the potential energy surface to improve the agreement between back-calculated and experimentally-measured observables when using techniques such as AMD and local elevation to guide the modification of force field terms. This will involve carrying out such simulations and comparisons with experimental data for a wide range of different proteins, however, to ensure that any changes made to force fields are not specific to a particular system. It would also be desirable to see measurement of more NMR data to improve the ratio of $N^{\mathrm{obs}}/N^{\mathrm{df}}$, allow conformational motion to be better accounted for, and increase the potential for cross-validation. Finally, increasing computer power and parallelisation should allow more accurate, perhaps even quantum-chemical, relationships between experimental observables and structural properties to be routinely applied for macromolecules such as proteins.

Acknowledgements I thank the very many people with whom I have discussed the issues involved in combining NMR data with MD simulations, in particular, Prof. Wilfred van Gunsteren, to whom this article is dedicated on the occasion of his 65th birthday.

References

Allison JR, van Gunsteren WF (2009) A method to explore protein side chain conformational variability using experimental data. Chem Phys Chem 10(18):3213–3228

Allison JR, Varnai P, Dobson CM, Vendruscolo M (2009) Determination of the free energy landscape of α-synuclein using spin label nuclear magnetic resonance measurements. J Am Chem Soc 131 (51):18,314–18,326

Allison JR, Bergeler M, Hansen N, van Gunsteren WF (2011) Current computer modeling cannot explain why two highly similar sequences fold into different structures. Biochemistry 50 (50):10,965–10,973

Allison JR, Hertig S, Missimer JH, Smith LJ, Steinmetz MO, Dolenc J (2012) Probing the structure and dynamics of proteins by combining molecular dynamics simulations and experimental NMR data. J Chem Theory Comput. doi:10.1021/ct300393b

Andronesi OC, Becker S, Seidel K, Heise H, Young HS, Baldus M (2005) Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy. J Am Chem Soc 127(37):12,965–12,974

Battiste JL, Wagner G (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear Overhauser effect data. Biochemistry 39(18):5355–5365

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

Beauchamp KA, Lin YS, Das R, Pande VS (2012) Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements. J Chem Theory Comput 8(4):1409–1414

Best RB, Hummer G (2009) Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. J Phys Chem B 113(26):9004–9015

Best RB, Vendruscolo M (2004) Determination of protein structures consistent with NMR order parameters. J Am Chem Soc 126 (26):8090–8091

Blackledge M (2005) Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. Prog Nucl Magn Reson Spectrosc 46(1):23–61

Boelens R, Koning TMG, Kaptein R (1988) Determination of biomolecular structures from proton-proton NOE's using a relaxation matrix approach. J Mol Struct 173:299–311

Bonvin AM, Brünger AT (1995) Conformational variability of solution nuclear magnetic resonance structures. J Mol Biol 250(1):80–93

Bonvin AMJJ, Boelens R, Kaptein R (1994) Time- and ensemble-averaged direct NOE restraints. J Biomol NMR 4(1):143–149

Borgias BA, James TL (1990) MARDIGRAS - a procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure. J Magn Reson 87(3):475–487

Borgias BA, Gochin M, Kerwood DJ, James TL (1990) Relaxation matrix analysis of 2D NMR data. Prog Nucl Magn Reson Spectrosc 22(1):83–100

Brown WM, Wang P, Plimpton SJ, Tharrington AN (2011) Implementing molecular dynamics on hybrid high performance computers - short range forces. Comp Phys Comm 182(4):898–911

Brünger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355(6359):472–475

Brünger AT (1993) Assessment of phase accuracy by cross validation: the free R value. Methods and applications. Acta Crystallogr D Biol Crystallogr 49(Pt 1):24–36

Brünger AT, Clore GM, Gronenborn AM, Saffrich R, Nilges M (1993) Assessing the quality of solution nuclear magnetic resonance structures by complete cross-validation. Science 261(5119):328–331

Brüschweiler R, Case DA (1994) Adding harmonic motion to the Karplus relation for spin-spin coupling. J Am Chem Soc 116 (24):11,199–11,200

Brüschweiler R, Roux B, Blackledge M, Griesinger C, Karplus M, Ernst RR (1992) Influence of rapid intramolecular motion on NMR cross-relaxation rates. A moleculardynamics study of antamanide in solution. J Am Chem Soc 114(7):2289–2302

Bryant RG (1983) The NMR time scale. J Chem Edu 60(11):933

Buck M, Bouguet-Bonnet S, Pastor RW, MacKerell AD (2006) Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. Biophys J 90(4):L36–L38

Bühl M, van Mourik T (2011) NMR spectroscopy: quantum-chemical calculations. WIREs: Comput Mol Sci 1(4):634–647

Bürgi R, Pitera J, van Gunsteren WF (2001) Assessing the effect of conformational averaging on the measured values of observables. J Biomol NMR 19(4):305–320

Camilloni C, Robustelli P, Simone AD, Cavalli A, Vendruscolo M (2012) Characterization of the conformational equilibrium between the two major substates of RNase A using NMR chemical shifts. J Am Chem Soc 134(9):3968–3971

Case DA (1995) Calibration of ring-current effects in proteins and nucleic acids. J Biomol NMR 6(4):341–346

Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, Oschkinat H (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. Nature 420 (6911):98–102

Chandrasekhar I, Clore GM, Szabo A, Gronenborn AM, Brooks BR (1992) A 500 ps molecular dynamics simulation study of interleukin-1β in water: correlation with nuclear magnetic resonance spectroscopy and crystallography. J Mol Biol 226 (1):239–250

Choy WY, Mulder FA, Crowhurst KA, Muhandiram DR, Millett IS, Doniach S, Forman-Kay JD, Kay LE (2002) Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. J Mol Biol 316(1):101–112

Christen M, Keller B, van Gunsteren WF (2007) Biomolecular structure refinement based on adaptive restraints using local-elevation simulation. J Biomol NMR 39(4):265–273

Clore GM, Iwahara J (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. Chem Rev 109(9):4108–4139

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(9):2923–2938

Clore GM, Schwieters CD (2006) Concordance of residual dipolar couplings, backbone order parameters and crystallographic *B*-factors for a small α/β protein: a unified picture of high probability, fast atomic motions in proteins. J Mol Biol 355(5):879–886

Clore G, Starich M, Gronenborn A (1998) Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. J Am Chem Soc 120(40):10,571–10,572

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13(3):289–302

Daura X, Antes I, van Gunsteren WF, Thiel W, Mark AE (1999) The effect of motional averaging on the calculation of NMR-derived structural properties. Proteins Struct Funct Bioinform 36(4):542–555

Daura X, Gademann K, Schäfer H, Jaun B, Seebach D, van Gunsteren WF (2001) The β-peptide hairpin in solution: conformational study of a β-hexapeptide in methanol by NMR spectroscopy and MD simulation. J Am Chem Soc 123(10):2393–2404

de Dios A, Pearson J, Oldfield E (1993) Secondary and tertiary structural effects on protein NMR chemical shifts: an ab initio approach. Science 260(5113):1491–1496

De Simone A, Richter B, Salvatella X, Vendruscolo M (2009) Toward an accurate determination of free energy landscapes in solution states of proteins. J Am Chem Soc 131(11):3810–3811

Dedmon MM, Lindorff-Larsen K, Christodoulou J, Vendruscolo M, Dobson CM (2005) Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. J Am Chem Soc 127(2):476–477

Dolenc J, Missimer JH, Steinmetz MO, van Gunsteren WF (2010) Methods of NMR structure refinement: molecular dynamics simulations improve the agreement with measured NMR data of a C-terminal peptide of GCN4-p1. J Biomol NMR 47(3):221–235

Donaldson LW, Skrynnikov NR, Choy WY, Muhandiram DR, Sarkar B, Forman-Kay JD, Kay LE (2001) Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy. J Am Chem Soc 123(40):9843–9847

Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24(16):1999–2012

Edmondson SP (1992) NOE R-factors and structural refinement using FIRM, an iterative relaxation matrix program. J Magn Reson 98 (2):283–298

Eghbalnia HR, Wang L, Bahrami A, Assadi A, Markley J (2005) Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. J Biomol NMR 32(1):71–81

Eichenberger AP, Smith LJ, van Gunsteren WF (2012) Ester-linked hen egg white lysozyme shows a compact fold in a molecular dynamics simulation - possible causes and sensitivity of experimentally observable quantities to structural changes maintaining this compact fold. FEBS J 279(2):299–315

Eriksson MAL, Berglund H, Härd T, Nilsson L (1993) A comparison of ^{15}N NMR relaxation measurements with a molecular dynamics simulation: backbone dynamics of the glucocorticoid receptor DNA-binding domain. Proteins Struct Funct Bioinform 17 (4):375–390

Esteban-Martín S, Fenwick RB, Salvatella X (2010) Refinement of ensembles describing unstructured proteins using NMR residual dipolar couplings. J Am Chem Soc 132(13):4626–4632

Fawzi NL, Phillips AH, Ruscio JZ, Doucleff M, Wemmer DE, Head-Gordon T (2008) Structure and dynamics of the Aβ21-30 peptide from the interplay of NMR experiments and molecular simulations. J Am Chem Soc 130(19):6145–6158

Fennen J, Torda AE, van Gunsteren WF (1995) Structure refinement with molecular dynamics and a Boltzmann-weighted ensemble. J Biomol NMR 6(2):163–70

Fenwick RB, Esteban-Martíın S, Richter B, Lee D, Walter KFA, Milovanovic D, Becker S, Lakomek NA, Griesinger C, Salvatella X (2011) Weak long-range correlated motions in a surface patch of ubiquitin involved in molecular recognition. J Am Chem Soc 133(27):10,336–10,339

Francis C, Lindorff-Larsen K, Best RB, Vendruscolo M (2006) Characterization of the residual structure in the unfolded state of the Δ131Δ fragment of staphylococcal nuclease. Proteins Struct Funct Bioinform 65(1):145–152

Frank A, Möller HM, Exner TE (2012) Toward the quantum chemical calculation of NMR chemical shifts of proteins. 2. Level of theory, basis set, and solvents model dependence. J Chem Theory Comput 8(4):1480–1492

Gillespie JR, Shortle D (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. J Mol Biol 268 (1):158–169

Glättli A, van Gunsteren WF (2004) Are NMR-derived model structures for β-peptides representative for the ensemble of structures adopted in solution? Angew Chem Int Ed 43(46):6312–6316

Haasnoot CAG, de Leeuw FAAM, de Leeuw HPM, Altona C (1979) Interpretation of vicinal proton-proton coupling constants by a generalized Karplus relation. conformational analysis of the exocyclic C4'-C5' bond in nucleosides and nucleotides: Preliminary communication. Recl Trav Chim Pays-Bas 98(12):576–577

Haasnoot CAG, De Leeuw FAAM, De Leeuw HPM, Altona C, (1981) Relationship between proton-proton NMR coupling constants and substituent electronegativities. III. Conformational analysis of proline rings in solution using a generalized Karplus equation. Biopolymers 20(6):1211–1245

Habeck M, Rieping W, Nilges M (2005) Bayesian estimation of Karplus parameters and torsion angles from three-bond scalar couplings constants. J Magn Reson 177(1):160–165

Hamelberg D, Mongan J, McCammon JA (2004) Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. J Chem Phys 120(24):11,919–11,929

Han B, Liu Y, Ginzinger S, Wishart D (2011) SHIFTX2: significantly improved protein chemical shift prediction. J Biomol NMR 50 (1):43–57

Hansen MR, Mueller L, Pardi A (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat Struct Mol Biol 5(12):1065–1074

Hess B, Scheek RM (2003) Orientation restraints in molecular dynamics simulations using time and ensemble averaging. J Mag Reson 164(1):19–27

Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins Struct Funct Bioinform 65(3):712–725

Huber T, Torda AE, van Gunsteren WF (1994) Local elevation: A method for improving the searching properties of molecular dynamics simulation. J Comput Aided Mol Des 8(6):695–708

Imai K, Osawa E (1990) An empirical extension of the Karplus equation. Magn Reson Chem 28(8):668–674

Kalk A, Berendsen HJC (1976) Proton magnetic relaxation and spin diffusion in proteins. J Magn Reson 24(3):343–366

Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B 105(28):6474–6487

Karplus M (1959) Contact electron-spin coupling of nuclear magnetic moments. J Chem Phys 30(1):11–15

Keepers JW, James TL (1984) A theoretical study of distance determinations from NMR. Two-dimensional nuclear Overhauser effect spectra. J Magn Reson 57(3):404–426

Keller B, Christen M, Oostenbrink C, van Gunsteren WF (2007) On using oscillating time-dependent restraints in MD simulation. J Biomol NMR 37(1):1–14

Kemmink J, van Mierlo CPM, Scheek RM, Creighton TE (1993) Local structure due to an aromatic-amide interaction observed by ^{1}H-nuclear magnetic resonance spectroscopy in peptides related to the N terminus of bovine pancreatic trypsin inhibitor. J Mol Biol 230(1):312–322

Kessler H, Griesinger C, Lautz J, Mueller A, van Gunsteren WF, Berendsen HJC (1988) Conformational dynamics detected by nuclear magnetic resonance NOE values and J coupling constants. J Am Chem Soc 110(11):3393–3396

Kjaergaard M, Poulsen FM (2012) Disordered proteins studied by chemical shifts. Prog Nucl Magn Reson Spectrosc 60:42–51

Klepeis JL, Lindorff-Larsen K, Dror RO, Shaw DE (2009) Long-timescale molecular dynamics simulations of protein structure and function. Curr Opin Struct Biol 19(2):120–127

Koenig B, Hu JS, Ottiger M, Bose S, Hendler R, Bax A (1999) NMR measurement of dipolar couplings in proteins aligned by transient binding to purple membrane fragments. J Am Chem Soc 121 (6):1385–1386

Koerdel J, Teleman O (1992) Backbone dynamics of calbindin D_{9k}: comparison of molecular dynamics simulations and ^{15}N NMR relaxation measurements. J Am Chem Soc 114(12):4934–4936

Kohlhoff KJ, Robustelli P, Cavalli A, Salvatella X, Vendruscolo M (2009) Fast and accurate predictions of protein NMR chemical shifts from interatomic distances. J Am Chem Soc 131 (39):13,894–13,895

Kristjansdottir S, Lindorff-Larsen K, Fieber W, Dobson CM, Vendruscolo M, Poulsen FM (2005) Formation of native and non-native interactions in ensembles of denatured ACBP molecules from paramagnetic relaxation enhancement studies. J Mol Biol 347(5):1053–1062

Lange OF, van der Spoel D, de Groot BL (2010) Scrutinizing molecular mechanics force fields on the submicrosecond timescale with NMR data. Biophys J 99(2):647–655

Leeflang B, Kroon-Batenburg L (1992) CROSREL: Full relaxation matrix analysis for NOESY and ROESY NMR spectroscopy. J Biomol NMR 2(5):495–518

Li DW, Brüschweiler R (2009) Certification of molecular dynamics trajectories with NMR chemical shifts. J Phys Chem Lett 1(1):246–248

Lindorff-Larsen K, Vendruscolo M, Paci E, Dobson CM (2004) Transition states for protein folding have native topologies despite high structural variability. Nat Struct Mol Biol 11(5):443–449

Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M (2005a) Simultaneous determination of protein structure and dynamics. Nature 433(7022):128–132

Lindorff-Larsen K, Best RB, Vendruscolo M (2005b) Interpreting dynamically-averaged scalar couplings in proteins. J Biomol NMR 32(4):273–280

Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins Struct Funct Bioinform 78(8):1950–1958

Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE (2012a) Systematic validation of protein force fields against experimental data. PLoS One 7(2):e32,131

Lindorff-Larsen K, Trbovic N, Maragakis P, Piana S, Shaw DE (2012b) Structure and dynamics of an unfolded protein examined by molecular dynamics simulation. J Am Chem Soc 134(8):3787–3791

Linge JP, Habeck M, Rieping W, Nilges M (2004) Correction of spin diffusion during iterative automated NOE assignment. J Magn Reson 167(2):334–342

Louhivuori M, Otten R, Lindorff-Larsen K, Annila A (2006) Conformational fluctuations affect protein alignment in dilute liquid crystal media. J Am Chem Soc 128(13):4371–4376

Louhivuori M, Otten R, Salminen T, Annila A (2007) Evidence of molecular alignment fluctuations in aqueous dilute liquid crystalline media. J Biomol NMR 39(2):141–152

Mackerell ADJ (2004) Empirical force fields for biological macromolecules: overview and issues. J Comput Chem 25(13):1584–1604

Mackerell AD, Bashford D, Bellott RL, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, 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(18):3586–3616

Mackerell AD, Feig M, Brooks CL (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J Comput Chem 25(11):1400–1415

Manolikas T, Herrmann T, Meier BH (2008) Protein structure determination from ^{13}C spin-diffusion solid-state NMR spectroscopy. J Am Chem Soc 130(12):3959–3966

Maragakis P, Lindorff-Larsen K, Eastwood MP, Dror RO, Klepeis JL, Arkin IT, Jensen MO, Xu H, Trbovic N, Friesner RA, Palmer AG, Shaw DE (2008) Microsecond molecular dynamics simulation shows effect of slow loop dynamics on backbone amide order parameters of proteins. J Phys Chem B 112(19):6155–6158

Marion D, Genest M, Ptak M (1987) Reconstruction of NOESY maps: a requirement for a reliable conformational analysis of biomolecules using 2D NMR. Biophys Chem 28(3):235–244

Markwick PRL, Sprangers R, Sattler M (2002) Dynamic effects on J-couplings across hydrogen bonds in proteins. J Am Chem Soc 125(3):644–645

Markwick PRL, Bouvignies G, Blackledge M (2007) Exploring multiple timescale motions in protein GB3 using accelerated molecular dynamics and NMR spectroscopy. J Am Chem Soc 129(15):4724–4730

Markwick PRL, Bouvignies G, Salmon L, McCammon JA, Nilges M, Blackledge M (2009) Toward a unified representation of protein structural dynamics in solution. J Am Chem Soc 131(46):16,968–16,975

Markwick PRL, Cervantes CF, Abel BL, Komives EA, Blackledge M, McCammon JA (2010) Enhanced conformational space sampling improves the prediction of chemical shifts in proteins. J Am Chem Soc 132(4):1220–1221

Marsh JA, Singh VK, Jia Z, Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: implications for fibrillation. Prot Sci 15 (12):2795–2804

McCammon J, Gelin B, Karplus M (1977) Dynamics of folded proteins. Nature 267:585–590

Meiler J (2003) PROSHIFT: protein chemical shift prediction using artificial neural networks. J Biomol NMR 26(1):25–37

Mierke DF, Scheek RM, Kessler H (1994) Coupling constants as restraints in ensemble distance driven dynamics. Biopolymers 34(4):559–563

Missimer JH, Dolenc J, Steinmetz MO, van Gunsteren WF (2010) Exploring the trigger sequence of the GCN4 coiled-coil: biased molecular dynamics resolves apparent inconsistencies in NMR measurements. Prot Sci 19(12):2462–2474

Mulder FAA, Filatov M (2010) NMR chemical shift data and ab initio shielding calculations: emerging tools for protein structure determination. Chem Soc Rev 39(2):578–590

Nanzer AP, Torda AE, Bisang C, Weber C, Robinson JA, van Gunsteren WF (1997) Dynamical studies of peptide motifs in the Plasmodium falciparum circumsporozoite surface protein by restrained and unrestrained MD simulations. J Mol Biol 267 (4):1012–1025

Neal S, Nip AM, Zhang H, Wishart DS (2003) Rapid and accurate calculation of protein ^{1}H, ^{13}C and ^{15}N chemical shifts. J Biomol NMR 26(3):215–240

Nielsen JT, Eghbalnia HR, Nielsen NC (2012) Chemical shift prediction for protein structure calculation and quality assessment using an optimally parameterized force field. Prog Nucl Magn Reson Spectrosc 60:1–28

Niggli DA, Ebert MO, Lin Z, Seebach D, van Gunsteren WF (2012) Helical content of a β3-octapeptide in methanol: molecular dynamics simulations explain a seemingdiscrepancy between conclusions derived from CD and NMR data. Chem Eur J 18(2):586–593

Oldfield E (2002) Chemical shifts in amino acids, peptides, and proteins: from quantum chemistry to drug design. Annu Rev Phys Chem 53(1):349–378

Osapay K, Case DA (1991) A new analysis of proton chemical shifts in proteins. J Am Chem Soc 113(25):9436–9444

Paci E, Vendruscolo M, Dobson CM, Karplus M (2002) Determination of a transition state at atomic resolution from protein engineering data. J Mol Biol 324(1):151–163

Paci E, Friel CT, Lindorff-Larsen K, Radford SE, Karplus M, Vendruscolo M (2004) Comparison of the transition state ensembles for folding of Im7 and Im9 determined using all-atom molecular dynamics simulations with φ value restraints. Proteins Struct Funct Bioinform 54 (3):513–525

Paci E, Greene LH, Jones RM, Smith LJ (2005) Characterization of the molten globule state of retinol-binding protein using a molecular dynamics simulation approach. FEBS J 272(18):4826–4838

Palmer AG, Case DA (1992) Molecular dynamics analysis of NMR relaxation in a zinc-finger peptide. J Am Chem Soc 114 (23):9059–9067

Pardi A, Billeter M, Wüthrich K (1984) Calibration of the angular dependence of the amide proton-Cα proton coupling constants, $^{3}J_{HN\alpha}$, in a globular protein: use of $^{3}J_{HN\alpha}$ for identification of helical secondary structure. J Mol Biol 180(3):741–751

Pastore A, Saudek V (1990) The relationship between chemical shift and secondary structure in proteins. J Magn Reson 90(1):165–176

Pérez C, Löhr F, Rüterjans H, Schmidt J (2001) Self-consistent Karplus parametrization of ^{3}J-couplings depending on the polypeptide side-chain torsion χ_1. J Am Chem Soc 123(29):7081–7093

Peter C, Rueping M, Wörner HJ, Jaun B, Seebach D, van Gunsteren WF (2003) Molecular dynamics simulations of small peptides: can one derive conformational preferences from ROESY spectra? Chem Eur J 9(23):5838–5849

Peti W, Smith L, Redfield C, Schwalbe H (2001) Chemical shifts in denatured proteins: Resonance assignments for denatured

ubiquitin and comparisons with other denatured proteins. J Biomol NMR 19(2):153–165
Philippopoulos M, Mandel AM, Palmer AG, Lim C (1997) Accuracy and precision of NMR relaxation experiments and MD simulations for characterizing protein dynamics. Proteins: Struct Funct Bioinform 28(4):481–493
Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26(16):1781–1802
Piana S, Lindorff-Larsen K, Shaw D (2011) How robust are protein folding simulations with respect to force field parameterization? Biophys J 100(9):L47–L49
Plimpton S (1995) Fast parallel algorithms for short-range molecular dynamics. J Comp Phys 117(1):1–19
Post CB, Meadows RP, Gorenstein DG (1990) On the evaluation of interproton distances for three-dimensional structure determination by NMR using a relaxation rate matrix analysis. J Am Chem Soc 112(19):6796–6803
Richter B, Gsponer J, Várnai P, Salvatella X, Vendruscolo M (2007) The MUMO (minimal under-restraining minimal over-restraining) method for the determination of native state ensembles of proteins. J Biomol NMR 37(2):117–135
Robustelli P, Kohlhoff K, Cavalli A, Vendruscolo M (2010) Using NMR chemical shifts as structural restraints in molecular dynamics simulations of proteins. Structure 18(8):923–933
Robustelli P, Stafford KA, Palmer AG (2012) Interpreting protein structural dynamics from NMR chemical shifts. J Am Chem Soc 134(14):6365–6374
Sanders CR, Hare BJ, Howard KP, Prestegard JH (1994) Magnetically-oriented phospholipid micelles as a tool for the study of membrane-associated molecules. Prog Nucl Magn Reson Spectrosc 26(Part 5):421–444
Sass J, Cordier F, Hoffmann A, Rogowski M, Cousin A, Omichinski J, Lowen H, Grzesiek S (1999) Purple membrane induced alignment of biological macromolecules in the magnetic field. J Am Chem Soc 121(10):2047–2055
Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S (2000) Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. J Biomol NMR 18(4):303–309
Saupe A (1968) Recent results in the field of liquid crystals. Angew Chem Int Ed 7(2):97–112
Scheek R, Torda A, Kemmink J, van Gunsteren W (1991) Computational aspects of the study of biological macromolecules by NMR, NATO ASI Series A225. Plenum Press, New York, pp 209–217
Schmid N, Bötschi M, Van Gunsteren WF (2010) A GPU solvent-solvent interaction calculation accelerator for biomolecular simulations using the GROMOS software. J Comput Chem 31(8):1636–1643
Schmid N, Eichenberger A, Choutko A, Riniker S, Winger M, Mark A, van Gunsteren W (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7. Eur Biophys J 40(7):843–856
Schmidt J (2007) Asymmetric Karplus curves for the protein side-chain ^{3}J couplings. J Biomol NMR 37(4):287–301
Schmidt JM, Blümel M, Löhr F, Rüterjans H (1999) Self-consistent ^{3}J-coupling analysis for the joint calibration of Karplus coefficients and evaluation of torsion angles. J Biomol NMR 14(1):1–12
Schwalbe H, Grimshaw SB, Spencer A, Buck M, Boyd J, Dobson CM, Red-field C, Smith LJ (2001) A refined solution structure of hen lysozyme determined using residual dipolar coupling data. Prot Sci 10(4):677–688
Scott W, Mark A, van Gunsteren WF (1998) On using time-averaging restraints in molecular dynamics simulation. J Biomol NMR 12:501–508
Shaw DE, Dror RO, Salmon JK, Grossman J, Mackenzie KM, Bank JA, Young C, Deneroff MM, Batson B, Bowers KJ, Chow E, Eastwood MP, Ierardi DJ, Klepeis JL, Kuskin JS, Larson RH, Lindorff-Larsen K, Maragakis P, Moraes MA, Piana S, Shan Y, Towles B (2009) Millisecond-scale molecular dynamics simulations on anton. In: Proceedings of the Conference on High Performance Computing, Networking, Storage and Analysis (SC09)
Shen Y, Bax A (2007) Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. J Biomol NMR 38(4):289–302
Shen Y, Bax A (2010) SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. J Biomol NMR 48(1):13–22
Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR 44(4):213–223
Showalter SA, Brüschweiler R (2007) Quantitative molecular ensemble interpretation of NMR dipolar couplings without restraints. J Am Chem Soc 129(14):4158–4159
Showalter SA, Johnson E, Rance M, Brüschweiler R (2007) Toward quantitative interpretation of methyl side-chain dynamics from NMR by molecular dynamics simulations. J Am Chem Soc 129 (46):14,146–14,147
Siemer AB, Ritter C, Ernst M, Riek R, Meier BH (2005) High-resolution solid-state NMR spectroscopy of the prion protein HET-s in its amyloid conformation. Angew Chem Int Ed 44(16):2441–2444
Smith LJ, Sutcliffe MJ, Redfield C, Dobson CM (1991) Analysis of φ and χ_1 torsion angles for hen lysozyme in solution from proton NMR spin-spin coupling constants. Biochemistry 30(4):986–996
Smith LJ, Mark AE, Dobson M, van Gunsteren WF (1995) Comparison of MD simulations and NMR experiments for hen lysozyme. Analysis of local fluctuations, cooperative motions, and global changes. Biochemistry 34(34), 10,918-10,931
Smith PE, van Schaik RC, Szyperski T, Wüthrich K, van Gunsteren WF (1995b) Internal mobility of the basic pancreatic trypsin inhibitor in solution: a comparison of NMR spin relaxation measurements and molecular dynamics simulations. J Mol Biol 246 (2):356–365
Solomon I (1955) Relaxation processes in a system of two spins. Phys Rev 99(2):559–565
Spera S, Bax A (1991) Empirical correlation between protein backbone conformation and Cα and Cβ ^{13}C nuclear magnetic resonance chemical shifts. J Am Chem Soc 113(14):5490–5492
Steiner D, Allison JR, van Gunsteren WF (2012) On the calculation of $^{3}J_{\alpha\beta}$ -coupling constants for side chains in proteins. J Biomol NMR. doi:10.1007/s10858-012-9634-5
Stone JE, Hardy DJ, Ufimtsev IS, Schulten K (2010) GPU-accelerated molecular modeling coming of age. J Mol Graph Model 29 (2):116–125
Suardıaz R, García de la Vega JM, Fabián JS, Contreras RH (2007) Theoretical Karplus relationships for vicinal coupling constants around χ_1 in valine. Chem Phys Lett 442(13):119–123
Teilum, K, Kragelund BB, Poulsen FM (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling. J Mol Biol 324(2), 349–57. 0022–2836
Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278(5340):1111–1114
Tong Y, Ji CG, Mei Y, Zhang JZH (2009) Simulation of NMR data reveals that proteins local structures are stabilized by electronic polarization. J Am Chem Soc 131(24):8636–8641
Torda A, Scheek R, van Gunsteren WF (1989) Time-dependent distance restraints in molecular dynamics simulations. Chem Phys Lett 157(4):289–294

Torda AE, Scheek RM, van Gunsteren WF (1990) Time-averaged nuclear overhauser effect distance restraints applied to tendamistat. J Mol Biol 214(1):223–235

Torda AE, Brunne RM, Huber T, Kessler H, van Gunsteren WF (1993) Structure refinement using time-averaged J-coupling constant restraints. J Biomol NMR 3(1):55–66

Trbovic N, Kim B, Friesner RA, Palmer AG (2008) Structural analysis of protein dynamics by MD simulations and NMR spin-relaxation. Proteins Struct Funct Bioinform 71(2):684–694

Tropp J (1980) Dipolar relaxation and nuclear Overhauser effects in nonrigid molecules: the effect of fluctuating internuclear distances. J Chem Phys 72(11):6035–6043

Trzesniak D, Glättli A, Jaun B, van Gunsteren WF (2005) Interpreting NMR data for β-peptides using molecular dynamics simulations. J Am Chem Soc 127(41):14,320–14,329

Tycko R, Blanco F, Ishii Y (2000) Alignment of biopolymers in strained gels: a new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR. J Am Chem Soc 122(38):9340–9341

Van Melckebeke H, Wasmer C, Lange A, Ab E, Loquet A, Bockmann A, Meier BH (2010) Atomic-resolution three-dimensional structure of HET-s(218–289) amyloid fibrils by solid-state NMR spectroscopy. J Am Chem Soc 132(39):13,765–13,775

Vendruscolo M, Paci E, Dobson CM, Karplus M (2001) Three key residues form a critical contact network in a protein folding transition state. Nature 409(6820):641–645

Vögeli B, Ying J, Grishaev A, Bax A (2007) Limits on variations in protein backbone dynamics from precise measurements of scalar couplings. J Am Chem Soc 129(30):9377–9385

Wagner G, Pardi A, Wüthrich K (1983) Hydrogen bond length and proton NMR chemical shifts in proteins. J Am Chem Soc 105(18):5948–5949

Wang Y, Jardetzky O (2002) Investigation of the neighboring residue effects on protein chemical shifts. J Am Chem Soc 124(47):14,075–14,084

Williamson MP, Asakura T (1993) Empirical comparisons of models for chemical-shift calculation in proteins. J Magn Reson B 101(1):63–71

Wishart DS, Sykes BD, Richards FM (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. J Mol Biol 222(2):311–333

Wishart DS, Sykes BD, Richards FM (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry 31(6):1647–1651

Wishart D, Bigam C, Holm A, Hodges R, Sykes B (1995) ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. J Biomol NMR 5(1):67–81

Wrabl JO, Shortle D, Woolf TB (2000) Correlation between changes in nuclear magnetic resonance order parameters and conformational entropy: molecular dynamics simulations of native and denatured staphylococcal nuclease. Proteins Struct Funct Bioinform 38(2):123–133

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

Xu XP, Case D (2001) Automated prediction of ^{15}N, ^{13}Cα, ^{13}Cβ and ^{13}C' chemical shifts in proteins using a density functional database. J Biomol NMR 21(4):321–333

Yao J, Chung J, Eliezer D, Wright PE, Dyson HJ (2001) NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. Biochemistry 40(12):3561–3571

Yip P, Case DA (1989) A new method for refinement of macromolecular structures based on nuclear overhauser effect spectra. J Magn Reson 83(3):643–648

Zagrovic B, van Gunsteren WF (2006) Comparing atomistic simulation data with the NMR experiment: how much can NOEs actually tell us? Proteins Struct Funct Bioinform 63(1):210–8