Structural biology by NMR: structure, dynamics, and interactions.

Phineus R L Markwick, Thérèse Malliavin, Michael Nilges

OPEN ACCESS Freely available online

PLoS COMPUTATIONAL BIOLOGY

Review

Structural Biology by NMR: Structure, Dynamics, and Interactions

Phineus R. L. Markwick, Thérèse Malliavin, Michael Nilges*

Institut Pasteur, Département de Biologie Structurale et Chimie, Unité de Bio-Informatique Structurale, CNRS URA 2185, Paris, France

Abstract: The function of bio-macromolecules is determined by both their 3D structure and conformational dynamics. These molecules are inherently flexible systems displaying a broad range of dynamics on time-scales from picoseconds to seconds. Nuclear Magnetic Resonance (NMR) spectroscopy has emerged as the method of choice for studying both protein structure and dynamics in solution. Typically, NMR experiments are sensitive both to structural features and to dynamics, and hence the measured data contain information on both. Despite major progress in both experimental approaches and computational methods, obtaining a consistent view of structure and dynamics from experimental NMR data remains a challenge. Molecular dynamics simulations have emerged as an indispensable tool in the analysis of NMR data.

Introduction

The function of bio-macromolecules is determined by both their 3D structure and conformational dynamics. These molecules are inherently flexible systems displaying a broad range of dynamics on time-scales from picoseconds to seconds. Nuclear Magnetic Resonance (NMR) spectroscopy has emerged as the method of choice for studying both protein structure and dynamics in solution. The principle behind NMR-based structure determination is to obtain a set of empirical structural parameters, such as inter-atomic distances and dihedral angles, which are implemented in the form of restraints in a molecular modeling algorithm to obtain a representation of the 3D structure of the bio-molecule. The empirical data are acquired from the study of different NMR-based "relaxation channels" or mechanisms that are sensitive to both molecular structure and dynamics. Several recent reviews discuss methods for NMR protein structure calculation [1–5]. The present review focuses on NMR methods that both determine and assess biomolecular structure or dynamics in the most objective way and focuses on the challenges involved in obtaining a consistent view of structure and dynamics from the available experimental NMR data.

3D Structure Determination from NMR Data

The NOE (Nuclear Overhauser Effect), a through-space relaxation mechanism involving the transfer of tranverse magnetisation between local spin-active nuclei, provides information about average inter-atomic distances up to about 6 Å [6], and remains the most important data for structure determination by NMR. More long-range–distance information can be obtained from paramagnetic relaxation [7]. Distance information is supplemented by the measurement of torsion angles from through-bond scalar J-couplings. Residual Dipolar Couplings (RDCs) [8,9] provide additional structural information concerning the orientation of inter-atomic vectors with respect to a reference frame. The principal data types are illustrated in Figure 1. The primary data of any NMR experiments, the chemical shifts (the resonating frequencies of the nuclei), depend on the local magnetic field and hence reflect the local molecular environment [10]. Combined with a library of short fragments of known 3D structure, the chemical shift alone can be sufficient to determine the 3D structure [11,12].

In general, structure determination from NMR relies on data from liquid samples, and it has only recently been demonstrated that structure determination of proteins by solid state NMR is feasible [13,14]. This has potential applications for molecules that neither are soluble nor form three-dimensional crystals easily, such as membrane proteins.

The experimental information on structure and dynamics is intricately mixed. For example, the NOE depends on the distance (a structural parameter, Figure 1) and on the angular fluctuation, which, if the distance is known and fixed, can be used to characterise local dynamics. Structural parameters extracted from NMR data are therefore rather approximate in nature; for the NOE derived distances, the error may be on the order of 2 Å. Due to the difficulties in the interpretation of the data, structure determination and analysis of dynamical information are in general performed independently.

Increasing speed and reliability of NMR structure determination. NMR structure determination still presents some considerable challenges: the method is limited to systems of relatively small molecular mass, data collection times are long, data analysis remains a lengthy procedure, and it is difficult to evaluate the quality of the final structures. These issues are particularly apparent when using NMR in structural genomics projects [15], and advances have been made in all areas. The calculation of a structure itself has become extremely rapid [16], and new labelling methods [17] have significantly improved both spectral quality and automated analysis, whilst rigourous standards and data formats afford compatibilty of different software packages [18].

The most significant advances in efficiency have been gained through automation of data analysis and structure calculation [16,19–21]. An analysis of the network of NOEs reduces the

Citation: Markwick PRL, Malliavin T, Nilges M (2008) Structural Biology by NMR: Structure, Dynamics, and Interactions. PLoS Comput Biol 4(9): e1000168. doi:10.1371/journal.pcbi.1000168

Editor: Johanna McEntyre, National Center for Biotechnology Information (NCBI), United States of America

Published September 26, 2008

Funding: ANR-06-CIS6-012-01.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: nilges@pasteur.fr

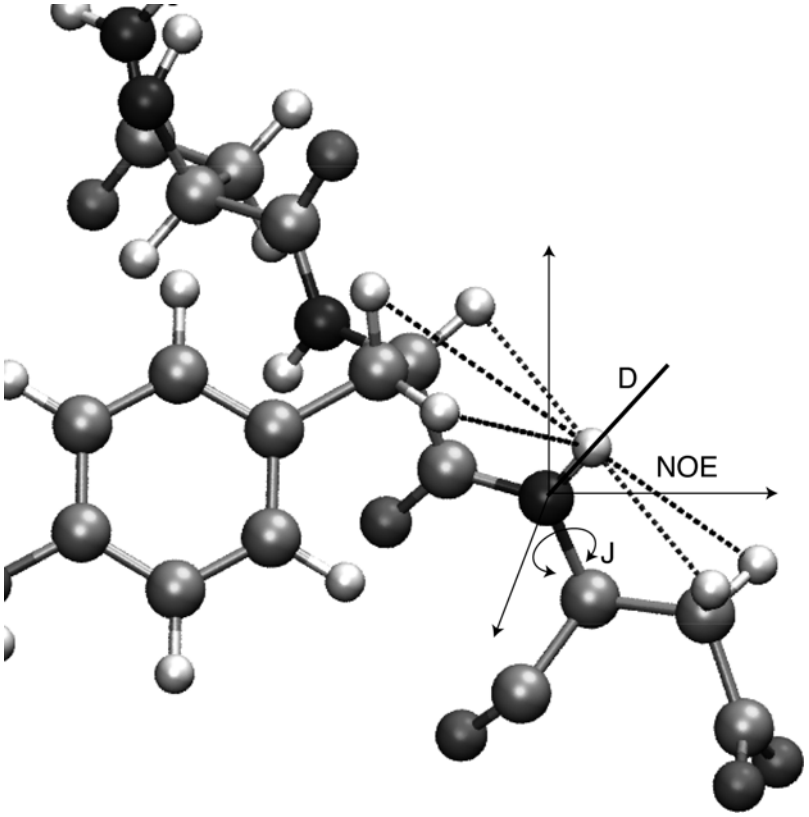

Figure 1. Illustration of structural data that can be obtained from NMR experiments, with the example of a backbone NH group. Dashed line: short inter-proton distances based on NOE measurements; arrow: torsion angle from scalar couplings (J); heavy line: orientation of a bond vector in a coordinate frame rigidly attached to the molecule from residual dipolar couplings (D).
doi:10.1371/journal.pcbi.1000168.g001

number of possible assignments of an NOE prior to the structure calculation [22,23] and leads to a considerable improvement in algorithm performance. Such protocols require prior knowledge of the frequencies of all nuclei (^{1}H, ^{13}C, ^{15}N). Automated chemical shift assignment has become more reliable [24], and with good data may allow fully automatic structure calculation [25].

An attractive idea is to dispense completely with the chemical shift assignment and to calculate "clouds" of (covalently unconnected) protons directly from the NOE data. Ambiguities in the NOE data make this a difficult task, even for the most successful implementations [26]. Results with a partial implementation of this idea limited to missing chemical shift assignments in automated structure calculations [27] are encouraging. RDCs allow one to obtain the fold of a protein [28–30], and in combination with ab initio fold prediction software and subsequent filtering with NMR data, allow for rapid automated assignment and fast fold determination [31].

When implemented in combination with other data, such as intermolecular NOEs, RDCs are also extremely useful in the determination of molecular complexes [32] or multi-domain proteins. Similarly, the combination of NMR and small angle X-ray scattering (SAXS) methods can be applied [33]. Chemical shift variations may also be employed in the form of additional intermolecular distance restraints [34]. Using appropriate labelling techniques, the application of NMR methods can be extended to the characterisation of molecular interactions in very large molecular assemblies [35,36], and, importantly, NMR allows the study of transient interactions and of low-affinity complexes [37,38].

Assessment of the quality of data and structures is of utmost importance, particularly when automation protocols are employed [39]. Validation tools can assess the consistency of the data directly by information theoretical methods [40]. The Bayesian structure calculation method discussed below inherently validates both structures and data [41,42]. New NMR quality assessments based on statistical methods [43] provide a global measure of the agreement of the calculated structures with the NOESY peak lists. The importance of using state-of-the-art structure calculation protocols and adapted force fields [44] has been demonstrated by a systematic re-calculation of a large number of structures [45].

Towards objective NMR structures. In order to calculate a 3D structure from data, it is necessary to use a model or a theory to calculate the data from the atomic coordinates. Incompleteness of the data, experimental errors, and approximations in the theory make it, strictly speaking, impossible to calculate the structure directly from the experimental observables. A further problem of the direct approach is that most theories contain unmeasurable parameters that need to be estimated before the calculation.

The standard approach to structure determination is to set up a so-called hybrid energy $E_{hybrid} = E_{phys} + w_{data} E_{data}$ [46,47], where the non-physical term E_{data} assesses the consistency between the experimental data and the coordinates, and the physical energy E_{phys} complements the experimental values with information known a priori, such as covalent bond lengths [47]. Conceptual problems with this approach are circumvented by introducing ad hoc assumptions and modifications, for example, by introducing bounds on NOE-based distances [6] rather than using the specific experimental values.

Repeated minimisation with the same empirical data results in structural divergence, which reflects the incompleteness of the data. The divergence is used as an ad hoc measure of the uncertainty of the solution, its "precision" [48,49]. Incorrectly, it is also sometimes regarded as a measure of the extent of molecular motion. Coordinates and their precision are influenced by many factors, such as the choice of auxiliary parameters and the limitations in the minimisation algorithm. In consequence, structure determination by NMR is often perceived as less objective than X-ray crystallography. A recent analysis of errors in published NMR structures [50] highlights the danger of subjective elements in structure determination procedures.

In contrast to the standard approach, the Inferential Structure Determination (ISD) method considers structure determination as an "inference problem," which is appropriate since the empirical data are incomplete and uncertain [41]. It differs from the standard approach both in the way it uses the data and in the way it generates and evaluates the resulting structures. Data enters the calculation as close to raw data as is feasible; the theory to calculate data from the structure and an error model is explicitly formulated within the formalism. The error model accounts for deviations between calculated and measured observables. A force field E_{phys} provides the prior knowledge concerning biomolecular structural parameters. These two terms resemble the two terms in the hybrid energy function introduced above. In contrast, however, probability calculus is used to assign and rank a "posterior probability" to every conformation. All unknown parameters are estimated during the calculation, making ISD a method without free parameters. This includes the unknown error of the empirical measurements, which is related to the weight of the experimental data [51].

Inferential structure determination consists of exploring the conformational space to obtain a probability distribution for the structural ensemble. Since systematic exploration of all possible structures is unfeasible, ISD uses an appropriate sampling strategy (replica-exchange Monte Carlo [52]) that visits regions of conformational space with a frequency proportional to the posterior probability. Compared to conventional structure calculation techniques, a Bayesian approach is computationally more challenging, because distributions of structures need to be explored.

The importance of the approach does not lie primarily in improved convergence for sparse data but in the fact that it puts calculation from experimental data on a sound theoretical basis: no ad hoc rules, nor empirical estimations of parameters, are necessary. It is in principle applicable to the interpretation of all forms of experimental data, and also for other applications such as comparative modelling.

Probing Structural Dynamics by NMR

The most severe approximation to structure determination is the general assumption that the experimental data can be represented by a single structure, neglecting the effects of internal dynamics. The ISD approach deals with statistical uncertainties in a rigourous manner, maintaining, however, this single copy model. Any ensemble generated by ISD or repeated minimisation cannot represent true dynamics, but only lack of information. It is therefore not meaningful to try to optimise the precision of the ensemble to some expected (or measured) dynamical property.

Attempts to go beyond the single structure approximation by ensemble or trajectory averaging have a long history [53–55]. In these studies, the molecular dynamics (MD) force field is not only used to complement the absence of structural information, but also the motion observed in the MD trajectory is employed as a model to explain dynamical features. Adding a pseudo-potential for the experimental data in an MD simulation perturbs the dynamics in a non-predictable manner, making a detailed analysis of the resulting trajectories difficult. For example, the ensemble of structures generated in such a way cannot be expected to have the correct relative free energy weighting for each ensemble member.

NMR is an ideal tool for probing dynamics occurring across a broad hierarchy of time-scales. Difficulties occur in the specific interpretation and quantification of the dynamic processes being observed. Raw experimental NMR data can provide detailed information concerning dynamically active regions in the system occuring on a particular time-scale (see Figure 2). However, this information is encoded in a complex manner and does not directly provide specific information about the molecular motions. To this end, experimental NMR data is complemented by the use of geometric models and increasingly by MD simulation to characterize at an atomistic level local dynamic processes and complex structural transitions.

Fast time-scale local motions. Precise information about local dynamics on pico- to nano-second time-scales can be obtained by spin relaxation measurements. Spin relaxation has been used to study fast time-scale dynamics in proteins for several decades, but the study of fast motions remains a rapidly developing and exciting field employing an increasing variety of experimental and theoretical methods. The importance of fast time-scale dynamics is often under-estimated: fast time-scale motions act to stabilize the protein in its folded state, and their presence is a necessary pre-requisite for slower time-scale dynamics involving large-scale collective motions [56].

The traditional "model-free" [57] approach describes local internal motions using two parameters (an order parameter characterising spatial restriction of the motion, and a relaxation time) without making reference to a specific motional model. On the other hand, numerous analytical models have been developed to describe fast time-scale local dynamic fluctuations [58–60]. One of the most popular anisotropic models is the 3D-Gaussian Axial Fluctuation (GAF) model [61] based on the observation of peptide plane motions extracted from an MD trajectory. Alternative approaches to interpreting spin relaxation make use of the strong relationship between structure and local dynamics [62,63] to rapidly predict ^{15}N order parameters from a known structure.

In order to provide a more complete description of fast time-scale dynamics, numerous experiments have been developed to obtain order parameters complementary to the N-H vector in both the backbone and side-chains [64]. Cross-correlated relaxation (CCR), which arises from the interference of two relaxation mechanisms such as the chemical shift anisotropy (CSA) and dipole–dipole interaction, has emerged as a powerful tool to study local anisotropic dynamics. For example, by combining a CSA model based on density functional theory (DFT) calculations [65] with MD simulation, it was possible to reproduce to a high degree of accuracy a complete set of CO CSA/DD CCR rates [66]. In the framework of the 3D-GAF model, it was shown that local anisotropic motions can be accurately estimated from a single CO/N-H CCR rate.

The study of local dynamic fluctuations using MD simulation is now routine. Experimentally determined order parameters are regularly used to gauge the accuracy of MD simulations [67]. Continued research in the area of force-field development [68] has resulted in a marked improvement in the prediction of order parameters [69]. The inclusion of polarization and quantum effects of the atomic nuclei in the next generation of force-fields will no doubt lead to further improvement. However, discrepan-

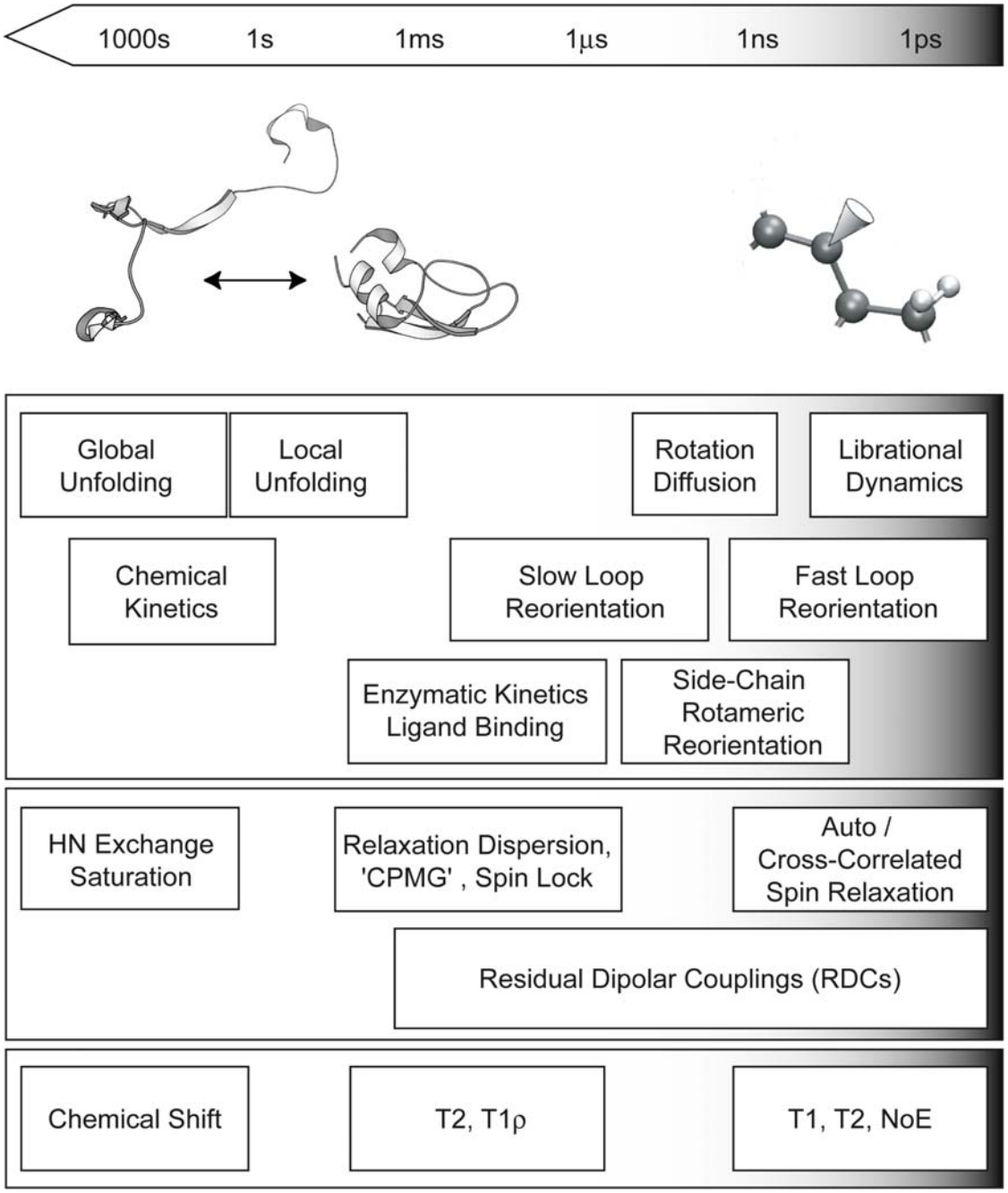

Figure 2. Time-scales, molecular motions (first panel), NMR experiments (second panel), and NMR parameters (bottom panel). Molecular motions at a particular time-scale can be probed by analysing the NMR observables in the bottom panel by using the appropriate NMR experiments (see text). Indicated are the extreme cases, folding/unfolding (which can be observed with hydrogen exchange saturation via the chemical shift) and fast local dynamics (observed with spin relaxation experiments through measurements of the relaxation times T_1, T_2, and the NOE between the amide nitrogen and the amide proton). The cone indicates, in a qualitative way, the spatial extension of the motion described by the N-H group.
doi:10.1371/journal.pcbi.1000168.g002

cies between experimental and simulated order parameters may not be solely due to inadequacies in the force-fields, but to incomplete conformational sampling [70].

Despite the maturity of the field, many issues remain unresolved. For example, different models of molecular motion are equally capable of reproducing the experimental results [71]. Many long-held assumptions concerning the local molecular geometry of the peptide plane, and in particular the position of the amino-proton are being revisited. Also, the generally accepted idea that fast internal motion and overall molecular tumbling can be treated independently has been brought into question [72].

Going beyond the nano-second limit. Many biologically important processes, such as enzyme catalysis, signal transduction, ligand binding, and allosteric regulation occur on the micro- to milli-second time-scale. Despite their obvious importance, the study of these slow motions remains a challenge to both

experimentalists and theoreticians alike. The study of dynamics at these longer time-scales are centred mostly on relaxation dispersion and RDC measurements.

In the presence of a suitable alignment medium, RDCs are averages over all orientations of the magnetic dipolar interaction tensor up to a time-scale defined by the inverse of the alignment-induced coupling. This makes them sensitive to dynamic processes up to several milli-seconds. That RDCs can probe dynamics on extended time-scales was recognised early [8]. However, even today there are conflicting views concerning the sensitivity of RDCs to slow time-scale motions and the ability to separate the contributions to RDCs arising from structural and dynamic properties of the system. Thus, several studies on model protein systems [73,74] have concluded that a single copy representation of the molecule is in general sufficient to explain the data, and that only a small subset of residues exhibit large amplitude fluctuations on slower time-scales. In contrast to this, simultaneous structure–dynamics determination approaches have suggested the presence of significant slow time-scale molecular motions. Independent studies performed on ubiquitin using model-free approaches [75–77] showed an effective homogeneous distribution of long time-scale dynamics across the molecule. A 3D-GAF–based RDC analysis of the protein GB3 suggested a heterogeneous distribution of highly anisotropic long time-scale dynamics [78,79]. In part, the discrepancy between different analyses can be ascribed to the very small number of systems studied in detail to date, and no general trends can be expected as yet. However, considering the fact that the two proteins studied in most detail (GB3 and ubiquitin) show a similar fold, it is surprising that the observed distribution of slow motions appear to be so different.

RDCs provide a detailed quantitative view of the time- and ensemble-averaged protein structure and the amplitude and direction of slow time-scale motions; recently, RDCs have also been obtained for excited protein states [80]. Considering the wealth of structural and dynamic information encoded in experimental RDC data as discussed previously in this Review, the observation of RDCs in these low-populated states may well provide a completely new direction for the study of long time-scale dynamics in proteins.

The characterisation of motion by relaxation dispersion involves measuring the excess transverse relaxation rate caused by the exchange of nuclei between different conformational states or sites with different characteristic chemical shifts. Recent methodological advances in experimental techniques have extended both the time-scale of observable dynamic processes [81] and the sensitivity [82] of the experiments to exchange processes involving enzyme catalysis [83–85], regulation [86], and ligand binding [87,88]. Lewis Kay and co-workers have developed an entire suite of relaxation dispersion experiments allowing the study of "invisible," low-populated excited states in proteins probing exchange processes in both the backbone and side-chains [87,89–91].

Relaxation dispersion experiments provide information concerning the location of dynamically active sites in a molecule and the exchange rates between the different conformational states as well as their relative free energies (and thus their populations). Unfortunately, relaxation dispersion experiments do not provide any direct structural information about the different conformational states, and a structural model of the dynamic processes observed is difficult to extract. This makes it necessary to combine the information with other experiments [92] or MD simulations.

Despite the continual increase in both available computational power and efficiency of contemporary algorithms, the simulation of slow motions in proteins involving stochastic transitions over large energy barriers on the rugged and highly structured potential energy surface remains a challenging and active field of research. Considerable progress has been made in the development of new methods to sample the conformational space of proteins more efficiently such as conformational flooding [93], accelerated MD [94], and many others reviewed recently [95]. "Biased potential" MD simulations have successfully identifed large-scale slow collective motions in proteins [96,97]. A 0.2 μs "brute-force" MD simulation of ubiquitin showed considerable dynamics occurring on time-scales beyond those probed by spin-relaxation measurements [98], and, very recently, accelerated MD simulations of the GB3 domain reliably reproduced RDC-based order parameters [70]. In light of these early successes, the study of long time-scale dynamics using a combination of MD simulation and experimental NMR holds great promise for the future.

Conclusions and Outlook

The fundamental challenge to NMR remains to combine and reconcile all the available information, both structural and dynamic, into a complete, and therefore intrinsically more accurate, representation of the conformational space sampled by biomolecular systems, with the aim of resolving the relationship between structure, dynamics, and function.

One of the most interesting aspects of NMR is that it is not limited to the study of highly structured systems, and an exciting new application of NMR-based experiments has emerged in the field of natively unstructured proteins. Fully or partially natively unstructured proteins make up a substantial part of protein sequences coded in eukaryotic genomes [99], and they play a key role in some of the most important biological processes and degenerative pathology. It is possible to measure small but finite RDCs from natively unstructured or unfolded proteins [100]. The interpretation of these RDCs is rather complex, since a single structure is certainly no longer appropriate in this case; rather a large ensemble of interchanging structures is required to fully describe the conformational behavior of the system, generated by random sampling [101,102] or by accelerated MD [103].

The close connection between experimental NMR, molecular modeling, and MD simulation has a long history. Molecular modeling approaches and simulations are necessary to interpret the data, whilst NMR experiments serve to act as a guide for the improvement of force fields. The study of long time-scale dynamics and unstructured proteins provides new and exciting challenges to both theoreticians and experimentalists.

References

1. Wishart DS (2005) NMR spectroscopy and protein structure determination: Applications to drug discovery and development. Curr Pharm Biotechnol 6: 105–120.
2. Liu HL, Hsu JP (2005) Recent developments in structural proteomics for protein structure determination. Proteomics 5: 2056–2068.
3. Altieri AS, Byrd RA (2004) Automation of NMR structure determination of proteins. Curr Opin Struct Biol 14: 547–553.
4. Tugarinov V, Hwang PM, Kay LE (2004) Nuclear magnetic resonance spectroscopy of highmolecular-weight proteins. Annu Rev Biochem 73: 107–146.
5. Bouvignies G, Markwick PRL, Blackledge MJ (2007) Simultaneous definition of high resolution protein structure and backbone conformational dynamics using NMR residual dipolar couplings. Chemphyschem 8: 1901–1909.
6. Wüthrich K (1986) NMR of proteins and nucleic acids. New York: John Wiley.
7. Iwahara J, Clore GM (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR. Nature 440: 1227–1230.
8. Tolman JR, Flanagan JM, Kennedy MA, Prestegard JH (1997) NMR evidence for slow collective motions in cyanometmyoglobin. Nature Struct Biol 4: 292–297.

9. Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278: 1111–1114.
10. Wishart DS, Case DA (2001) Use of chemical shifts in macromolecular structure determination. Methods Enzymol 338: 3–34.
11. Cavalli A, Salvatella X, Dobson CM, Vendruscolo M (2007) Protein structure determination from NMR chemical shifts. Proc Natl Acad Sci U S A 104: 9615–9620.
12. Shen Y, Lange O, Delaglio F, Rossi P, Aramini JM, et al. (2008) Consistent blind protein structure generation from NMR chemical shift data. Proc Natl Acad Sci U S A 105: 4685–4690.
13. Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, et al. (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. Nature 420: 98–102.
14. Loquet A, Bardiaux B, Gardiennet C, Blancher C, Baldus M, et al. (2008) 3D structure determination of the crh protein from highly ambiguous solid-state NMR restraints. JACS. pp 3579–3589.
15. Yee A, Gutmanas A, Arrowsmith CH (2006) Solution NMR in structural genomics. Curr Opin Struct Biol 2006: 611–617.
16. Güntert P (2004) Automated NMR structure calculation with CYANA. Methods Mol Biol 278: 353–378.
17. Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Ono AM, et al. (2006) Optimal isotope labelling for NMR protein structure determinations. Nature 440: 52–57.
18. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, et al. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59: 687–696.
19. Huang YJ, Moseley HN, Baran MC, Arrowsmith C, Powers R, et al. (2005) An integrated platform for automated analysis of protein NMR structures. Methods Enzymol 394: 111–141.
20. Kuszewski J, Schwieters CD, Garrett DS, Byrd RA, Tjandra N, et al. (2004) Completely automated, highly error-tolerant macromolecular structure determination from multidimensional nuclear Overhauser enhancement spectra and chemical shift assignments. J Am Chem Soc 126: 6258–6273.
21. Rieping W, Habeck M, Bardiaux B, Bernard A, Malliavin TE, et al. (2007) ARIA2: Automated NOE assignment and data integration in NMR structure calculation. Bioinformatics 23: 381–382.
22. Herrmann T, Güntert P, Wüthrich K (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol 319: 209–227.
23. Huang YJ, Tejero R, Powers R, Montelione GT (2006) A topology-constrained distance network algorithm for protein structure determination from NOESY data. Proteins 62: 587–603.
24. Jung YS, Zweckstetter M (2004) Mars—Robust automatic backbone assignment of proteins. J Biomol NMR 30: 11–23.
25. Lopez-Mendez B, Güntert P (2006) Automated protein structure determination from NMR spectra. J Am Chem Soc 128: 13112–13122.
26. Grishaev A, Llinas M (2005) Protein structure elucidation from minimal NMR data: The CLOUDS approach. Methods Enzymol 394: 261–295.
27. Ab E, Pugh DJ, Kaptein R, Boelens R, Bonvin AM (2006) Direct use of unassigned resonances in NMR structure calculations with proxy residues. J Am Chem Soc 128: 7566–7571.
28. Hus JC, Marion D, Blackledge M (2001) Determination of protein backbone structure using only residual dipolar couplings. J Am Chem Soc 123: 1541–1542.
29. Kontaxis G, Delaglio F, Bax A (2005) Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining. Methods Enzymol 394: 42–78.
30. Wang L, Mettu RR, Donald BR (2006) A polynomial-time algorithm for de novo protein backbone structure determination from nuclear magnetic resonance data. J Comput Biol 13: 1267–1288.
31. Meiler J, Baker D (2003) Rapid protein fold determination using unassigned NMR data. Proc Natl Acad Sci U S A 100: 15404–15409.
32. Clore GM, Schwieters CD (2003) Docking of protein-protein complexes on the basis of highly ambiguous intermolecular distance restraints derived from $^{1}H/^{15}N$ chemical shift mapping and backbone 15n-1h residual dipolar couplings using conjoined rigid body/torsion angle dynamics. J Am Chem Soc 125: 2902–2912.
33. Grishaev A, Wu J, Trewhella J, Bax A (2005) Refinement of multidomain protein structures by combination of solution small-angle X-ray scattering and NMR data. J Am Chem Soc 127: 16621–16628.
34. Dominguez C, Boelens R, Bonvin A (2003) HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc 125: 1731–1737.
35. Horst R, Wider G, Fiaux J, Bertelsen EB, Horwich AL, et al. (2006) Proton–proton Overhauser NMR spectroscopy with polypeptide chains in large structures. Proc Natl Acad Sci U S A 103: 15445–15450.
36. Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature 445: 618–622.
37. Williams DC Jr, Cai M, Suh JY, Peterkofsky A, Clore GM (2005) Solution NMR structure of the 48-kda IIAMannose-HPr complex of the *Escherichia co*li mannose phosphotransferase system. J Biol Chem 280: 20775–20784.
38. Tang C, Iwahara J, Clore GM (2006) Visualization of transient encounter complexes in protein–protein association. Nature 444: 383–386.
39. Bhattacharya A, Tejero R, Montelione GT (2007) Evaluating protein structures determined by structural genomics consortia. Proteins 66: 778–795.
40. Nabuurs SB, Spronk CA, Krieger E, Maassen H, Vriend G, et al. (2003) Quantitative evaluation of experimental NMR restraints. J Am Chem Soc 125: 12026–12034.
41. Rieping W, Habeck M, Nilges M (2005) Inferential structure determination. Science 309: 303–306.
42. Nicastro G, Habeck M, Masino L, Svergun DI, Pastore A (2006) Structure validation of the Josephin domain of ataxin-3: Conclusive evidence for an open conformation. J Biomol NMR 36: 267–277.
43. Huang YJ, Powers R, Montelione GT (2005) Protein NMR recall, precision, and f-measure scores (RPF scores): Structure quality assessment measures based on information retrieval statistics. J Am Chem Soc 127: 1665–1674.
44. Linge JP, Williams MA, Spronk CA, Bonvin AM, Nilges M (2003) Refinement of protein structures in explicit solvent. Proteins 50: 496–506.
45. Nederveen AJ, Doreleijers JF, Vranken W, Miller Z, Spronk CA, et al. (2005) RECOORD: A recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. Proteins 59: 662–672.
46. Jack A, Levitt M (1978) Refinement of large structures by simultaneous minimization of energy and *R* factor. Acta Cryst A34: 931–935.
47. Brunger AT, Nilges M (1993) Computational challenges for macromolecular structure determination by X–ray crystallography and solution NMR spectroscopy. Quart Rev BioPhys 26: 49–125.
48. Spronk CA, Nabuurs SB, Bonvin AM, Krieger E, Vuister GW, et al. (2003) The precision of NMR structure ensembles revisited. J Biomol NMR 25: 225–234.
49. Snyder DA, Montelione GT (2005) Clustering algorithms for identifying core atom sets and for assessing the precision of protein structure ensembles. Proteins 59: 673–686.
50. Nabuurs SB, Spronk CA, Vuister GW, Vriend G (2006) Traditional biomolecular structure determination by NMR spectroscopy allows for major errors. PLoS Comput Biol 2: e9. doi:10.1371/journal.pcbi.0020009.
51. Habeck M, Rieping W, Nilges M (2006) Weighting of experimental evidence in macromolecular structure determination. Proc Natl Acad Sci U S A 103: 1756–1761.
52. Habeck M, Nilges M, Rieping W (2005) Replica-exchange Monte Carlo scheme for Bayesian data analysis. Phys Rev Lett 94: 018105.
53. Kim Y, Prestegard JH (1989) A dynamic model for the structure of acyl carrier protein in solution. Biochemistry 28: 8792–8797.
54. Torda AE, Scheek RM, van Gunsteren WF (1990) Time-averaged nuclear Overhauser effect distance restraints applied to tendamistat. J Mol Biol 214: 223–235.
55. Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M (2005) Simultaneous determination of protein structure and dynamics. Nature 433: 128–132.
56. Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, et al. (2007) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. Nature 450: 913–916.
57. 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.
58. Woessner DE (1965) Nuclear magnetic dipole–dipole relaxation in molecules with internal motion. J Chem Phys 42: 1855–1859.
59. Daragan VA, Mayo KH (1997) Motional model analyses of protein and peptide dynamics using 13C and 15N NMR relaxation. Prog NMR Spectrosc 31: 63–105.
60. Chang SL, Tiandra N (2001) Molecular dynamics and NMR spin relaxation in proteins. J Am Chem Soc 123: 11484–11485.
61. Bremi T, Bruschweiler R (1997) Locally anisotropic internal polypeptide backbone dynamics by NMR relaxation. J Am Chem Soc 119: 6672–6673.
62. Zhang F, Brüschweiler R (2002) Contact model for the prediction of NMR N-H order parameters in globular proteins. J Am Chem Soc 124: 12654–12655.
63. Abergel D, Bodenhausen G (2005) Predicting internal protein dynamics from structures using coupled networks of hindered rotators. J Chem Phys 123: 204901.
64. Bruschweiler R (2003) New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins. Curr Opion Struct Biol 13: 175–183.
65. Markwick PRL, Sattler M (2004) Site-specific variations of cabonyl chemical shift anisotropies in proteins. J Am Chem Soc 126: 11424–11425.
66. Markwick PRL, Sprangers R, Sattler M (2005) Local structure and anisotropic backbone dynamics from cross-correlated NMR relaxation in proteins. Ang Chem Int Ed 44: 3232–3237.
67. Case DA (2002) Molecular dynamics and NMR spin relaxation in proteins. Acc Chem Res 35: 325–331.
68. Mackerell AD Jr (2004) Empirical force fields for biological macromolecules: Overview and issues. J Comput Chem 25: 1584–1604.
69. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, et al. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins: Struct Funct Bioinf 65: 712–725.

70. Markwick PR, Bouvignies G, Blackledge M (2007) Exploring multiple timescale motions in protein GB3 using accelerated molecular dynamics and NMR spectroscopy. J Am Chem Soc 129: 4724–4730.
71. Loth K, Pelupessy P, Bodenhausen G (2005) Chemical shift anisotropy tensors of carbonyl, nitrogen and amide proton nuclei in proteins through cross-correlated relaxation in NMR spectroscopy. J Am Chem Soc 127: 6062–6068.
72. Tugarinov V, Liang Z, Shapiro YE, Freed JH, Meirovitch E (2001) A structural mode-coupling approach to 15N NMR relaxation in proteins. J Am Chem Soc 123: 3055–3063.
73. Ulmer TA, Ramirez BE, Delaglio F, Bax A (2003) Evaluation of backbone proton postions and dynamics in a small protein by liquid crystal NMR spectroscopy. J Am Chem Soc 125: 9179–9191.
74. 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.
75. Peti W, Meiler J, Bruschweiler R, Griesinger C (2002) Model-free analysis of protein backbone motion from residual dipolar couplings. J Am Chem Soc 124: 5822–5833.
76. Lakomek NA, Carlomagno T, Becker S, Griesinger C (2006) A thorough dynamic interpretation of residual dipolar couplings in ubiquitin. J Biomol NMR 34: 101–115.
77. Briggman KB, Tolman JR (2003) De novo determination of bond orientations and order parameters from residual dipolar couplings with high accuracy. J Am Chem Soc 125: 10164–10165.
78. Bouvignies G, Bernado P, Meier S, Cho K, Grzesiek S, et al. (2005) Identification of slow correlated motions in proteins using residual dipolar couplings and hydrogen-bond scaler couplings. Proc Nat Acad Sci 102: 13885–13890.
79. Bouvignies G, Markwick PRL, Bruschweiler R, Blackledge M (2006) Simultaneous determination of protein backbone structure and dynamics from residual dipolar couplings. J Am Chem Soc 128: 15100–15101.
80. Vallurupalli P, Hansen DF, Stollar E, Meirovitch E, Kay LE (2007) Measurement of bond vector orientations in invisible excited states of proteins. Proc Natl Acad Sci U S A 104: 18473–18477.
81. Loria JP, Rance M, Palmer AG (1999) A relaxation-compensated carr–purcell–meilboomgill sequence for characterizing chemical exchange by NMR spectroscopy. J Am Chem Soc 121: 2331–2332.
82. Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Nat Acad Sci U S A 94: 12366–12371.
83. Osborne MJ, Schnell J, Benkovic SJ, Dyson HJ, Wright PE (2001) Backbone dynamics in dihydrofolate reductase complexes: Role of loop flexibility in the catalytic mechanism. Biochemistry 40: 9846–9859.
84. Cole R, Loria JP (2002) Evidence for flexibility in the function of ribonuclease A. Biochemistry 41: 6072–6081.
85. Eisenmesser EZ, Bosco DA, Akke M, Kern D (2002) Enzyme dynamics during catalysis. Science 295: 1520–1523.
86. Feher VA, Cavanagh J (1999) Millisecond-timescale motions contribute to the function of the bacterial response regulator protein SpoOF. Nature 400: 289–293.
87. Mulder FAA, Mittermaier AA, Hon B, Dahlquist FW, Kay LE (2001) Studying excited states of proteins by NMR spectroscopy. Nat Struct Biol 8: 932–935.
88. Volkman BF, Lipson D, Wemmer D, Kern D (2001) Two-state allosteric behaviour in a single-domain signaling protein. Science 291: 2429–2433.
89. Korzhnev DM, Salvatella X, Vendruscolo M, Nardo AAD, Davidson AR, et al. (2004) Lowpopulated folding intermediates of fyn sh3 characterized by relaxation dispersion NMR. Nature 430: 586–590.
90. Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, et al. (2005) Intrinsic dynamics of an enzyme underlies catalysis. Nature 438: 117–121.
91. Vallurupalli P, Hansen DF, Kay LE (2008) Probing structure in invisible protein states with anisotropic NMR chemical shifts. J Am Chem Soc 130: 2734–2735.
92. Wang L, Pang Y, Holder T, Brender JR, Kurochkin AV, et al. (2001) Functional dynamics in the active site of the ribonuclease binase. Proc Nat Acad Sci U S A 98: 7684–7689.
93. Grubmüller H (1995) Predicting slow structural transistions in macromolecular systems: Conformational flooding. Phys Rev E 52: 2893–2907.
94. Hamelberg D, Mongan J, McCammon JA (2004) Fast peptidyl cis-trans isomerization within the flexible Gly-rich flaps of HIV-1 protease. J Chem Phys 120: 11919–11929.
95. Elber R (2005) Long-timescale simulation methods. Curr Opin Struct Biol 15: 151–156.
96. Schulze BG, Grubmüller H, Evanseck JD (2000) Functional significance of hierachical tiers in carbonmonoxy myoglobin: Conformational substates and transitions studied by conformational flooding simulations. J Am Chem Soc 122: 8700–8711.
97. Hamelberg D, McCammon JA (2005) Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. J Am Chem Soc 127: 13778–13779.
98. Nederveen AJ, Bonvin AMJJ (2005) NMR relaxation and internal dynamics of ubiquitin from a 0.2 μs MD simulation. J Chem Theory and Comput 1: 363–374.
99. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6: 197–208.
100. Shortle D, Ackerman MS (2001) Persistence of native-like topology in a denatured protein in 8 m urea. Science 293: 487–489.
101. Bernardo P, Blanchard L, Timmins P, Marion D, Ruigrok WH, et al. (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. Proc Nat Acad Sci U S A 102: 17002–17007.
102. Jha AK, Colubri A, Freed K, Sosnick T (2005) Statistical coil model of the unfolded state: Resolving the reconciliation problem. Proc Nat Acad Sci 102: 13099–13104.
103. Mukrasch MD, Markwick PRL, Biernat J, von Bergen M, Bernado P, et al. (2007) Highly populated turn conformations in natively unfolded tau protein identified from residual dipolar couplings and molecular simulation. J Am Chem Soc 129: 5235–5243.