

Quantitative Analysis of Biomolecular NMR Spectra: A Prerequisite for the Determination of the Structure and Dynamics of Biomolecules

Thérèse E. Malliavin*

Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie PhysicoChimique, 13 rue P. et M. Curie, 75 005 Paris, France.

Abstract: Nuclear Magnetic Resonance (NMR) became during the two last decades an important method for biomolecular structure determination. NMR permits to study biomolecules in solution and gives access to the molecular flexibility at atomic level on a complete structure: in that respect, it is occupying a unique place in structural biology. During the first years of its development, NMR was trying to meet the requirements previously defined in X-ray crystallography. But, NMR then started to determine its own criteria for the definition of a structure. Indeed, the atomic coordinates of an NMR structure are calculated using restraints on geometrical parameters (angles and distances) of the structure, which are only indirectly related to atom positions: in that respect, NMR and X-ray crystallography are very different. The indirect relation between the NMR measurements and the molecular structure and dynamics makes critical the precision and the interpretation of the NMR parameters and the development of quantitative analysis methods. The methods published since 1997 for liquid-NMR of proteins are reviewed here. First, methods for structure determination are presented, as well as methods for spectral assignment and for structure quality assessment. Second, the quantitative analysis of structure mobility is reviewed.

1. INTRODUCTION

Nuclear Magnetic Resonance (NMR) became at the beginning of 90's, a method of choice for structural biology in solution. NMR was facing a previously developed and well-established technique: the X-ray crystallography, and the use of NMR had to be developed in structural biology. To favor this development, NMR should keep characteristics similar to X-ray crystallography, ie: (i) use as much as possible the protocols, softwares and force-field parameters developed for X-ray crystallography, (ii) obtain a convergence of the atomic coordinates similar to those obtained for crystal structures.

Since the end of the 90's, the definition of a biomolecular structure evolved. The unfolded or the flexible state of a protein was becoming a possible biologically significant state, and the conformational transition of proteins was proposed to be the basis of the neuro-degenerative diseases. The ability of NMR to observe internal mobility of molecules in solution thus became a major advantage of the method.

The structure determination by NMR is based on the measurement of geometrical parameters (distances and angles) between atoms. The influence of the NMR measurements on the structure determination is a key point for several reasons. (i) The parameters measured by NMR are qualitatively very different from the electronic density from which the crystallographic structures are determined, and the scientific knowledge accumulated in X-ray crystallography cannot thus be transferred directly to NMR. (ii) The relationship between the geometrical and the NMR

parameters is fuzzy, because each NMR parameter is depending not only on the structure, but also on the internal dynamics of the molecule, and these two aspects are difficult to separate.

The development of methods to quantitatively investigate structure and/or internal mobility from NMR measurements, is thus playing an important role for the development of NMR in structure biology, and the present review is oriented towards the presentation of these methods.

The review will focus on proteins, on liquid-NMR, and on articles published since 1997. The experiments developed for the measurement of new NMR phenomena and parameters will not be described. This review is not intended to present the processing methods used to transform the free-induction decay signal to spectral signal.

Other reviews were recently published about specific aspects of the quantitative analysis of NMR spectra. Two reviews are dealing with the validation of protein models determined by NMR [1, 2]. A larger number of reviews are presenting methods for automatic spectral assignment and structure calculation [3-8]. Reviews are focusing [9, 10] on the methods for structure calculation and refinement: long-range orientational and distance restraints, rigid-body dynamics, database potentials. The Ref. 11 is presenting methods for structure and assignment determination using dipolar couplings. Two reviews are presenting the interpretation of chemical shifts and coupling constants in macromolecules [12] and the theory of chemical shift anisotropy [13].

Several reviews are presenting the analysis of internal motions using NMR relaxation experiments [14-16], the analysis of the Brownian tumbling [17] and the prediction of NMR relaxation data from protein structures [18]. Two more recent reviews are dealing more specifically with the

*Address correspondence to this authors at the Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie PhysicoChimique, 13 rue P. et M. Curie, 75 005 Paris, France; Tel: (33) 1 58 41 51 68; Fax: (33) 1 58 41 50 26; E-mail: Therese.Malliavin@ibpc.fr

applications of the studies of protein internal dynamics: analyzing the protein disorder [19] or the molecular recognition [20].

Protein liquid-NMR is now facing two main challenges: (i) to make the structure determination easier, more precise and of better quality, (ii) to put together structure and internal mobility, in order to obtain a complete view of biomolecules. The present review is thus organized in two main parts, one devoted to the structure determination, and the other to the analysis of internal mobility.

2. STRUCTURE DETERMINATION:

The NMR structure determination requires first the assignment of NMR spectra, i.e. the assignments of each signal resonance frequency to a least one nucleus in the molecule. Such an assignment is usually solved in a progressive way, by determining first the spin systems (clustering the chemical shifts according to the residues), then by determining the sequential assignment (ordering the spin systems in the sequence), and finally by performing the assignments of the nuclear Overhauser effects (NOEs).

As an NMR molecular structure is mainly defined by interatomic distances, obtained from the NOE measurements, the first attempts at the end of the 80's to calculate atomic coordinates of a structure from distance restraints were based on the use of the Distance Geometry. This method is supplementing the NMR distance restraints by other restraints derived from the properties that should have an Euclidian object as the molecular structure in the 3D space. Unfortunately, it turned out that the problem of calculating the coordinates from the distances, is quite under-determined, because the number of NMR restraints is small with respect to the number of degrees of freedom, and also because the NMR restraints have a small upperbound, i.e. they usually correspond to distances smaller than 5 Å. It is thus difficult to define an Euclidian object from these distances, and that results in distortions of the obtained structures, which should then be removed by running molecular dynamics simulations in order to overcome energy barriers. Because of these disadvantages, the majority of the structures is now calculated by optimization methods, among them the most popular is the simulated annealing.

The NMR structure calculation is based on the optimization of atomic coordinates with respect to restraints derived from NMR measurements, in the frame of a molecular modeling force-field. The main source of restraints is usually the interatomic distances, evaluated from the intensity of magnetization transfer between the spins through the 3D space (nuclear Overhauser effect). This evaluation is imprecise, because of the existence of indirect pathways for magnetization transfer (spin diffusion), and because of the influence of internal mobility on the phenomenon. Additional restraints are provided by the J-coupling constants, which permit an estimation of the dihedral angles. The residual dipolar couplings, which can be measured on an aligned biomolecule, for example in a liquid-crystalline medium, are giving access to angle values between internuclear vectors and orientational tensor of the molecule [21]. The angles between interatomic vectors can be determined by using the effects of dipole-dipole cross-correlated relaxation [22].

2.1. Spectral Assignment

The assignment of protein NMR spectra is classically performed in two steps: the sequential assignment, which determines the chemical shifts of the backbone nuclei, and the NOE assignment, which is usually performed in parallel with the structure calculation, and which provides the distance restraints determining the fold. If these two steps are performed manually, they are demanding several months of an human expertise. Thus, following the development of structural proteomics [23], the search for fast and automatic methods to assign NMR spectra, was increasing.

The NOE assignment, performed in parallel with the structure determination, is the subject of many developments [24-31]. A seminal work in that direction is the ARIA approach [32, 33]. The general framework of the automatic NOE assignment is the following. Using the sequential assignment information, the NOE cross-peaks can be automatically assigned to several spin pairs. But, many of these assignments are false, and the automatic approaches consist to sort out the false and possible assignments by running iteratively structure calculation and removing false assignments according to their lack of consistency with the structure. The lack of consistency can be evaluated from statistics on the restraint violation (ARIA: Ref. 32, NOAH/DIAMOND: Ref. 25), by network-anchoring score (CANDID: Ref. 28) and by Bayesian (AUREMOL: Ref. 27) or other probabilistic analysis (PASD: Ref. 31). The effect of the chemical shift tolerance on the result of ARIA was investigated [34], and it was shown that ARIA protocol can deal with a large number of assignment possibilities for each peak, provided the correct option is present. Further step in the direction of automatic NOE assignment was proposed, in the case of incomplete assignment of sidechains [35], by simulating the missing chemical shifts from preliminary structures and by introducing the simulated chemical shift information into the ARIA protocol.

Automatic methods for sequential assignment are based on algorithms previously developed in computer science, as artificial intelligence (AUTOASSIGN: Ref 36) or pattern recognition [37]. Other approaches (TATAPRO: Ref. 38, MAPPER: Ref. 39) are based on statistical models and are requiring the measurements of ^{13}C and ^{15}N chemical shifts for TATAPRO, and ($^{13}\text{C}\alpha$, $^{13}\text{C}\beta$) chemical shifts for MAPPER. Another method, MARS, [40] is using only the $^{13}\text{C}\alpha/^{13}\text{C}\beta$ connectivity information, and was extensively tested with respect to missing peaks and distortions into the peak alignment. A hierarchical algorithm, HYPER [41], was proposed to perform stereospecific assignment, and two Bayesian approaches, SPI and BACUS, [42, 43] are determining (i) the spin systems from homonuclear and ^{15}N heteronuclear spectra, (ii) the probabilistic identities of NOESY cross-peaks in terms of the chemical shifts provided by SPI. Bayesian general frames for backbone resonance assignment were proposed [44-46]. Methods for automatic sequential assignment based on the use of ^{15}N connectivities observed on HNN and HN(C)N experiments were proposed [47], based on the different typical peak patterns observed in these experiments. The computational complexity of the sequential assignment problem using only $^{13}\text{C}\alpha$ chemical shift data and $\text{C}\alpha$ (i, i - 1) sequential connectivity information was explored [48].

An approach to automatic assignment is focusing on the direct determination of the protein tertiary [49, 50] or secondary [51] structure from the unassigned NMR spectra. The assignment is then a consequence of the structure determination. Up to now, the best result in that direction was obtained using the CLOUD approach [52] which is determining the protein structure as a density of protons (cloud) from the NOEs. The protein structure is then determined by threading [53] from the proton densities. An algorithm was proposed [54] for the NMR-constrained threading of a protein on a given structure. Several tools for computer-aided spectral assignment are available [55-60].

A complete set of tools, partially based on AUTOASSIGN, was developed [61] for rapid and automatic determination of medium-accuracy protein backbone structures. Simulations of 3D NOESY-HSQC were proposed to help the spectral assignment [62].

2.2. Structure Calculation

The calculation of biomolecular structure under NMR restraints is facing the problem of finding the global energy minimum of multi-dimensional conformational space, without performing an exhaustive search. The simulated annealing procedures are using high temperatures (ie. high kinetic energies) to overcome the energy barriers of the potential hypersurface, but, in Cartesian coordinates, the high-frequency vibrations of the bond and angle bond potentials are introducing dynamics instabilities at high temperatures, and are thus limiting the efficiency of the simulated annealing. An alternative to Cartesian coordinates is to perform the calculations in the torsion angle space, where the bonds and bond angles are intrinsically defined, and cannot give rise to high-frequency vibrations. During the last years, several algorithms were described for implementing the torsion angle space dynamics in CNS [63, 64], CYANA [65], XPLOR [66], XPLOR-NIH [67, 68] and ROSETTA [69]. The use of the torsion angle space was shown [63] to dramatically improve the convergence radius of a DNA duplex and to alleviate the steric hindrance observed in structures calculated in Cartesian coordinates. The torsion angle space was also used to perform a systematic search in the conformational space [70, 71], and to determine the relative orientation of covalently linked protein domains using dipolar coupling restraints [72].

Additional knowledge obtained from the analysis of previously determined structures, can be included into the algorithm of structure determination. The analysis of X-ray structures databases was used to propose potential energies to bias the biomolecular conformation during the calculation of NMR structures. These methods were applied to the definition of a torsion angle potential [73-75], and to the definition of a Ramachandran potential [76]. Similarly, the use of the protein gyration radius calculated from the protein size [77, 78] was proposed as an additional restraint in the structure calculation.

During the simulated annealing procedures, the molecule is undergoing steep conformational transitions. Moreover, the potentials between non-bonded atoms, used in the first stages of the simulated annealing are simplified. For these reasons, the structures obtained by the simulated annealing are often presenting steric hindrances, or chemical parameters

which disagree with the knowledge based on protein structure databases. Short molecular dynamics simulation in water have been shown to improve the quality of the calbindin D9k [79] and Interleukine 4 [80] structures. Such an approach was then more extensively explored [81, 82] on IL-4, crambin and ubiquitin and was proved to improve some of the parameters used to determine the structure quality. The influence of the non-bonded force field parameters on the quality of NMR structure was studied [80] on the case of Interleukine 4, and the PROLSQ non-bonded energy function was shown to achieve a higher structure quality than other non-bonded representations. A correcting factor was derived to take into account the bias induced by the spin diffusion in the distance restraints [83].

Once the spectral assignment is performed, ambiguity still remains for some distance restraints. Two largely encountered cases are concerning the stereospecific assignments and the disulfide bridges. It was shown [84] that the application of a floating restraint on the methylene groups permits to obtain structures of a quality comparable to those obtained using experimental stereospecific assignments. A similar approach was used [85] for the assignments of disulfide bridges.

A small RMSD between the conformers of an NMR structure is usually considered as a sign of a good convergence of the simulated annealing procedure. Nevertheless, it is possible [86] to maximize the RMSD of an ensemble of structures, while maintaining the accordance with the experimentally measured restraints. This results indicates that the RMSD of NMR structures is not a good estimate of the true uncertainty in the atomic coordinates.

The distributed computing was applied [87] to the determination of the NMR structure of the bio-active peptide endothelin-1: the number of generated conformers was 100 times the number usually generated in a structure determination in order to allow a better exploration of the conformational space. An implementation of the torsion potential permitting to decrease the CPU time needed for structure calculation, was proposed [88].

A general sampling algorithm was recently proposed [89] to explore the probability densities arising in Bayesian data analysis problems. This algorithm was shown to decrease significantly the backbone RMSD between the NMR conformers of the SH3 domain.

2.3. Structure Determination from a Minimal Set of Restraints

A disadvantage of the structure determination by NMR is that it relies on a heterogeneous set of redundant experimental restraints. This makes difficult to assess the precision of the obtained structure from the precision of the measured NMR parameters. The search of the minimal set of restraints permitting to obtain a reasonably accurate structure, is thus a question explored since the first years of biomolecular NMR.

The effect of the number of restraints on the quality of obtained structure was examined, during the last years, in the frame of the development of structural proteomics. The purpose of this analysis was the prediction of a 3D structure from a minimal set of restraints. This problem was approached by several ways. A genetic algorithm, adapted to

the identification of solutions in combinatorial optimization problems, was applied [90] to the NMR structure determination, which the intention to use it in the case of large proteins and few restraints. The precision of long-range HN-HN distances measured on deuterated samples was tested [91] as well as the efficiency of these restraints for rapid protein fold determination [92].

A search approach in a protein fragment database was shown [93] to be an efficient way to generate a backbone fold from residual coupling data. The variation of dipolar couplings along the protein sequence was shown [94-97] to allow the prediction of protein structural motifs and topology. The threading of a sequence through a library of candidate folds using secondary structure information from NMR, is successful [98], provided that the candidate folds contain the correct protein fold.

The TOUCHSTONEX approach [99] is performing protein structure prediction using a very limited set of NMR distance restraints: N/8 long-range restraints between sidechains, N being the number of residues. The conformational search is reduced by using a lattice model and a reduced representation of the residues. Additional restraints predicted from a threading of the protein sequence, are used. A branch and bound algorithm was used [100] for protein structure refinement along with a reduced representation of the protein, in which each residue is represented by six atoms: N, H, C α , C β , C and O.

The prediction of the protein structure directly from unassigned NMR data was also explored [101] in the frame of the ROSETTA algorithm for *ab initio* 3D structure prediction. The method was shown to produce correct folded model, provided that a least 4 % of the backbone atoms are correctly assigned in the initial models. The RMSD of the protein conformations obtained was in the 4.4-6.5 Å range. The method was then evaluated in a larger scale by applying it to a benchmark set of the Protein Data Bank [102].

2.4. J-Coupling Restraints

The number of J-coupling restraints used for a structure determination is usually one order of value smaller than the number of NOE restraints, and they are thus having a smaller influence on the structure definition. Nevertheless, as for the NOE restraints, the determination of precise dihedral angle values from J-coupling constants is not straightforward, in particular because of the intrinsic degeneracy of the Karplus relation. The Refs. 103 and 104 proposed a method to perform a self-consistent analysis of J-coupling constants to improve the determination of dihedral angles inside a protein structure: this approach permits also a new parametrization of the Karplus curve. A software, MULDER [105], was developed to extract the angle torsion information from NMR data: ³J-coupling constants and sugar pucker data. The precision of protein structure calculated using only angle restraints, was evaluated [106], and a protocol to calculate the structure using restraints on secondary structures, hydrogen bonds and distances between hydrophobic core residues, was proposed.

2.5. Dipolar Coupling Restraints

The values of the residual dipolar couplings (RDCs) are related to the orientation of chemical bonds with respect to

the alignment tensor of the molecule. In certain cases, in presence of axial symmetry or if the structure already know, the alignment tensor is known. But, in general, it has to be determined in order to use the measurements of RDCs. A protocol of structure calculation was proposed [107] where a simulated annealing refinement against the RDCs along with a grid search is used to simultaneously refine the structure and determine the axial and rhombic component of the tensor. This method requires nevertheless that a number of NOE restraints sufficient to determine the fold, are available. Another method was proposed [108] to determine the alignment tensor directly from the values of the RDCs, without knowledge of the structure. Provided that the vector bonds for which RDCs were measured, are uniformly distributed in the 3D space, the histogram of their RDCs approximates a powder pattern, from which the components of the alignment tensor can be extracted. This approach was recently continued and amplified in Ref. 109. The alignment of a well-defined domain in a protein can be also determined from a few RDCs [110] using an inversion of the order matrix. The precision of the determination of the alignment tensor was found is determined not only by the accuracy of the measured couplings, but also by the uncertainty on structure [111]. This uncertainty is leading to an underestimation of the magnitude of the alignment, if large numbers of dipolar couplings are available, and is leading to errors in alignment, if few couplings were measured. If ligand-protein complexes are studied, it was shown [112] that the symmetry properties of molecular complexes can aid in the definition of reference frame. Molecular alignment tensors of two partners in a complex can be determined prior to backbone assignment, using dipolar couplings and characteristic C α /C β chemical shifts [113].

The RDCs were also used in fast and/or automatic structure determination. It was shown [114, 115] that the fold of a protein can be calculated using RDCs and few long-range NOEs, by determining the alignment tensor on rigid fragments or on peptide plane. This determination is rapid, as it does not require the complete assignment of the NOEs. A strategy based on the RDCs was proposed [116] to simultaneously assign the spectral resonances and determine the structure. Another approach [117] is based on the measurement of sequential resonances between amide ¹H and ¹⁵N and ¹⁵C α (corresponding to the intra-residue and to the sequential connectivities), along with ¹³Ca-H α dipolar couplings. This single-step determination of protein structure is intended to provide an efficient tool for the structure determination in structural proteomics. The software DipoCoup was designed [118] for 3D-structure homology comparison based on RDCs and pseudo-contact shifts, in order to recognize protein fold motifs. The use of dipolar couplings to recognize protein structural motifs was proposed [94] in order to be included in annotation in the frame of proteomics projects. The RDCs and the chemical shifts are sufficient [119, 120] for the spectral assignment of a protein of known structure, without using any sequential NMR connectivity information. An algorithm called Nuclear Vector Replacement (NVR) [121, 122] was introduced to perform the assignment of a protein of known structure, based on RDCs and NOEs. The RCD were also used as restraints to improve the *ab initio* protein structure prediction [123, 124]. The ϕ and ψ angles can be determined

from the dipolar couplings, and the subsequent use of a program for ab initio structure prediction (ROSETTA) permits to determine the structure of ubiquitin [125].

Structure determination based only on the use of dipolar couplings attracted much interest, as the dipolar couplings values are easier to quantitatively measure than the NOE intensities. This approach became attractive, as it was shown [126, 127] that it is possible to eliminate the alignment tensor from the penalty function incorporating restraints from the RDCs. Similarly, the orientation restraints with respect to the alignment tensor can be replaced by inter-vector projection angles [128, 129], which are independent of the orientation of the alignment tensor.

The orientations of the dipolar vectors in the alignment tensor cannot be uniquely determined by a single set of RDCs, because of the uncertainty of the axis of alignment tensor [130]. A method was developed [131] to incorporate the RDCs into structure calculation in case of near axial symmetry of alignment. A two-step approach, MECCANO, was proposed [132] for the determination of protein backbone structure using only RDCs measured for two alignments of the molecule: the parameters of the alignment tensors are first determined from a least-square search algorithm, and the orientations of peptide planes are then constructed from the alignment information determined previously and refined by RDC-restrained molecular dynamics (software SCULPTOR). This approach was then applied to the structure determination of the reaction site of methionine sulfoxide reductase [133]. Similarly, from RDCs measured in two media, the direction of an internuclear angles and the backbone angles ϕ and ψ can be exactly computed [134]: this result was used to propose a systematic search algorithm for the backbone structure. An interactive tool was developed [135] for rigid-body modeling of multi-domain macromolecules using RDCs. A principal component analysis applied to RDC sets measured in at least six media with different alignment tensors is sensitive to structural heterogeneity effects induced by the media [136] and was applied to the analysis of ubiquitin [137]. The graphical analysis of the relative values of the RDCs in two media, permits [138] to determine the relative orientation of molecular alignment tensors.

Several methods were proposed [139-141] to predict the sterically induced alignment from the general shape of a molecule. A model taking into account the short-range steric and long-range electrostatic interactions was recently proposed [142] and permits to predict the orientation of a protein in the liquid crystals used for biological NMR.

An approach to calculate the alignment tensor from the orientation-dependent ^{15}N TROSY chemical shift changes [143] gives orientation angles consistent with those determined using the RDCs, and may be useful for large proteins. A strategy for structure determination using RDCs along with long-range order restraints available from paramagnetic systems [144] has permitted the *de novo* determination of a cytochrome structure.

Quantum chemical calculations were performed [145] to evaluate the vibrational averaging effects on the dipolar coupling between directly bonded nuclei. These effects can be expressed as effective bond lengths that are 0.3-5 % larger than the true bond lengths.

2.6. Chemical Shift Restraints

The chemical shift values are depending on the electronic environment of the observed nuclei. During the first years of NMR, they were seldom used as restraints in structure calculation, but, their efficiency to reduce the exploration of the conformational space, became more apparent later. Indeed, the comparison of chemical shifts inside homologous protein families, showed [146] that it is possible to predict the chemical shift values in case of large identity between the sequences, and to use this informations to help protein assignment [147]. A method for automatic prediction of chemical shifts was then proposed [148], based on local sequence alignment between the analyzed protein, and sequences from the BioMagResBank.

The protein chemical shifts were analyzed, from a statistical point of view, with respect to several structural and chemical parameters: the amino-acid types, the secondary structures, the role of the nearest-neighbor residue [149], the type of β -sheet structure [150]. From these analyses, methods were proposed to predict the amino-acid type and the secondary structure [151-156]. The chemical shift mean values were used to predict the structural class [157] or the secondary structure content [158]. A method was also proposed [159] to predict ^{15}N chemical shifts in proteins using preceding residue-specific individual shielding surfaces or torsion angles. The prediction methods are most often based on neural networks, Bayesian or training approaches. These predictions were extended to the structure determination via the use of restraints on the backbone angles ϕ and ψ [160]. A database of experimental chemical shifts was used along with homology of amino-acid triplets [161] to extract ϕ and ψ angle restraints from the measured chemical shifts of backbone atoms. From the statistics on the ^{13}C chemical shifts, the conformation of the peptidic bond Xaa-Pro could be predicted [162]. The effect of sequence in the random coil chemical shift values was also statistically analyzed [163].

The pseudo-contact shifts, depending on the magnetic susceptibility anisotropy produced by the presence of a metal ion in a protein structure, were used to include new restraints for the structure calculation through molecular dynamics [164], or to orient helices in the 3D space [165]. Structure refinement under CSA restraints was performed in XPLOR-NIH [166, 167]. Recently, the determination of macromolecular structure was proposed through the measurement of paramagnetic relaxation enhancement due to covalently attached paramagnetic groups [168].

Quantum chemistry computations were used to improve the prediction of chemical shifts from atomic coordinates, in the frame of empirical [169] models or DFT [170, 171] databases. The program SHIFTS is based on such an empirical expression of chemical shifts from the atomic coordinates. The effects of atomic positions on the ^{15}N chemical shift values were probed [172] using DFT calculations. SHIFTCALC [173] is calculating of chemical shifts from atomic coordinates, by using empirical correlations for ^1H and ^{13}C nuclei. The program SHIFTX, for predicting ^1H , ^{13}C and ^{15}N chemical shifts, is based [174] on a hybrid predictive approach that employs pre-calculated,

empirically derived chemical shift hypersurfaces in combination with classical or semi-classical equations. Less commonly predicted parameters are calculated using quantum chemistry approaches, as the CSA [175] or the dipole shielding polarizabilities [176]. The relationship between the chemical shifts values and the partial atomic charges in the molecular modeling approach was investigated [177], and linear correlations are found between the average NMR chemical shifts and the average partial charges.

The ^{13}C NMR chemical shifts can be used to predict disulfide bond formation [178]. A theoretical model was developed [179] to describe the variation of amide chemical shifts along the temperature and to relate it to a variation of conformation. The variation of chemical shifts with the temperature, and its relationship with the existence of hydrogen bonds was also studied [180, 181].

Variations of experimental chemical shifts were used to refine the BPTI [182], the lysozyme [183] and the melittin [184] structures at high pressure. The starting conformation of the protein was a conformation determined at a pressure of 1 atm, and this conformation was deformed according to the variations of chemical shifts observed with increasing pressure.

2.7. Structure Quality

The 3D NMR structures are typically represented as objects in cartesian space, calculated from geometrical restraints. Restraints inconsistency as well as the protocol of structure calculation can induce distortions in the structure. It is therefore important to develop independent methods to analyze the quality of the obtained structure and of the restraints. Methods for analyzing the structure quality are developed since the beginning of structural biology: they are based on statistical analysis of the databases of determined structure [185-187]. The basic estimators for NMR structure quality are related to the violation of NMR restraints, but other parameters were recently proposed, as the T1/T2 distribution [188]. Sidechain packing prediction was demonstrated [189] to serve as a complementary method to validate sidechain conformations in NMR structures. Relationships between the restraints and the influence they can have to each other are also analyzed, as the internal consistency of RDCs [190] and of the NMR restraints [191], the relationship between the hydrogen bonds and the RDCs [192]. The internal consistency of NMR restraints was studied from the point of view of information theory [193]. The maximum likelihood model provides a frame [194] to analyze the propagation of errors from the restraints to the atomic coordinates. The structure quality can be also described through the calculation of an R factor based on the Bayesian analysis of the spectra [195]. A Web service for the quantitative evaluation of protein structure is available [196]. NMR quality assessment scores, derived from the methods of information retrieval statistics, were recently introduced [197] to provide global measures of the goodness-of-fit of the 3D structures with NOESY peak lists.

The NMR restraints can be not equally distributed on a biomolecular structure, and some regions of the molecule are thus less defined than others. A method for visualizing the

conformational variability as atomic densities was proposed [198].

The measurements of deuterium exchange and of proton chemical shifts were shown [199] to be well correlated with the hydrogen bond distance up to a precision of 0.05 Å. The comparison of hydrogen bond lengths measured using $^3\text{hJ}_{\text{NC}\gamma}$ coupling constants with hydrogen bonds in X-ray crystallographic structures revealed [200] a increasing RMSD as the resolution of the structure decreases. An analysis of backbone hydrogen bonds, performed on nine high-resolution crystal structure, showed [201] a correlation between the hydrogen bond distance and the hydrogen bond angle. This correlation was substantiated by *ab initio* calculations, and was proposed to be a criterium for evaluating the structure quality. Such an approach was enlarged [202], by analyzing the backbone hydrogen bonds on a database of 1500 protein chains: potentials of mean force describing the hydrogen bond geometry, were derived from these data. These potentials were included in the protocol of structure calculation for 10 proteins: they improve their Ramachandran diagrams and decrease the RMSD between the coordinates of the NMR and the X-ray crystallographic structure.

2.8. Structure Determination of Complexes

The determination of multimeric structures of biomolecules by NMR is a challenge facing the NMR spectroscopists, because of the size of the studied systems, and of the ambiguities between inter- and intra-molecule restraints. NMR measures the chemical shift variation of one interaction partner in presence of the others, which gives information on the interaction surface between the two molecules. Also, the observation of RDCs can permit to define the relative orientation of the two interaction partners. The methods recently proposed to determine dimeric protein structure are based on these parameters [203, 204]. The same parameters were also proposed to be used to score the solutions proposed by *ab initio* docking programs [205].

The determination of an hexameric complex of insulin was performed [206] based on the measurement of ambiguous NOEs. In that case, the level of restraint ambiguity was higher than in the previously discussed cases, but this ambiguity can be overcome using the hexameric symmetry. The ambiguous restraints used in NMR structure calculation were also proposed to be used in protein-protein docking in a more general biochemical context [207], taking into account mutagenesis data,

Ligand-receptor interaction are classically studied by exchange-transferred NOEs, and several articles are dealing with the accuracy of structures calculated from transferred-NOE [208], with the effect of NMR relaxation on the transferred-NOEs [209] and with the docking of flexible ligands to proteins using transferred NOEs [210, 211].

Restraints issued from chemical shift mapping were used for the protein-protein [212] docking, for the spatial localization of a ligand binding site [213], or to place a ligand with respect to a receptor [214]. An automatic method for the processing of chemical shift perturbation on NMR spectra was proposed [215].

The structure of the melittin tetramer was determined at different temperatures using NOE restraints along with

chemical shift refinement to analyze the exchange between monomers and tetramers [216].

2.9. Data Model, Databases

With the development of biomolecular structure determination by NMR, several specific databases have been opened for the storage and the downloading of NMR data. The first database opened is the BioMagResbank (BMRB) [217]: it contains the chemical shifts of the nuclei in assigned proteins. This database is now also containing restraint files used for the structure calculation [218]. The BMRB data (chemical shifts and restraints) are pointed from the Protein Data Bank (PDB) pages. A database of $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ carbon-13 chemical shifts measured in proteins [169] was used to explore the effects of structure geometry to the ^{13}C shift values. RefDB, a secondary database of reference-corrected protein chemical shifts derived from the BioMagResBank (BMRB), was opened later [219]. More recently, the NMR community realizes that it might be useful to store in databases raw NMR data, for the purposes of method testing, or literature investigation. The behavior to deposit in databases raw NMR data is, up to now, not widely encountered in the NMR community. But, two projects proposing the storage of raw NMR data as a relational database are existing (SPINS: Ref. 220, NMRb: Ref. 221).

The databases of recalculated protein structures are also developed. Indeed, the NMR structures were calculated using successive sets of methods, and the quality of NMR structures now available in the PDB is heterogeneous. The availability of standard-quality structures will certainly improve their utility for molecular modeling studies. The databases DRESS (Database of REfined Solution NMR Structures) [222] and RECOORD (REcalculated COORDinates Database) [223] contain protein structures for which distance and dihedral angle restraint files were available in PDB, re-calculated using the current states of the protocols based on the softwares CYANA and CNS.

Recommendations were published by IUPAC-IUBMB-IUPAB [224-226] to describe a standard in protein NMR structure, in particular for the atom naming, the NMR spectra referencing, the determination of restraints, and the determination and validation of structures. Similarly, the development of numerous methods to deal with all aspects of biomolecular NMR has led to the creation of a collaborative project: the Collaborative Computing Project for NMR (CCPN) which proposed a general data model [227, 228] permitting to handle all NMR objects from NMR raw data up to NMR structures.

3. INTERNAL MOBILITY

The NMR measurements are closely depending on molecular internal mobility. Moreover, NMR is permitting the measurement of the mobility at the atomic level, and for a set of probes equally distributed in a molecule. In that way, NMR is providing a unique opportunity to study the internal mobility of molecules in the nanosecond-millisecond range. The NMR parameters which have been first used to analyze internal mobility, are the nuclear Overhauser effects measured between azote and amide hydrogens, and the relaxation times T1 and T2 of the amide

hydrogens. These parameters are related to the internal mobility as they are linear combination of values of the spectral density function of the motion. In a rigid tumbling molecule, the spectral density function is a Lorentzian depending only on the tumbling correlation time of the molecule. But, as biomolecules are flexible, the model of a rigid tumbling molecule is usually not fitting the data, and numerous models have been developed to introduce internal flexibility. The first proposed model is the Lipari-Szabo model [229]: it is characterized by a overall tumbling correlation time τ_c , and by an internal correlation time τ_e and an order parameter S^2 for the internal flexibility of each observed residue. Similar models have been proposed with several levels of internal mobility timescale, as the 'extended model-free approach' [230].

Another possibility, which was initiated by the Ref. 231 is to make use of the linear relationship between NMR relaxation parameters and spectral density function, and to invert this relation to obtain spectral density values. One is then faced with the interpretation of these values in terms of molecular flexibility.

The characteristics of these approaches is that the number of model parameters is of the same order of value than the number of measured NMR parameters. The system to solve is thus under-determined: the determination of mobility parameters is usually done using a multi-step protocol, and the choice at each step is critical. For example, in the case of the Lipari-Szabo model, the global correlation time τ_c is determined from a global analysis of all observed residues, and the τ_e and S^2 parameters are then determined using this τ_c value. In order to obtain reliable τ_e and S^2 values, residues undergoing exchange have to be removed from the set of residues from which τ_c is calculated.

The information on mobility extracted from NMR relaxation parameters is mainly qualitative and difficult to relate to instantaneous molecular conformations. The molecular dynamics simulations can be used to interpret the results on the molecular flexibility in terms of conformational space.

3.1. Overhauser Effect Analysis

Methods for the simulation of NOESY-HSQC spectra [232] and for the simulation of relaxation and conformational exchange [233] were developed. The nuclear Overhauser effects (NOE), produced by magnetization transfer between protons, are depending on the structure, but also on the internal dynamics of the observed molecule. The quantitative analysis of the homonuclear NOE, as well as the inversion of the intensity matrix, was extensively studied since the beginning of the 90's. The measurement of spin-diffusion-suppressed NOE data and the analysis of heteronuclear relaxation data in order to introduce a dynamical correction to ^1H - ^1H crossrelaxation rates, were used [234] to interpret various discrepancies between the NMR and crystal structures of *E. coli* thioredoxin. More recently, the interest focused on the case of partially deuterated protein [235], and on the case of the quantitative analysis of longitudinal (NOE) and transverse (ROE) relaxation [236-238]. This analysis of NOE and ROE relaxation permits, for homonuclear NMR, to separate the information on the molecular internal dynamics and on the structure.

3.2. Describing the Relaxation Measurements in Terms of Internal Motion

The NMR relaxation measurements permits the analysis of molecular internal dynamics. This analysis is classically performed by measuring NOE, T1 and T2 on a set of molecule protons: the amide protons, if ^{15}N relaxation is used, or the aliphatic protons if ^{13}C relaxation is used. Once the relaxation rates are measured, these have to be analyzed to describe the internal motions in the frame of a model. The first proposed model, the Lipari-Szabo model is based on the hypothesis that the studied molecule is globular, and the same global tumbling time is observed for all molecule regions. In the reality, this is often not the case, because of the elongated shape of the studied molecule, or because an equilibrium between monomeric and dimeric states. Furthermore, it was shown [239] that the inclusion of T2 relaxation parameters can induce an overestimation of the τ_m value. Because of these problems, numerous methods were proposed to determine an unbiased value of the tumbling correlation time. A straightforward one is selecting the subset of spins having a NOE value larger than a threshold to calculate the τ_c value [240]. The correlation time τ_c can be also determined, in the highly structured molecule regions, as the value optimizing the fit of the parameters in the Lipari-Szabo model [241]. A graphical method was proposed [242] to analyze the spectral density functions, in order to extract the τ_c value. The Ref. 243 proposed a protocol to determine iteratively the overall tumbling correlation time, even in the presence of large conformational exchange. Bayesian methods are also used for the determination of correlation times and order parameters [244].

Another research axis is complementary to the determination of the overall tumbling correlation time, and concerns the determination of the anisotropy of the molecule rotational diffusion. The correlated changes in T1 and T2 induced by anisotropic tumbling can be used [245] to distinguish between the effect of rotational diffusion anisotropy and the effect of chemical exchange. Methods have been proposed to determine this anisotropy using heteronuclear [246] or homonuclear [247] relaxation rates, in the absence of any structural information. If the structure of the observed molecule is known, the anisotropy of diffusion can be determined from the relaxation data by simulated annealing [248, 249], by optimization [250], by a Bayesian approach [251] or by a graphical method [252].

One major problem in the analysis of NMR relaxation data is the choice of the model used to fit the relaxation measurements. The propagation of the experimental errors can be assessed using a graphical method [253], or by simulations on synthetic data [254]. The influence of the spectral peak-fitting method on the value of the relaxation rates was also studied [255]. An automatic analysis method is used [256] to avoid user-dependent biases in the choice of a motion model. The model can also be chosen from information criteria [257].

As described above, the values of the spectral density functions of the motion can be calculated directly from the relaxation parameters, through the resolution of a linear system. At least three relaxation parameters (NOE, T1, T2) are measured for the relaxation of a nucleus at the frequency ω in a magnetic field. But, these relaxation parameters are

depending on values of the spectral density function at five different frequencies. To overcome this underdetermination, a solution is to use reduced spectral densities, i.e. to make the hypothesis that the values of the spectral density function at several frequencies are equal. A graphical approach was proposed [258] to perform the Lipari-Szabo analysis with reduced spectral density mapping for ^{15}N relaxation, but the reduced spectral density hypothesis is not valid in the case of ^{13}C relaxation [259]. The use of ^{13}C and ^{15}N relaxation measurements at four different magnetic fields was shown [260] to permit the calculation of all the spectral density function values, without reduction. On the other hand, in the frame of ^{15}N relaxation, the use of reduced spectral density functions, combined with the Lipari-Szabo model of motion, permits [261] the expression of the order parameter and of the correlation times in function of the relaxation parameters. The Ref. 262 introduced an order parameter depending on the resonance frequency of the observed nuclei. This parameter allows to determine the optimal number of Lorentzian to be included in the spectral density function.

In the last years, models describing the motion of biomolecules from the point of view of the hydrodynamics started to be used in the field of NMR relaxation. From these models, a software, HYDRONMR, was developed to predict NMR relaxation parameters [263]. The prediction of T1/T2 values in the frame of HYDRONMR [264] was proposed to be used to distinguish the residues undergoing chemical exchange from those undergoing diffusion anisotropy. The combined use of NMR relaxation measurements and hydrodynamic calculations was used [265] to study protein oligomerisation. Similarly, protein self-diffusion was studied in parallel with ^{15}N relaxation [266].

Several methods were proposed to predict motional or relaxation parameters from the structure of a protein. An empirical relation ship was derived [267] between the overall tumbling τ_c and the solvent accessible atom surfaces. The set of published NMR relaxation parameters was statistically analyzed [268] in order to detect empirical laws governing the order parameters. A formula to predict the order parameter of a protein residue from its distances to other heavy atoms of the protein, was empirically derived [269]. The relationship between the dynamics parameters measured on methyl groups and protein local structure was investigated [270], but no clear relation was found. On the other hand, the residues conserved by the evolution seem to be the less mobile ones [271].

The relative movements of domains inside the same biomolecule are now starting to be analyzed. A wobble-in-a-cone model was proposed [272] to describe the slow internal motions of nucleic acid fragments as well as multidomain proteins. This model was then used to analyze relaxation data measured on calmodulin [273].

The internal dynamics of protein unfolded state was studied for the bovine pancreatic trypsin inhibitor [274]. The entropy change between the folded and unfolded forms of SNase was calculated [275] in the frame of a model relating the amplitude of bond vector motions to conformational entropy. A spectral density model based on a truncated Lorentzian distribution of correlation times was used [276] to characterize the internal dynamics of protein unfolded states.

The internal motions of proteins at a semi-local level (secondary structure elements) were studied from different points of view. A motion model was proposed for describing the rotations of protein sidechain using ^{13}C relaxation [277]. The study of a α -helical peptide [278] validated a model of correlated internal motions of $\text{C}\alpha$ -H and N-H bonds. It was also proposed to measure, for the protein backbone, the relaxation of amide ^{15}N by the ^{15}N - ^1H dipolar interaction and the relaxation of carbonyl $^{13}\text{C}'$ by the $^{13}\text{C}'$ - $^{13}\text{C}\alpha$ dipolar interaction [279], to detect semi-local motion of secondary structure elements. The covariation of the backbone internal dynamics among a set of protein mutants was experimentally studied [280]. The covariation of internal motion was also studied by comparing the dynamics for different internuclear vectors in sidechains [281]. The decoupling hypothesis, i.e. neglecting the correlation between internal and overall molecular motions in the model-free approach, was investigated [282], and a new expression was derived for the spectral density function. The effect of the dynamical coupling between two interaction-carrying vectors on the internal auto- and cross-correlation functions were also investigated in the limit of small amplitude motions [283]. A huge majority of approaches are using a model-free analysis of the relaxation rates and/or of the spectral density functions. Another approach, the two-body Slowly Relaxing Local Structure (SRLS) model permits [284], to correct the underestimation of local correlation times produced by the model-free approach.

3.3. Residual Dipolar Coupling and Mobility

The residual dipolar couplings (RDCs) were first proposed to improve the quality of NMR structures or to speed-up the structure determination, but they are also depending on the internal motion of the molecule, and are motionally averaged in the picosecond to the millisecond timescale. Nevertheless, it was demonstrated [285] that, for most practical applications, the refinement of NMR structure against dipolar couplings and using a single structure representation is adequate. An approach for the simultaneous retrieval of structural and dynamical information from RDCs measured in several oriented media was proposed [286-288]: it is based on the inversion by singular value decomposition of the linear relationship between the RDCs and the dynamical and structural information. RDCs were analyzed in the frame of a model-free model [289], and motions slower than the global correlation time were then revealed in ubiquitin [290]. The detailed analysis of B3 IgG domain of streptococcal protein G [291] showed that an agreement between calculated and experimental dipolar couplings is obtained with a two-structure ensemble representation. The order parameter derived from the dipolar couplings is not exactly a measure of motion, but is convoluted by the hindered sampling of orientations [292]. The comparison of the internal mobility extracted from dipolar couplings with classical order parameters was proposed to elucidate the biomolecular slower motions [293].

Non-vanishing RDCs were measured in short peptides [294] and in unfolded proteins [295], which seems paradoxical as it was implicitly assumed that a fully denatured protein will give zero RDCs. A model of random flight chain was used [296] to calculate spatial probability distributions and deduce from them RDCs values

in agreement with the experimental measurements in unfolded state.

3.4. Molecular Dynamics and Ensemble Refinement

NMR studies are often coupled to molecular modeling and molecular dynamics simulations. The parallel use of the simulations is supported by the insights into molecular internal dynamics obtained by NMR. The articles reviewed here are not supposed to show a complete view of the molecular modeling related to NMR measurements, but to give main directions.

A model to analyze the motions in MD simulations was proposed [297, 298], based on the diagonalization of the covariance matrix describing the reorientation of the ^{15}N - ^1H or ^{13}C - ^1H bonds. NMR relaxation rates can be then calculated from the diagonalization parameters.

This method was applied to NMR relaxation measurements performed on the terminal RNA-binding domain of the human U1A protein [299], and revealed correlated motions important for the protein interaction with RNA. It was also used for investigating the internal motions of ubiquitin in the native form and in a partially folded state [300].

The molecular dynamics simulations related to NMR are often intended to compare experimental and calculated NMR parameters. The comparison of three experimental sets of ^{15}N relaxation parameters along with two theoretical sets calculated from molecular dynamics simulations showed [301] that the NMR and MD generalized parameters are of comparable accuracy for residues exhibiting motions on a fast timescale. The comparison of the conformational entropy and the NMR order parameters on native and denatured staphylococcal nuclease [302] revealed that the general behavior of the experimental order parameters between native and unfolded states was reproduced by the simulations, but that the majority of theoretical order parameters disagree with the corresponding experimental ones. Theoretical values of the ^{15}N spectral densities functions were compared [303] with the experimental ones in the case of the bRK1 PH domain. The spectral density value $J(0.87\omega_{\text{H}})$, which is more sensitive to the sub-nanosecond time scale of internal motion, shows larger discrepancies with the experimental values. A model-free approach was applied on synthetic ^{15}N relaxation data [304], calculated from a 10 nanosecond molecular dynamics trajectory of a dihydrofolate reductase ternary complex. The results obtained from the synthetic data were analogous to the experimental ones, except those concerning the nanosecond time scale motions with small amplitude.

The interest is also focusing on the relationship between the measured NMR parameters (NOE, J coupling values and chemical shifts) and the conformations sampled by the molecule. A molecular dynamics simulation of a β -heptapeptide showed [305] that NMR parameter values close to those of the folded state can be recalculated from trajectory even if the peptide was exploring a significant proportion of the conformational space and undergoing unfolding and refolding transitions during the trajectory. Similarly, multiple loop conformations of peptides were explored by MD simulations and shown [306] to be compatible with NMR data. A formalism was proposed [307] to estimate the

sensitivity of a spectroscopic parameter to the composition of the conformational ensemble, and permits to show that the mean NOE distances and 3J coupling constants are quite insensitive to the initial distribution in the conformational space. The effect of averaging during a molecular dynamics trajectory was previously estimated [308] by defining synthetic NOE restraints from an MD trajectory, and by comparing the structure obtained from these NOEs with the average trajectory structure. Many of the distance values derived from the synthetic NOE intensities, were significantly effected by the internal dynamics, and could lead to structure distortions, depending on the bound used for distance restraints. But, the major issue about this comparison was the restraint inconsistency: removing the inconsistent restraints permits to improve the accuracy and the precision of the structure. Another similar study [309] showed that for a non-negligible percentage of distance restraints, the assumption of a rigid structure is not valid.

The effect of averaging on the values of J coupling constants was studied through the analysis of MD trajectories. The precision of the calculated J-coupling constants across the hydrogen bonds was improved [310] by using conformations obtained from trajectories including several hundreds of picoseconds. The type of helical structure of a peptide was analyzed simultaneously [311] by measuring J-coupling constants across the hydrogen bonds and by recording an MD trajectory: both approaches gives similar results, and are supporting each other.

Molecular mechanics investigation of the conformational space of tetra- and pentapeptides was used [312, 313] to calculate the NMR parameters (3J vicinal coupling constant, $H\alpha$ and $^{13}C\alpha$ chemical shifts). The calculated NMR parameter values were found in good agreement with the experimental ones. The conformational distribution of the peptides shows also that the NMR parameter values are depending on the charge distribution and on the peptide sequence.

Molecular dynamics studies are also widely used to determine information on the behavior of a biomolecule, on which only partial information was measured by NMR. Typical examples concern the folding of peptides or proteins [314-316], the conformational equilibrium of peptides [317-320], the inter-molecular interactions [321], the peptide solvation [322], and the conformational transitions induced in a protein by high hydrostatic pressure [323].

Simulations were also used to compare NMR and X-ray structures of the same protein. X-ray structure at two resolutions were compared to an NMR structure and to an MD trajectory [324]: the NMR ensemble is more compatible with the high resolution X-ray structure than with the low resolution structure and the MD trajectory. Similarly, series of MD simulations were performed [325] on 34 proteins, for which the structure was determined by NMR and/or by X-ray crystallography. The simulations show that NMR structure have generally, higher internal strain than X-ray structures, and that a significant proportion of them are unstable and diverge rapidly in simulations.

Validation of molecular modeling force-fields is performed by testing them with respect to NMR measurements. Molecular dynamics simulations on the protein hen egg white lysozyme with the GROMOS87 and

GROMOS96 force-fields were used [326] to calculate theoretical NMR parameters and compare them to the experimental ones. The GROMOS96 force-field shows better agreement with the NMR data than with the X-ray structure.

Coarse-grained models were also used to modelize NMR structure. The Gaussian network model is able [327] to predict ^{15}N relaxation data. The analysis of NMR conformers and MD trajectory in the frame of essential dynamics [328] shows a significant overlap between NMR and MD ensembles of conformers.

The time and/or ensemble averaging of restraints was tested on MD trajectories for Jcoupling and NOE restraints [329] and for orientational restraints derived from RDCs or chemical-shift anisotropies [330]. The ensemble averaging restraints for J-couplings and NOEs was using [329] an elliptic definition of the restraint, to avoid overshooting in molecular dynamics refinement. The averaging of J-coupling and NOE restraints permitted to fit the experimental results in the case of a peptide for which no single conformation could explain the measured NOE [329]. The averaging of orientational restraints [330] is essential to not reduce artificially the fluctuations of the restrained vectors.

The conformational transition in a protein, based on a conformational change of an aromatic sidechain, was studied [331] by adiabatic energy mapping, free energy perturbation and molecular dynamics simulations. These techniques permitted to calculate the structure of the minor conformation, which could not have been determined from NOE restraints. A maximum entropy method was used [332] to determine from MD trajectories an ensemble of flexible polypeptide conformation consistent with the experimental NMR measurements.

A Monte-Carlo algorithm was used to obtain an ensemble of conformers describing the unfolded state of the bovine acyl-coenzyme A binding protein [333]. Ensemble molecular dynamics simulations along with spin-label NMR was used to obtain a view [334] of the long-range interactions in the partially denatured state of α -synuclein, where the secondary structure elements are not fully formed.

Molecular dynamics simulations were run in parallel for an ensemble of copies of a protein [335] in order to determine an ensemble of conformers consistent with the NMR dynamics. A similar approach was recently [336] applied to the analysis of sidechain order parameters, to explore their internal motions. This method was then applied for the simultaneous determination of the ubiquitin structure and dynamics [337], under NOE and S^2 restraints.

4. CONCLUSION

The present review was intended to describe the methods developed since 1997 to quantitatively analyze the parameters measured for proteins in solution by NMR. It was organized in two main sections: the structure determination and the analysis of plasticity. The quantitative analysis is exploring several directions to push away the limits of NMR structure determination: (i) the analysis of complexes, (ii) the automatization and the speed-up of structure determination, (iii) the use of all experimental information available (chemical shifts, residual dipolar couplings). Due to the development of bioinformatics, new algorithms coming from computer science, are developed,

and it is possible to rely on databases of NMR and structural parameters, in order to improve the *a priori* knowledge of the studied system.

The analysis of internal mobility is characterized by a large range of NMR parameters analyzed, including, in addition to the classical relaxation parameters, the RDCs. The models used for describing the internal motion are evolving, in order to describe the unfolded states of biomolecules, to include hydrodynamics models, and to become more robust. The molecular dynamics simulations are now more and more used to test hypotheses on the NMR measurements. The ensemble refinement of NMR structures is permitting to include information about the internal mobility of the molecule in the structure determination process.

ACKNOWLEDGMENTS

CNRS is acknowledged for funding.

REFERENCES

- [1] Laskowski, R.; MacArthur, M.; Thornton, J. *Curr Opin Struct Biol.*, **1998**, *8*, 631.
- [2] Spronk, C.; Nabuurs, S.; Krieger, E.; Vriend, G.; Vuister, G. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **2004**, *45*, 315.
- [3] Nilges, M.; O'Donoghue, S. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **1998**, *32*, 107.
- [4] Moseley, H.; Montelione, G. *Curr Opin Struct Biol.*, **1999**, *9*, 635.
- [5] Guntert, P. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **2003**, *43*, 105.
- [6] Altieri, A.; Byrd, R. *Curr Opin Struct Biol.*, **2004**, *14*, 547.
- [7] Gronwald, W.; Kalbitzer, H. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **2004**, *44*, 33.
- [8] Baran, M.; Huang, Y.; Moseley, H.; Montelione, G. *Chem Rev*, **2004**, *104*, 3541.
- [9] Clore, G.; Gronenborn, A., *Proc Natl Acad Sci U S A.*, **1998**, *95*, 5891.
- [10] Clore, G.; Schwieters, C. *Curr Opin Struct Biol.*, **2002**, *12*, 146.
- [11] Jung, Y.; Sharma, M.; Zweckstetter, M. *Angew Chem Int Ed Engl.*, **2004**, *43*, 3479.
- [12] Case, D. *Curr Opin Struct Biol.*, **2000**, *10*, 197.
- [13] Sitkoff, D.; Case, D. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **1998**, *32*, 165.
- [14] Daragan, V.; Mayo, K. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **1997**, *31*, 63.
- [15] Fischer, M.; Majumdar, A.; Zuiderweg, E. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **1998**, *33*, 207.
- [16] Barbar, E. *Biopolymers*, **1999**, *51*, 191.
- [17] Korzhnev, D.; Ibraghimov, I.; Billeter, M.; Orekhov, V. *J Biomol NMR.*, **2001**, *21*, 263.
- [18] Bruschweiler, R. *Curr Opin Struct Biol.*, **2003**, *13*, 175.
- [19] Bracken, C.; Iakoucheva, L.; Romero, P.; Dunker, A. *Curr Opin Struct Biol.*, **2004**, *14*, 570.
- [20] Atkinson, R.; Kieffer, B. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **2004**, *44*, 141.
- [21] Tjandra, N.; Bax, A. *Science*, **1997**, *278*, 1111.
- [22] Reif, B.; Hennig, M.; Griesinger, C. *Science*, **1997**, *276*, 1230.
- [23] Montelione, G.; Zheng, D.; Huang, Y.; Gunsalus, K.; Szyperski, T. *Nat Struct Biol.*, **2000**, *7*, 982.
- [24] Mumenthaler, C.; Guntert, P.; Braun, W.; Wuthrich, K. *J Biomol NMR.*, **1997**, *10*, 351.
- [25] Xu, Y.; Wu, J.; Gorenstein, D.; Braun, W. *Journal of Magnetic Resonance*, **1999**, *136*, 76.
- [26] Xu, Y.; Jablonsky, M.; Jackson, P.; Braun, W.; Krishna, N. *Journal of Magnetic Resonance*, **2001**, *148*, 35.
- [27] Gronwald, W.; Moussa, S.; Elsner, R.; Jung, A.; Gansmeier, B.; Trenner, J.; Kremer, W.; Neidig, K.; Kalbitzer, H. *J Biomol NMR.*, **2002**, *23*, 271.
- [28] Herrmann, T.; Guntert, P.; Wuthrich, K. *Journal of Molecular Biology*, **2002**, *319*, 209.
- [29] Herrmann, T.; Guntert, P.; Wuthrich, K. *J Biomol NMR*, **2002**, *24*, 171.
- [30] Linge, J.; Habeck, M.; Rieping, W.; Nilges, M. *Bioinformatics*, **2003**, *19*, 315.
- [31] Kuszewski, J.; Schwieters, C.; Garrett, D.; Byrd, R.; Tjandra, N.; Clore, G. *J Am Chem Soc.*, **2004**, *126*, 6258.
- [32] Nilges, M.; Macias, M.; O'Donoghue, S.; Oschkinat, H. *J Mol Biol*, **1997**, *269*, 408.
- [33] Habeck, M.; Rieping, W.; Linge, J.; Nilges, M. *Methods Mol Biol.*, **2004**, *278*, 379.
- [34] Fossi, M.; Linge, J.; Labbude, D.; Leitner, D.; Nilges, M.; Oschkinat, H. *J Biomol NMR.*, **2005**, *31*, 21.
- [35] Hare, B.; Wagner, G. *J Biomol NMR.*, **1999**, *15*, 103.
- [36] Zimmerman, D.; Kulikowski, C.; Huang, Y.; Feng, W.; Tashiro, M.; Shimotakahara, S.; Chien, C.; Powers, R.; Montelione, G. *Journal of Molecular Biology*, **1997**, *269*, 592.
- [37] Buchler, N.; Zuiderweg, E.; Wang, H.; Goldstein, R. *Journal of Magnetic Resonance*, **1997**, *125*, 34.
- [38] Atreya, H.; Sahu, S.; Chary, K.; Govil, G. *J Biomol NMR.*, **2000**, *17*, 125.
- [39] Guntert, P.; Salzmann, M.; Braun, D.; Wuthrich, K. *J Biomol NMR.*, **2000**, *18*, 129.
- [40] Jung, Y.; Zweckstetter, M. *J Biomol NMR.*, **2004**, *30*, 11.
- [41] Tejero, R.; Monleon, D.; Celda, B.; Powers, R.; Montelione, G. *J Biomol NMR.*, **1999**, *15*, 251.
- [42] Grishaev, A.; Llinas, M. *J Biomol NMR.*, **2002**, *24*, 203.
- [43] Grishaev, A.; Llinas, M. *J Biomol NMR.*, **2004**, *28*, 1.
- [44] Bartels, C.; Guntert, P.; Billeter, M.; Wuthrich, K. *J Comp Chem*, **1997**, *18*, 139.
- [45] Lukin, J.; Gove, A.; Talukdar, S.; Ho, C. *J Biomol NMR.*, **1997**, *9*, 151.
- [46] Vitek, O.; Vitek, J.; Craig, B.; Bailey-Kellogg, C. *Statistical Applications in Genetics and Molecular Biology*, **2004**, *1*, 1.
- [47] Bhavesh, N.; Panchal, S.; Hosur, R. *Biochemistry*, **2001**, *40*, 14727.
- [48] Andreu, M.; Levy, R. *J Biomol NMR*, **2002**, *23*, 263.
- [49] Malliavin, T.; Barthe, P.; Delsuc, M. *Theo Chem Accts*, **2001**, *106*, 91.
- [50] Atkinson, R.; Saudek, V. *FEBS Lett*, **2002**, *510*, 1.
- [51] Bailey-Kellogg, C.; Widge, A.; Kelley, J.; Berardi, M.; Bushweller, J.; Donald, B. *J Comput Biol.*, **2000**, *7*, 537.
- [52] Grishaev, A.; Llinas, M. *Proc Natl Acad Sci U S A.*, **2002**, *99*, 6707.
- [53] Grishaev, A.; Llinas, M. *Proc Natl Acad Sci U S A.*, **2002**, *99*, 6713.
- [54] Xu, Y.; Xu, D.; Crawford, O.; Einstein, J. *J Comput Biol.*, **2000**, *7*, 449.
- [55] Malliavin, T.; Pons, J.; Delsuc, M. *Bioinformatics*, **1998**, *14*, 624.
- [56] Hyberts, S.; Wagner, G. *J Biomol NMR*, **2003**, *26*, 335.
- [57] Slupsky, C.; Boyko, R.; Booth, V.; Sykes, B. *J Biomol NMR*, **2003**, *27*, 313.
- [58] Moseley, H.; Sahota, G.; Montelione, G. *J Biomol NMR.*, **2004**, *28*, 341.
- [59] Goddard, T. D.; Kneller, D. G. *University of California, San Francisco*, **2004**.
- [60] Pawley, N.; Gans, J.; Michalczyk, R. *Bioinformatics*, **2005**, *21*, 680.
- [61] Zheng, D.; Huang, Y.; Moseley, H.; Xiao, R.; Aramini, J.; Swapna, G.; Montelione, G. *Protein Sci.*, **2003**, *12*, 1232.
- [62] Gorler, A.; Gronwald, W.; Neidig, K.; Kalbitzer, H. *Journal of Magnetic Resonance*, **1999**, *137*, 39.
- [63] Stein, E.; Rice, L.; Brunger, A. *Journal of Magnetic Resonance*, **1997**, *124*, 154.
- [64] Brunger, A.; Adams, P.; Clore, G.; DeLano, W.; Gross, P.; Grosse-Kunstleve, R.; Jiang, J.; Kuszewski, J.; Nilges, M.; Pannu, N.; Read, R.; Rice, L.; Simonson, T.; Warren, G. *Acta Crystallogr D Biol Crystallogr*, **1998**, *54*, 905.
- [65] Guntert, P.; Mumenthaler, C.; Wuthrich, K. *Journal of Molecular Biology*, **1997**, *273*, 283.
- [66] Badger, J.; Kumar, R.; Yip, P.; Szalma, S. *Proteins*, **1999**, *35*, 25.
- [67] Schwieters, C.; Clore, G. *Journal of Magnetic Resonance*, **2001**, *152*, 288.
- [68] Schwieters, C.; Kuszewski, J.; Tjandra, N.; Clore, G. *J Magn Reson*, **2003**, *160*, 65.
- [69] Wedemeyer, W.; Baker, D. *Proteins*, **2003**, *53*, 262.
- [70] Gippert, G.; Wright, P.; Case, D. *J Biomol NMR.*, **1998**, *11*, 241.

- [71] Guntert, P.; Billeter, M.; Ohlenschlager, O.; Brown, L.; Wuthrich, K. *J Biomol NMR.*, **1998**, *12*, 543.
- [72] Clore, G.; Bewley, C. *Journal of Magnetic Resonance*, **2002**, *154*, 329.
- [73] Kuszewski, J.; Gronenborn, A.; Clore, G. *Journal of Magnetic Resonance*, **1997**, *125*, 171.
- [74] Kuszewski, J.; Clore, G. *Journal of Magnetic Resonance*, **2000**, *146*, 249.
- [75] Clore, G.; Kuszewski, J. *J Am Chem Soc.*, **2002**, *124*, 2866.
- [76] Bertini, I.; Cavallaro, G.; Luchinat, C.; Poli, I. *J Biomol NMR.*, **2003**, *26*, 355.
- [77] Kuszewski, J.; Gronenborn, A.; Clore, G. *J Am Chem Soc.*, **1999**, *121*, 2337.
- [78] Huang, X.; Powers, R. *J Am Chem Soc.*, **2001**, *123*, 3834.
- [79] Kordel, J.; Pearlman, D.; Chazin, W. *J Biomol NMR.*, **1997**, *10*, 231.
- [80] Linge, J.; Nilges, M. *J Biomol NMR.*, **1999**, *13*, 51.
- [81] Spronk, C.; Linge, J.; Hilbers, C.; Vuister, G. *J Biomol NMR*, **2002**, *22*, 281.
- [82] Linge, J.; Williams, M.; Spronk, C.; Bonvin, A.; Nilges, M. *Proteins*, **2003**, *50*, 496.
- [83] Linge, J.; Habeck, M.; Rieping, W.; Nilges, M. *J Magn Reson.*, **2004**, *167*, 334.
- [84] Folmer, R.; Hilbers, C.; Konings, R.; Nilges, M. *J Biomol NMR.*, **1997**, *9*, 245.
- [85] Boisbouvier, J.; Blackledge, M.; Sollier, A.; Marion, D. *J Biomol NMR.*, **2000**, *16*, 197.
- [86] Spronk, C.; Nabuurs, S.; Bonvin, A.; Krieger, E.; Vuister, G.; Vriend, G. *J Biomol NMR.*, **2003**, *25*, 225.
- [87] Takashima, H.; Mimura, N.; Ohkubo, T.; Yoshida, T.; Tamaoki, H.; Kobayashi, Y. *J Am Chem Soc.*, **2004**, *126*, 4504.
- [88] Neudecker, P.; Sticht, H.; Rosch, P. *J Biomol NMR.*, **2001**, *21*, 373.
- [89] Habeck, M.; Nilges, M.; Rieping, W. *Phys Rev Lett*, **2005**, *94*, in press.
- [90] Bayley, M.; Jones, G.; Willett, P.; Williamson, M. *Protein Sci.*, **1998**, *7*, 491.
- [91] Brierecheck, D.; Rule, G. *Journal of Magnetic Resonance*, **1998**, *134*, 52.
- [92] Koharudin, L.; Bonvin, A.; Kaptein, R.; Boelens, R. *Journal of Magnetic Resonance*, **2003**, *163*, 228.
- [93] Andrec, M.; Du, P.; Levy, R. *J Biomol NMR.*, **2001**, *21*, 335.
- [94] Andrec, M.; Du, P.; Levy, R. *J Am Chem Soc.*, **2001**, *123*, 1222.
- [95] Mesleh, M.; Veglia, G.; DeSilva, T.; Marassi, F.; Opella, S. *J Am Chem Soc.*, **2002**, *124*, 4206.
- [96] Mesleh, M.; Lee, S.; Veglia, G.; Thiriot, D.; Marassi, F.; Opella, S. *J Am Chem Soc.*, **2003**, *125*, 8928.
- [97] Mascioni, A.; Veglia, G. *J Am Chem Soc.*, **2003**, *125*, 12520.
- [98] Ayers, D.; Gooley, P.; Widmer-Cooper, A.; Torda, A. *Protein Sci.*, **1999**, *8*, 1127.
- [99] Li, W.; Zhang, Y.; Kihara, D.; Huang, Y.; Zheng, D.; Montelione, G.; Kolinski, A.; Skolnick, J. *Proteins*, **2003**, *53*, 290.
- [100] Standley, D.; Eyrich, V.; Felts, A.; Friesner, R.; McDermott, A. *Journal of Molecular Biology*, **1999**, *285*, 1691.
- [101] Meiler, J.; Baker, D. *Proc Natl Acad Sci U S A.*, **2003**, *100*, 15404.
- [102] Li, W.; Zhang, Y.; Skolnick, J. *Biophys J.*, **2004**, *87*, 1241.
- [103] Schmidt, J.; Blumel, M.; Lohr, F.; Ruterjans, H. *J Biomol NMR.*, **1999**, *14*, 1.
- [104] Perez, C.; Lohr, F.; Ruterjans, H.; Schmidt, J. *J Am Chem Soc.*, **2001**, *123*, 7081.
- [105] Padrta, P.; Sklenar, V. *J Biomol NMR.*, **2002**, *24*, 339.
- [106] Alexandrescu, A. *Proteins*, **2004**, *56*, 117.
- [107] Clore, G.; Gronenborn, A.; Tjandra, N. *Journal of Magnetic Resonance*, **1998**, *131*, 159.
- [108] Clore, G.; Gronenborn, A.; Bax, A. *Journal of Magnetic Resonance*, **1998**, *133*, 216.
- [109] Bryce, D.; Bax, A. *J Biomol NMR.*, **2004**, *28*, 273.
- [110] Losonczi, J.; Andrec, M.; Fischer, M.; Prestegard, J. *Journal of Magnetic Resonance*, **1999**, *138*, 334.
- [111] Zweckstetter, M.; Bax, A. *J Biomol NMR.*, **2002**, *23*, 127.
- [112] Al-Hashimi, H.; Bolon, P.; Prestegard, J. *Journal of Magnetic Resonance*, **2000**, *142*, 153.
- [113] Zweckstetter, M. *J Biomol NMR.*, **2003**, *27*, 41.
- [114] Fowler, C.; Tian, F.; Al-Hashimi, H.; Prestegard, J. *Journal of Molecular Biology*, **2000**, *304*, 447.
- [115] Mueller, G.; Choy, W.; Yang, D.; Forman-Kay, J.; Venters, R.; Kay, L. *Journal of Molecular Biology*, **2000**, *300*, 197.
- [116] Tian, F.; Valafar, H.; Prestegard, J. *J Am Chem Soc.*, **2001**, *123*, 11791.
- [117] Zweckstetter, M.; Bax, A. *J Am Chem Soc.*, **2001**, *123*, 9490.
- [118] Meiler, J.; Peti, W.; Griesinger, C. *J Biomol NMR.*, **2000**, *17*, 283.
- [119] Hus, J.; Prompers, J.; Bruschweiler, R. *Journal of Magnetic Resonance*, **2002**, *157*, 119.
- [120] Jung, Y.; Zweckstetter, M. *J Biomol NMR.*, **2004**, *30*, 25.
- [121] Langmead, C.; Yan, A.; Lilien, R.; Wang, L.; Donald, B. *J Comput Biol.*, **2004**, *11*, 277.
- [122] Langmead, C.; Donald, B. *J Biomol NMR.*, **2004**, *29*, 111.
- [123] Rohl, C.; Baker, D. *J Am Chem Soc.*, **2002**, *124*, 2723.
- [124] Haliloglu, T.; Kolinski, A.; Skolnick, J. *Biopolymers*, **2003**, *70*, 548.
- [125] Wedemeyer, W.; Rohl, C.; Scheraga, H. *J Biomol NMR.*, **2002**, *22*, 137.
- [126] Moltke, S.; Grzesiek, S. *J Biomol NMR.*, **1999**, *15*, 77.
- [127] Sass, H.; Musco, G.; Stahl, S.; Wingfield, P.; Grzesiek, S. *J Biomol NMR.*, **2001**, *21*, 275.
- [128] Meiler, J.; Blomberg, N.; Nilges, M.; Griesinger, C. *J Biomol NMR.*, **2000**, *16*, 245.
- [129] Skrynnikov, N.; Kay, L. *J Biomol NMR.*, **2000**, *18*, 239.
- [130] Ramirez, B.; Bax, A. *J Am Chem Soc.*, **1998**, *120*, 9106.
- [131] Mueller, G.; Choy, W.; Skrynnikov, N.; Kay, L. *J Biomol NMR.*, **2000**, *18*, 183.
- [132] Hus, J.; Marion, D.; Blackledge, M. *J Am Chem Soc.*, **2001**, *123*, 1541.
- [133] Beraud, S.; Bersch, B.; Brutscher, B.; Gans, P.; Barras, F.; Blackledge, M. *J Am Chem Soc.*, **2002**, *124*, 13709.
- [134] Wang, L.; Donald, B. *J Biomol NMR.*, **2004**, *29*, 223.
- [135] Dosset, P.; Hus, J.; Marion, D.; Blackledge, M. *J Biomol NMR.*, **2001**, *20*, 223.
- [136] Hus, J.; Bruschweiler, R. *J Biomol NMR.*, **2002**, *24*, 123.
- [137] Hus, J.; Peti, W.; Griesinger, C.; Bruschweiler, R. *J Am Chem Soc.*, **2003**, *125*, 5596.
- [138] Nomura, K.; Kainosho, M. *Journal of Magnetic Resonance*, **2002**, *154*, 146.
- [139] Zweckstetter, M.; Bax, A. *J Am Chem Soc.*, **2000**, *122*, 3791.
- [140] Fernandes, M.; Bernado, P.; Pons, M.; de la Torre, J. G. *J Am Chem Soc.*, **2001**, *123*, 12037.
- [141] Almond, A.; Axelsen, J. *J Am Chem Soc.*, **2002**, *124*, 9986.
- [142] Zweckstetter, M.; Hummer, G.; Bax, A. *Biophys J.*, **2004**, *86*, 3444.
- [143] Tate, S.; Shimahara, H.; Utsunomiya-Tate, N. *Journal of Magnetic Resonance*, **2004**, *171*, 284.
- [144] Hus, J.; Marion, D.; Blackledge, M. *Journal of Molecular Biology*, **2000**, *298*, 927.
- [145] Case, D. *J Biomol NMR.*, **1999**, *15*, 95.
- [146] Gronwald, W.; Boyko, R.; Sonnichsen, F.; Wishart, D.; Sykes, B. *J Biomol NMR.*, **1997**, *10*, 165.
- [147] Gronwald, W.; Willard, L.; Jellard, T.; Boyko, R.; Rajarathnam, K.; Wishart, D.; Sonnichsen, F.; Sykes, B. *J Biomol NMR*, **1998**, *12*, 395.
- [148] Wishart, D.; Watson, M.; Boyko, R.; Sykes, B. *J Biomol NMR.*, **1997**, *10*, 329.
- [149] Wang, Y.; Jardetzky, O. *J Am Chem Soc.*, **2002**, *124*, 14075.
- [150] Sharman, G.; Griffiths-Jones, S.; Jourdan, M.; Searle, M. *J Am Chem Soc.*, **2001**, *123*, 12318.
- [151] Huang, K.; Andrec, M.; Heald, S.; Blake, P.; Prestegard, J. *J Biomol NMR.*, **1997**, *10*, 45.
- [152] Labudde, D.; Leitner, D.; Kruger, M.; Oschkinat, H. *J Biomol NMR.*, **2003**, *25*, 41.
- [153] Pons, J.; Delsuc, M. *J Biomol NMR.*, **1999**, *15*, 15.
- [154] Marin, A.; Malliavin, T.; Nicolas, P.; Delsuc, M. *J Biomol NMR.*, **2004**, *30*, 47.
- [155] Hung, L.; Samudrala, R. *Protein Sci.*, **2003**, *12*, 288.
- [156] Wang, Y.; Jardetzky, O. *Protein Sci.*, **2002**, *11*, 852.
- [157] Mielke, S.; Krishnan, V. *Bioinformatics*, **2003**, *19*, 2054.
- [158] Sibley, A.; Cosman, M.; Krishnan, V. *Biophys J.*, **2003**, *84*, 1223.
- [159] Wang, Y.; Jardetzky, O. *J Biomol NMR.*, **2004**, *28*, 327.
- [160] Beger, R.; Bolton, P. *J Biomol NMR.*, **1997**, *10*, 129.
- [161] Cornilescu, G.; Delaglio, F.; Bax, A. *J Biomol NMR.*, **1999**, *13*, 289.
- [162] Schubert, M.; Labudde, D.; Oschkinat, H.; Schmieder, P. *J Biomol NMR.*, **2002**, *24*, 149.
- [163] Schwarzingler, S.; Kroon, G.; Foss, T.; Chung, J.; Wright, P.; Dyson, H. *J Am Chem Soc.*, **2001**, *123*, 2970.

- [164] Banci, L.; Bertini, I.; Savellini, G.; Romagnoli, A.; Turano, P.; Cremonini, M.; Luchinat, C.; Gray, H. *Proteins*, **1997**, *29*, 68.
- [165] Bertini, I.; Longinetti, M.; Luchinat, C.; Parigi, G.; Sgheri, L. *J Biomol NMR*, **2002**, *22*, 123.
- [166] Lipsitz, R.; Tjandra, N. *J Am Chem Soc.*, **2001**, *123*, 11065.
- [167] Lipsitz, R.; Tjandra, N. *Journal of Magnetic Resonance*, **2003**, *164*, 171.
- [168] Iwahara, J.; Schwieters, C.; Clore, G. *J Am Chem Soc.*, **2004**, *126*, 5879.
- [169] Iwadate, M.; Asakura, T.; Williamson, M. *J Biomol NMR*, **1999**, *13*, 199.
- [170] Sitkoff, D.; Case, D. *J Am Chem Soc.*, **1997**, *119*, 12262.
- [171] Xu, X.; Case, D. *J Biomol NMR*, **2001**, *21*, 321.
- [172] Xu, X.; Case, D. *Biopolymers*, **2002**, *65*, 408.
- [173] Williamson, M.; Refaee, M. SHIFTCALC, **2004**, University of Sheffield, United Kingdom.
- [174] Neal, S.; Nip, A.; Zhang, H.; Wishart, D. *J Biomol NMR*, **2003**, *26*, 215.
- [175] Markwick, P.; Sattler, M. *J Am Chem Soc.*, **2004**, *126*, 11424.
- [176] Boyd, J.; Domene, C.; Redfield, C.; Ferraro, M.; Lazzaretto, P. *J Am Chem Soc.*, **2003**, *125*, 9556.
- [177] Thomas, A.; Milon, A.; Brasseur, R. *Proteins*, **2004**, *56*, 102.
- [178] Sharma, D.; Rajarathnam, K. *J Biomol NMR*, **2000**, *18*, 165.
- [179] Williamson, M. *Proteins*, **2003**, *53*, 731.
- [180] Baxter, N.; Williamson, M. *J Biomol NMR*, **1997**, *9*, 359.
- [181] Cierpicki, T.; Otlewski, J. *J Biomol NMR*, **2001**, *21*, 249.
- [182] Williamson, M.; Akasaka, K.; Refaee, M. *Prot Sci*, **2003**, *12*, 1971.
- [183] Refaee, M.; Tezuka, T.; Akasaka, K.; Williamson, M. *J Mol Biol*, **2003**, *327*, 857.
- [184] Iwadate, M.; Asakura, T.; Dubovskii, P.; Yamada, H.; Akasaka, K.; Williamson, M. *J Biomol NMR*, **2001**, *19*, 115.
- [185] Doreleijers, J.; Rullmann, J.; Kaptein, R. *Journal of Molecular Biology*, **1998**, *281*, 149.
- [186] Doreleijers, J.; Raves, M.; Rullmann, T.; Kaptein, R. *J Biomol NMR*, **1999**, *14*, 123.
- [187] Doreleijers, J.; Vriend, G.; Raves, M.; Kaptein, R. *Proteins*, **1999**, *37*, 404.
- [188] de Alba, E.; Tjandra, N. *Journal of Magnetic Resonance*, **2000**, *144*, 367.
- [189] Chung, S.; Subbiah, S. *Proteins*, **1999**, *35*, 184.
- [190] Zidek, L.; Padrta, P.; Chmelik, J.; Sklenar, V. *Journal of Magnetic Resonance*, **2003**, *162*, 385.
- [191] Pearlman, D. *J Biomol NMR*, **1999**, *13*, 325.
- [192] Jensen, P.; Axelsen, J.; Lerche, M.; Poulsen, F. *J Biomol NMR*, **2004**, *28*, 31.
- [193] Nabuurs, S.; Spronk, C.; Krieger, E.; Maassen, H.; Vriend, G.; Vuister, G. *J Am Chem Soc.*, **2003**, *125*, 12026.
- [194] Dzakula, Z.; Juranic, N.; DeRider, M.; Westler, W.; Macura, S.; Markley, J. *Journal of Magnetic Resonance*, **1998**, *135*, 454.
- [195] Gronwald, W.; Kirchhofer, R.; Gorler, A.; Kremer, W.; Gansmeier, B.; Neidig, K.; Kalbitzer, H. *J Biomol NMR*, **2000**, *17*, 137.
- [196] Willard, L.; Ranjan, A.; Zhang, H.; Monzavi, H.; Boyko, R.; Sykes, B.; Wishart, D. *Nucl Acid Res.*, **2003**, *31*, 3316.
- [197] Huang, Y.; Powers, R.; Montelione, G. *J Am Chem Soc*, **2005**, *127*, 1665.
- [198] Schwieters, C.; Clore, G. *J Biomol NMR*, **2002**, *23*, 221.
- [199] Harris, T.; Mildvan, A. *Proteins*, **1999**, *35*, 275.
- [200] Alexandrescu, A.; Snyder, D.; Abildgaard, F. *Protein Sci.*, **2001**, *10*, 1856.
- [201] Lipsitz, R.; Sharma, Y.; Brooks, B.; Tjandra, N. *J Am Chem Soc.*, **2002**, *124*, 10621.
- [202] Grishaev, A.; Bax, A. *J Am Chem Soc.*, **2004**, *126*, 7281.
- [203] Clore, G. *Proc Natl Acad Sci U S A.*, **2000**, *97*, 9021.
- [204] Clore, G.; Schwieters, C. *J Am Chem Soc.*, **2003**, *125*, 2902.
- [205] Dobrodumov, A.; Gronenborn, A. *Proteins*, **2003**, *53*, 18.
- [206] O'Donoghue, S.; Chang, X.; Abseher, R.; Nilges, M.; Led, J. *J Biomol NMR*, **2000**, *16*, 93.
- [207] Dominguez, C.; Boelens, R.; Bonvin, A. *J Am Chem Soc.*, **2003**, *125*, 1731.
- [208] Eisenmesser, E.; Zhabell, A.; Post, C. *J Biomol NMR*, **2000**, *17*, 17.
- [209] Zhabell, A.; Post, C. *J Biomol NMR*, **2002**, *22*, 303.
- [210] Zhabell, A.; Post, C. *Proteins*, **2002**, *46*, 295.
- [211] Jayalakshmi, V.; Krishna, N. *Journal of Magnetic Resonance*, **2004**, *168*, 36.
- [212] McCoy, M.; Wyss, D. *J Am Chem Soc.*, **2002**, *124*, 2104.
- [213] McCoy, M.; Wyss, D. *J Am Chem Soc.*, **2002**, *124*, 11758.
- [214] Wang, B.; Raha, K.; Merz, K. *J Am Chem Soc.*, **2004**, *126*, 11430.
- [215] Peng, C.; Unger, S.; Filipp, F.; Sattler, M.; Szalma, S. *J Biomol NMR*, **2004**, *29*, 491.
- [216] Iwadate, M.; Asakura, T.; Williamson, M. *Eur J Biochem*, **1998**, *257*, 479.
- [217] Wimmer, R.; Muller, N.; Petersen, S. *J Biomol NMR*, **1997**, *9*, 101.
- [218] Doreleijers, J.; Mading, S.; Maziuk, D.; Sojourner, K.; Yin, L.; Zhu, J.; Markley, J.; Ulrich, E. *J Biomol NMR*, **2003**, *26*, 139.
- [219] Zhang, H.; Neal, S.; Wishart, D. *J Biomol NMR*, **2003**, *25*, 173.
- [220] Baran, M.; Moseley, H.; Sahota, G.; Montelione, G. *J Biomol NMR*, **2002**, *24*, 113.
- [221] Pons, J.; Malliavin, T.; Tramesel, D.; Delsuc, M. *Bioinformatics*, **2004**, *20*, 3707.
- [222] Nabuurs, S.; Nederveen, A.; Vranken, W.; Doreleijers, J.; Bonvin, A.; Vuister, G.; Vriend, G.; Spronk, C. *Proteins*, **2004**, *55*, 483.
- [223] Nederveen, A.; Doreleijers, J.; Vranken, W.; Miller, Z.; Spronk, C.; Nabuurs, S.; Guntert, P.; Livny, M.; Markley, J.; Nilges, M.; Ulrich, E.; Kaptein, R.; Bonvin, A. *Proteins*, **2005**, in press.
- [224] Markley, J.; Bax, A.; Arata, Y.; Hilbers, C.; Kaptein, R.; Sykes, B.; Wright, P.; Wuthrich, K. *J Biomol NMR*, **1998**, *12*, 1.
- [225] Markley, J.; Bax, A.; Arata, Y.; Hilbers, C.; Kaptein, R.; Sykes, B.; Wright, P.; Wuthrich, K. *Journal of Molecular Biology*, **1998**, *280*, 933.
- [226] Markley, J.; Bax, A.; Arata, Y.; Hilbers, C.; Kaptein, R.; Sykes, B.; Wright, P.; Wuthrich, K. *Eur J Biochem*, **1998**, *256*, 1.
- [227] Fogh, R.; Ionides, J.; Ulrich, E.; Boucher, W.; Vranken, W.; Linge, J.; Habeck, M.; Rieping, W.; Bhat, T.; Westbrook, J.; Henrick, K.; Gilliland, G.; Berman, H.; Thornton, J.; Nilges, M.; Markley, J.; Laue, E. *Nat Struct Biol*, **2002**, *9*, 416.
- [228] Fogh, R.; Boucher, W.; Vranken, W.; Pajon, A.; Stevens, T.; Bhat, T.; Westbrook, J.; Ionides, J.; Laue, E. *Bioinformatics*, **2005**, in press.
- [229] Lipari, G.; Szabo, A. *J Am Chem Soc*, **1982**, *104*, 4546.
- [230] Clore, G.; Szabo, A.; Bax, A.; Kay, L.; Driscoll, P. *J Am Chem Soc*, **1990**, *112*, 4989.
- [231] Peng, J.; Wagner, G. *Biochemistry*, **1992**, *31*, 8571.
- [232] Zhu, L.; Dyson, H.; Wright, P. *J Biomol NMR*, **1998**, *11*, 17.
- [233] Jayalakshmi, V.; Krishna, N. *Journal of Magnetic Resonance*, **2002**, *155*, 106.
- [234] LeMaster, D. *J Biomol NMR*, **1997**, *9*, 79.
- [235] Zolnai, Z.; Juranic, N.; Macura, S. *J Biomol NMR*, **1998**, *12*, 333.
- [236] Malliavin, T. E.; Desvaux, H.; Aumelas, A.; Chavanieu, A.; Delsuc, M. A. *Journal of Magnetic Resonance*, **1999**, *140*, 189.
- [237] Schleucher, J.; Wijmenga, S. *J Am Chem Soc.*, **2002**, *124*, 5881.
- [238] Feenstra, K.; Peter, C.; Scheek, R.; van Gunsteren, W.; Mark, A. *J Biomol NMR*, **2002**, *23*, 181.
- [239] Lee, A.; Wand, A. *J Biomol NMR*, **1999**, *13*, 101.
- [240] Yao, S.; Hinds, M.; Norton, R. *Journal of Magnetic Resonance*, **1998**, *131*, 347.
- [241] Mispelter, J.; Izadi-Pruneyre, N.; Quiniou, E.; Adjadj, E. *Journal of Magnetic Resonance*, **2000**, *143*, 229.
- [242] Idiyatullin, D.; Daragan, V.; Mayo, K. *Journal of Magnetic Resonance*, **2001**, *152*, 132.
- [243] Larsson, G.; Martinez, G.; Schleucher, J.; Wijmenga, S. *J Biomol NMR*, **2003**, *27*, 291.
- [244] McMahon, M.; Oldfield, E. *J Biomol NMR*, **1999**, *13*, 133.
- [245] Pawley, N.; Wang, C.; Koide, S.; Nicholson, L. *J Biomol NMR*, **2001**, *20*, 149.
- [246] Clore, G.; Gronenborn, A.; Szabo, A.; Tjandra, N. *J Am Chem Soc.*, **1998**, *120*, 4889.
- [247] Malliavin, T.; Giudice, E. *Biopolymers*, **2002**, *63*, 335.
- [248] Dossset, P.; Hus, J.; Blackledge, M.; Marion, D. *J Biomol NMR*, **2000**, *16*, 23.
- [249] Walker, O.; Varadan, R.; Fushman, D. *Journal of Magnetic Resonance*, **2004**, *168*, 336.
- [250] Ghose, R.; Fushman, D.; Cowburn, D. *Journal of Magnetic Resonance*, **2001**, *149*, 204.
- [251] Andrec, M.; Inman, K.; Weber, D.; Levy, R.; Montelione, G. *Journal of Magnetic Resonance*, **2000**, *146*, 66.
- [252] Millet, O.; Pons, M. *J Biomol NMR*, **2001**, *19*, 181.
- [253] Jin, D.; Andrec, M.; Montelione, G.; Levy, R. *J Biomol NMR*, **1998**, *12*, 471.
- [254] Pawley, N.; Gans, J.; Nicholson, L. *J Biomol NMR*, **2002**, *24*, 215.
- [255] Viles, J.; Duggan, B.; Zaborowski, E.; Schwarzinger, S.; Huntley, J.; Kroon, G.; Dyson, H.; Wright, P. *J Biomol NMR*, **2001**, *21*, 1.

- [256] Cole, R.; Loria, J. *J Biomol NMR.*, **2003**, *26*, 203.
- [257] d'Auvergne, E.; Gooley, P. *J Biomol NMR.*, **2003**, *25*, 25.
- [258] Andreac, M.; Montelione, G.; Levy, R. *J Biomol NMR.*, **2000**, *18*, 83.
- [259] Atkinson, R.; Lefevre, J. *J Biomol NMR.*, **1999**, *13*, 83.
- [260] Mayo, K.; Daragan, V.; Idiyatullin, D.; Nesmelova, I. *Journal of Magnetic Resonance*, **2000**, *146*, 188.
- [261] Renner, C.; Moroder, L.; Holak, T. *Journal of Magnetic Resonance*, **2001**, *151*, 32.
- [262] Idiyatullin, D.; Daragan, V.; Mayo, K. *Journal of Magnetic Resonance*, **2003**, *161*, 118.
- [263] de la Torre, J.; Huertas, M.; Carrasco, B. *Journal of Magnetic Resonance*, **2000**, *147*, 138.
- [264] Bernado, P.; de la Torre, J.; Pons, M. *J Biomol NMR.*, **2002**, *23*, 139.
- [265] Bernado, P.; Akerud, T.; de la Torre, J.; Akke, M.; Pons, M. *J Am Chem Soc.*, **2003**, *125*, 916.
- [266] Orekhov, V.; Korzhnev, D.; Pervushin, K.; Hoffmann, E.; Arseniev, A. *J Biomol Struct Dyn.*, **1999**, *17*, 157.
- [267] Krishnan, V.; Cosman, M. *J Biomol NMR.*, **1998**, *12*, 177.
- [268] Goodman, J.; Pagel, M.; Stone, M. *Journal of Molecular Biology*, **2000**, *295*, 963.
- [269] Zhang, F.; Bruschweiler, R. *J Am Chem Soc.*, **2002**, *124*, 12654.
- [270] Mittermaier, A.; Kay, L.; Forman-Kay, J. *J Biomol NMR.*, **1999**, *13*, 181.
- [271] Mittermaier, A.; Davidson, A.; Kay, L. *J Am Chem Soc.*, **2003**, *125*, 9004.
- [272] Chang, S.; Tjandra, N. *J Am Chem Soc.*, **2001**, *123*, 11484.
- [273] Baber, J.; Szabo, A.; Tjandra, N. *J Am Chem Soc.*, **2001**, *123*, 3953.
- [274] Barbar, E.; Hare, M.; Daragan, V.; Barany, G.; Woodward, C. *Biochemistry*, **1998**, *37*, 7822.
- [275] Yang, D.; Mok, Y.; Forman-Kay, J.; Farrow, N.; Kay, L. *J Mol Biol*, **1997**, *272*, 790.
- [276] Ochsenbein, F.; Neumann, J.; Guittet, E.; van Heijenoort, C. *Protein Sci.*, **2002**, *11*, 957.
- [277] Daragan, V.; Mayo, K. *Journal of Magnetic Resonance*, **1998**, *130*, 329.
- [278] Idiyatullin, D.; Krushelnitsky, A.; Nesmelova, I.; Blanco, F.; Daragan, V.; Serrano, L.; Mayo, K. *Protein Sci.*, **2000**, *9*, 2118.
- [279] Fischer, M.; Zeng, L.; Majumdar, A.; Zuiderweg, E. *Proc Natl Acad Sci U S A.*, **1998**, *95*, 8016.
- [280] Mayer, K.; Earley, M.; Gupta, S.; Pichumani, K.; Regan, L.; Stone, M. *Nat Struct Biol.*, **2003**, *10*, 962.
- [281] Lemaster, D. *J Am Chem Soc.*, **1999**, *121*, 1726.
- [282] Vugmeyster, L.; Raleigh, D.; Palmer, A.; Vugmeyster, B. *J Am Chem Soc.*, **2003**, *125*, 8400.
- [283] Abergel, D.; Bodenhausen, G. *J Chem Phys.*, **2004**, *121*, 761.
- [284] Tugarinov, V.; Liang, Z.; Shapiro, Y.; Freed, J.; Meirovitch, E. *J Am Chem Soc.*, **2001**, *123*, 3055.
- [285] Clore, G.; Schwieters, C. *J Am Chem Soc.*, **2004**, *126*, 2923.
- [286] Tolman, J.; Al-Hashimi, H.; Kay, L.; Prestegard, J. *J Am Chem Soc.*, **2001**, *123*, 1416.
- [287] Tolman, J. *J Am Chem Soc.*, **2002**, *124*, 12020.
- [288] Briggman, K.; Tolman, J. *J Am Chem Soc.*, **2003**, *125*, 10164.
- [289] Meiler, J.; Prompers, J.; Peti, W.; Griesinger, C.; Bruschweiler, R. *J Am Chem Soc.*, **2001**, *123*, 6098.
- [290] Peti, W.; Meiler, J.; Bruschweiler, R.; Griesinger, C. *J Am Chem Soc.*, **2002**, *124*, 5822.
- [291] Clore, G.; Schwieters, C. *Biochemistry*, **2004**, *43*, 10678.
- [292] Fredriksson, K.; Louhivuori, M.; Permi, P.; Annala, A. *J Am Chem Soc.*, **2004**, *126*, 12646.
- [293] Bernado, P.; Blackledge, M. *J Am Chem Soc.*, **2004**, *126*, 7760.
- [294] Ohnishi, S.; Shortle, D. *Proteins*, **2003**, *50*, 546.
- [295] Shortle, D.; Ackerman, M. *Science*, **2001**, *293*, 487.
- [296] Louhivuori, M.; Paakkonen, K.; Fredriksson, K.; Permi, P.; Lounila, J.; Annala, A. *J Am Chem Soc.*, **2003**, *125*, 15647.
- [297] Prompers, J.; Bruschweiler, R. *J Am Chem Soc.*, **2001**, *123*, 7305.
- [298] Prompers, J.; Bruschweiler, R. *J Am Chem Soc.*, **2002**, *124*, 4522.
- [299] Showalter, S.; Hall, K. *Journal of Molecular Biology*, **2002**, *322*, 533.
- [300] Prompers, J.; Bruschweiler, R. *J Am Chem Soc.*, **2002**, *124*, 4522.
- [301] Philippopoulos, M.; Mandel, A.M.; Palmer, A.G.; Lim, C. *Proteins*, **1997**, *28*, 481.
- [302] Wrabl, J.; Shortle, D.; Woolf, T. *Proteins*, **2000**, *38*, 123.
- [303] Pfeiffer, S.; Fushman, D.; Cowburn, D. *J Am Chem Soc.*, **2001**, *123*, 3021.
- [304] Chen, J.; Brooks, C.; Wright, P. *J Biomol NMR.*, **2004**, *29*, 243.
- [305] Daura, X.; Antes, I.; van Gunsteren, W.; Thiel, W.; Mark, A. *Proteins*, **1999**, *36*, 542.
- [306] Carstens, H.; Renner, C.; Milbradt, A.; Moroder, L.; Tavan, P. *Biochemistry*, **2005**, *44*, 4829.
- [307] Burgi, R.; Pitera, J.; van Gunsteren, W. *J Biomol NMR.*, **2001**, *19*, 305.
- [308] Chalaux, F.; O'Donoghue, S.; Nilges, M. *Proteins*, **1999**, *34*, 453.
- [309] Schneider, T.; Brunger, A.; Nilges, M. *J Mol Biol*, **1999**, *285*, 727.
- [310] Markwick, P.; Sprangers, R.; Sattler, M. *J Am Chem Soc.*, **2003**, *125*, 644.
- [311] Freedberg, D.; Venable, R.; Rossi, A.; Bull, T.; Pastor, R. *J Am Chem Soc.*, **2004**, *126*, 10478.
- [312] Vila, J.; Ripoll, D.; Baldoni, H.; Scheraga, H. *J Biomol NMR.*, **2002**, *24*, 245.
- [313] Vila, J.; Baldoni, H.; Ripoll, D.; Scheraga, H. *J Biomol NMR*, **2003**, *26*, 113.
- [314] Cregut, D.; Serrano, L. *Protein Sci.*, **1999**, *8*, 271.
- [315] Colombo, G.; Roccatano, D.; Mark, A. *Proteins*, **2002**, *46*, 380.
- [316] Smith, L.; Jones, R.; van Gunsteren, W. *Proteins*, **2005**, *58*, 439.
- [317] Groth, M.; Malicka, J.; Rodziewicz-Motowidlo, S.; Czaplowski, C.; Klaudel, L.; Wicz, W.; Liwo, A. *Biopolymers*, **2001**, *60*, 79.
- [318] Floquet, N.; Hery-Huynh, S.; Dauchez, M.; Derreumaux, P.; Tamburro, A.; Alix, A. *Biopolymers*, **2004**, *76*, 266.
- [319] Bosques, C.; Tschampel, S.; Woods, R.; Imperiali, B. *J Am Chem Soc.*, **2004**, *126*, 8421.
- [320] Mihailescu, D.; Reed, J.; Smith, J. *Biopolymers*, **2003**, *70*, 121.
- [321] Yu, H.; Daura, X.; van Gunsteren, W.F. *Proteins*, **2004**, *54*, 116.
- [322] Fioroni, M.; Diaz, M.; Burger, K.; Berger, S. *J Am Chem Soc.*, **2002**, *124*, 7737.
- [323] Canalia, M.; Malliavin, T.; Kremer, W.; Kalbitzer, H. *Biopolymers*, **2004**, *74*, 377.
- [324] Philippopoulos, M.; Lim, C. *Proteins*, **1999**, *36*, 87.
- [325] Fan, H.; Mark, A. *Proteins*, **2003**, *53*, 111.
- [326] Stocker, U.; van Gunsteren, W. *Proteins*, **2000**, *40*, 145.
- [327] Haliloglu, T.; Bahar, I. *Proteins*, **1999**, *37*, 654.
- [328] Abseher, R.; Horstink, L.; Hilbers, C.; Nilges, M. *Proteins*, **1998**, *31*, 370.
- [329] Scott, W.; Mark, A.; van Gunsteren, W. *J Biomol NMR.*, **1998**, *12*, 501.
- [330] Hess, B.; Scheek, R. *Journal of Magnetic Resonance*, **2003**, *164*, 19.
- [331] Frimurer, T.; Peters, G.; Sorensen, M.; Led, J.; Olsen, O. *Protein Sci.*, **1999**, *8*, 25.
- [332] Groth, M.; Malicka, J.; Czaplowski, C.; Oldziej, S.; Lankiewicz, L.; Wicz, W.; Liwo, A. *J Biomol NMR.*, **1999**, *15*, 315.
- [333] Lindorff-Larsen, K.; Kristjansdottir, S.; Teilum, K.; Fieber, W.; Dobson, C.; Poulsen, F.; Vendruscolo, M. *J Am Chem Soc.*, **2004**, *126*, 3291.
- [334] Dedmon, M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. *J Am Chem Soc.*, **2005**, *127*, 476.
- [335] Best, R.; Vendruscolo, M. *J Am Chem Soc.*, **2004**, *126*, 8090.
- [336] Best, R.; Clarke, J.; Karplus, M. *J Am Chem Soc.*, **2004**, *126*, 7734.
- [337] Lindorff-Larsen, K.; Best, R.; Depristo, M.; Dobson, C.; Vendruscolo, M. *Nature*, **2005**, *433*, 128.