ELSEVIER

Progress in Nuclear Magnetic Resonance Spectroscopy 45 (2004) 315–337

PROGRESS IN NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

www.elsevier.com/locate/pnmrs

Validation of protein structures derived by NMR spectroscopy

Chris A.E.M. Spronk[a], Sander B. Nabuurs[a], Elmar Krieger[a], Gert Vriend[a], Geerten W. Vuister[b,*]

[a]*Centre for Molecular and Biomolecular Informatics, IMM, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands*
[b]*Department of Biophysical Chemistry, IMM, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands*

Received 15 July 2004

Contents

* Corresponding author. Address: Department of Biophysical Chemistry, Radboud University Nijmegen, Toernooiveld 1, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands. Tel.: +31 24 3652321; fax: +31 24 3652112.
E-mail address: g.vuister@science.ru.nl (G.W. Vuister).

0079-6565/$ - see front matter
doi:10.1016/j.pnmrs.2004.08.003

1. Introduction

The availability of three-dimensional (3D) structures of proteins and nucleic acids has greatly improved our insights into biological processes and macromolecular structure–function relationships. The two main experimental techniques used to derive structure models of biological macromolecules are X-ray diffraction and solution NMR spectroscopy. X-ray diffraction techniques were introduced for studying biomolecular structures some 5 decades ago with the determination of the 3D structures of DNA [1] and myoglobin [2]. It has become a well-established and developed technique, yielding ever more accurate structures with the use of improved equipment, automation and analysis tools [3]. The vast number of X-ray structures that are now available from the Protein Data Bank (PDB) [4,5] allows for the selection of well-determined and validated protein structure models. These models provide a reference set from which structural preferences of proteins can be derived and which in turn can be used for the analysis of newly derived structure models.

High-resolution NMR spectroscopy is a much younger technique and was first used as a tool for structure determination of proteins and nucleic acids at atomic resolution some 20 years ago [6,7]. Since then, approximately 4000 NMR-derived structures have been deposited at the PDB, which amounts to $\sim$15% of the total number of structures. More importantly, however, the number of unique folds determined by NMR spectroscopy is much higher, i.e. about 25% [8], and NMR structures thus represent a significant source of information on protein folds. Apart from structural information, NMR also provides information about the dynamics of proteins and nucleic acids. This unique combination of structure and dynamics renders NMR a powerful tool for the analysis of dynamic events, such as folding transitions in proteins, which are of major importance for biological function.

One of the most important aspects of all structural models is their reliability or accuracy [9,10]. Structure models are not merely to be published and deposited in public databases, but are supposed to explain biological function and serve as guides for further study. Hence, their usefulness depends on the quality of the model. Structure models can contain errors at many different levels, even to the extent that the overall fold can be completely wrong [8]. Other, less serious errors may occur in the details of the structure model, rendering it less precise, and too inaccurate for a molecular modeling type of experiment, yet of sufficient relevance to help in understanding the biological function.

It is important always to realize that a 3D representation of a molecule is not a real *structure*, but a *model* of that structure built to represent experimental data [8]. The model of the structure may or may not be a useful representation of the truth, depending on the quality of the data, and the skill of the experimenter who translates the experimental data into structural coordinates. In this light we must consider the fact that in any experiment errors will occur. These errors may be random, which will influence the reproducibility and the precision of the data and resulting structures, or systematic, which will influence the accuracy of the structure model. It is important to distinguish between these two types of errors, as well as to differentiate carefully between precision and accuracy.

Historically, the quality of NMR-derived structures is lower when compared to X-ray structures, which is largely caused by the much lower information-content of the NMR data. Until some years ago, distance and dihedral angle restraints derived from NOE and *J*-coupling experiments were the main sources of information in NMR structure determination. However, technological and methodological developments have provided new means of improving the quality of NMR structures. Valuable structural restraints can be obtained from experiments that allow measurement of scalar ^{3}J-couplings across hydrogen bonds [11,12], residual dipolar couplings [13], carbonyl chemical shift anisotropies [14,15] and paramagnetism [16,17]. In addition, backbone dihedral angle constraints can be derived from chemical shifts in combination with database searches [18].

Parallel to the developments in experimental NMR, powerful automation and computational methods for spectrum interpretation have become available. Automated NOE assignment and structure calculation [19–23], and structure calculations using the concept of proton-densities [24–26] have increased the amount of information that can be extracted from experimental data and added to the reliability of NMR structures by reducing the need for manual interpretation. Finally, refinement of NMR structures in explicit solvent (see Section 6) has been shown to be an effective method to improve structure quality and is now routinely employed as a final step in NMR structure determination [27,28].

In this review we will discuss both traditional and more recent methods that are commonly used to address the quality of NMR derived structure models. We have chosen to follow the main steps in a typical NMR structure determination project. In doing so, a brief introduction on NMR structure calculation and selection procedures is given, followed by validation in terms of experimental data and an assessment of the precision and accuracy of structure models. A major part of this review is dedicated to validation of the overall geometric quality of NMR structures, which is based on comparison to high-resolution X-ray structures. All aspects of structure validation are illustrated in detail using selected actual examples.

2. NMR structure determination

2.1. Structure calculation procedures

Structure determination using NMR data is fundamentally different from that using X-ray data, and relies on many more assumptions. Here, we will only briefly discuss the major steps involved in NMR structure derivation. For a more elaborate discussion on this topic we refer to two reviews on structure calculation protocols and methods for automated structure determination [29,30].

NMR structure determination involves the following steps (cf. Fig. 1): acquisition of experimental data, assignment of resonances to atoms in the molecular structure, assignment of observed interactions to atom pairs and translation of these interactions into conformational restraints, such as distance restraints, torsion angle restraints and orientational restraints. From the resulting set of conformational restraints an initial set of structures is calculated, which is then manually or automatically analyzed to identify errors in the restraints and assignments. Such errors are typically identified by a violation analysis of the restraints in the ensemble of structures, followed by checking the restraints against the experimental data. The structure and error analysis is then used to decide on whether the restraints should be modified, i.e. whether false assignments were made or whether the bounds of the restraints should be adjusted. It is here that NMR structure determination is especially vulnerable to human error and possible bias, such as unjustified removal of violating restraints, particularly in non-automated methods.

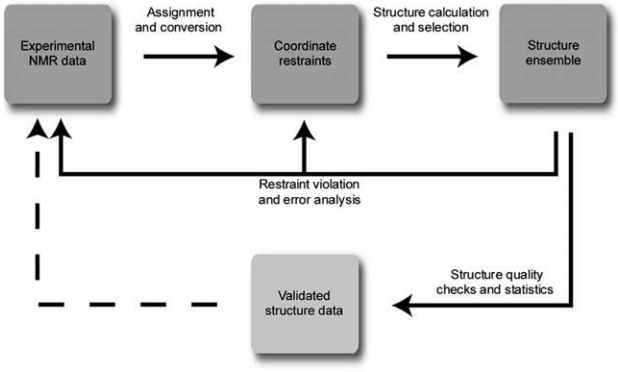

Fig. 1. Schematic representation of the basic steps in NMR structure determination.

The steps from assignment to structure and error analysis are generally performed in an iterative manner, until an optimized data set is obtained from which, hopefully, all errors in the assignment procedure have been removed. This optimized data set is then used to calculate a final ensemble of structure models that is subsequently subjected to validation in order to obtain an indication of the quality and structural statistics such as a measure of the fit of the structures to the experimental data, positional RMSD values, Ramachandran plot scores and others. As will be pointed out in this review, structure validation is a very important step in the NMR structure determination process as it provides an indication of the reliability of the model. In practice it appears that validation is often only used to obtain structural statistics, i.e. a description of the resulting structure. However, validation often points to possible problems, unique conformations that may either be real or resulting from errors, and as such should be used to re-examine carefully the original data or structure calculation procedures.

2.2. Structure selection

The first step in structure validation is the selection of NMR structures from a large ensemble of calculated structures. This is a crucial step, since it will affect all further analyses. Unfortunately, the absence of a consensus on structure selection procedures introduces a subjective bias in the final result. The two most widely used procedures are based on either the number of experimental restraint violations or the energy of the structure. Both methods involve arbitrary cutoff criteria, which can be user-defined.

For both procedures, the interplay between the terms in the applied force field and the experimental restraints used to calculate the structures presents a non-trivial factor. Provided that the force field itself is internally consistent, a low energy conformation can always be found in the absence of any experimental restraints. In the presence of experimental restraints and assuming the molecule has a single global conformation, we distinguish the following four situations:

1. Ideally, the experimental restraints are internally consistent and compatible with the terms in the force field.

It should be possible to obtain a solution that satisfies all restraints without introducing geometric distortions into the structure.
2. The experimental restraints are internally inconsistent due to false assignments or conversion. All calculated structures will contain experimental restraint violations and an error analysis is required.
3. The experimental restraints are incompatible with the terms in the force field. The result depends on the balance between the terms in the force field and the weighting of the experimental restraints. Increasing the relative weight on the experimental restraints may force the structure to fit the data, but introduces geometric strain in the molecule.
4. A combination of situations 2 and 3 above.

Clearly, an analysis of the resulting structures should not be confined to experimental restraint violations alone since this may result in structures with poor geometric quality, albeit without any experimental restraint violations. The geometric quality of a structure can be checked by inspecting the energy of the different force field terms. Alternatively, direct geometric validation can be performed using structure calculation software or by specialized structure validation software.

Structure selection using an energy cutoff criterion is expected to yield similar ensembles as compared to using violation cutoff criteria, since the energy of the structures is correlated with the number and size of the violations. The distinction between properly folded and misfolded structures is usually manifested by a sharp increase in energy (cf. Fig. 2). In this respect the following should be noted: low energy ensembles often contain a few violations above the violation-cutoff for restraints, since they were not selected using this criterion. A structure without violations above the cutoff actually may have a (overall) higher energy when compared to a structure with a few violations above the cutoff, since the latter could have a lower number of total violations. It must also be noted that ensembles with a few violations are not necessarily wrong, as the violations could result from complicated dynamical processes. Fig. 2 illustrates the effect of the different selection procedures. Sorting the ensemble of the PDZ2-AS protein on the basis of NOE energies yields a different ordering to that which would have been obtained using NOE violations as a criterion. However, it is also clear that the first 22 structures are well-folded and would have been selected irrespective of the criterion used. Conversely, the sharp rise in all parameters observed for the last five structures indicates that these represent misfolded structures which are incompatible with the experimental data.

3. Validation of experimental data

In this section we will discuss various methods that can be used to describe the measure of fit of a structure model to the experimental data, their accompanying quality indicators and the information content of NMR derived restraints. We will focus on the validation of distance restraints, dihedral angle restraints and orientational restraints derived from residual dipolar couplings.

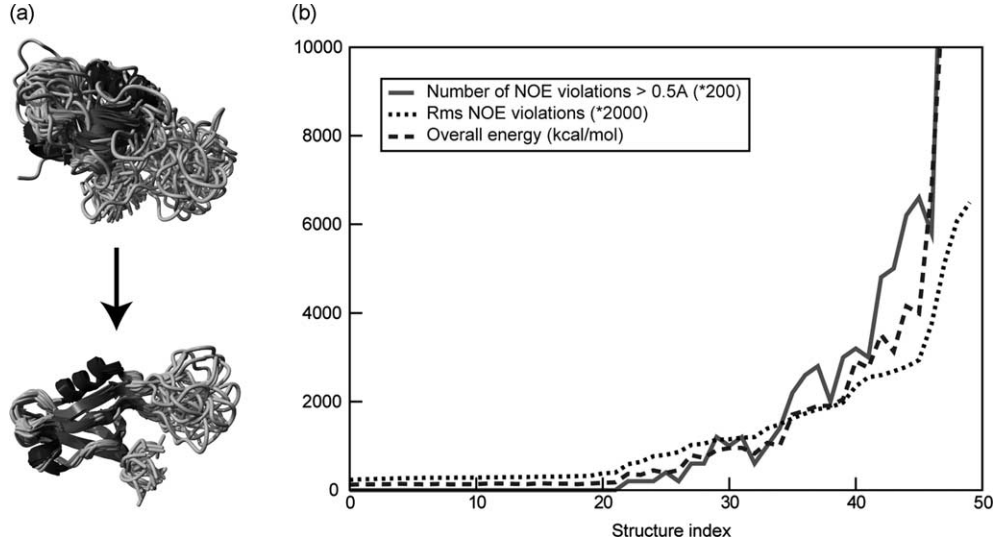

Fig. 2. Structure calculation and selection, illustrated using an alternatively spliced form of the second PDZ domain of PTP-BL (PDZ2-AS) [96] (PDB entry 1OZI). (a) Fifty structures were calculated using a simulated annealing protocol in XPLOR-NIH [97] based on distance and dihedral angle restraints (top). Twenty-two structures were selected that contained no distance restraint violations larger than 0.5 Å and no dihedral angle restraint violations larger than 10° (bottom). Figure generated with YASARA (http://www.yasara.org). (b) Number of NOE violations (solid line), rms NOE violations (dotted line) and overall energies (dashed line), for the 50 calculated structures in (a). Structures were sorted based on NOE energies.

Table 1
Data set information and restraint validation for the PDZ2-AS data set [96]

	Distance restraints	Torsion angle restraints	Dipolar coupling restraints
Data set information			
Number of restraints	1413 (39/22/10/29%)[a]	76	65
Restraints per residue	14.4	0.8	0.7
NOE completeness (4 Å shell)	50%	–	–
QUEEN set information	99.3% (0.1/0.5/12.7/86.7%)[a]	0.7%	–
Non-redundant distance restraints	565 (0.1/0.5/7.7/91.7%)[a]	–	–
Structure calculations[b]			
NOE cross-validation[b1]			
Number of restraint violations[c]	7.6±19.4	–	–
Cross-validated average number of violations[c]	14.1±4.5	–	–
RMS violations	0.03±0.04 Å	–	–
Cross-validated RMS violations	0.31±0.09 Å	–	–
Independent RDC *Q*-factor	–	–	73±8%
RDC independent validation[b2]			
Number of restraint violations[c]	0.2±0.8	0.1±0.2	–
RMS violations	0.008±0.003 Å	0.1±0.1°	–
Independent RDC *Q*-factor	–	–	70±9%
Full data set[b3]			
Number of restraint violations[c]	2±1	2±2	–
RMS violations	0.020±0.002 Å	0.6±0.2°	–
RDC *Q*-factor	–	–	1±1%

[a] The numbers in brackets relate to the distribution (expressed in percentages) over the intra-residual, sequential, medium-range and long-range restraint classes, respectively.

[b] Three structure calculations were performed: [b1]complete NOE cross-validation using only distance restraints. Structures were calculated with working sets of 90%, and test sets of 10% of the distance restraints; [b2]independent validation of RDC restraints using all distance and dihedral angle restraints, excluding RDC restraints; [b3]structure calculation using all restraints. All structure calculations yielded 50 structures with no distance restraint violations >0.5 Å (b1–b3) and dihedral angle restraint violations >10° (b2–b3).

[c] A violation threshold of 0.2 Å (b1–b3) and 2° (b2–b3) was used for the violation analysis.

3.1. Fit of structures to experimental restraints

3.1.1. Restraint violations

Various measures exist to describe the accordance of structure models with the experimental data (cf. Table 1). The most widely used method is the number and magnitude of restraint violations. The interpretation of the number of violations of a structure ensemble is somewhat problematic, since the selection of structures is either directly based on the number of violations, or indirectly by using energy criteria. In the process of structure calculation and refinement, violating restraints may have been removed at some point. Restraint removal can be based on rational grounds by re-interpretation of the experimental data, which does not necessarily pose a problem. However, it cannot be ruled out that, on occasion, violating restraints were removed from a restraint list simply because an ensemble free of restraint violations was deemed desirable. Further, the existence of multiple conformations may result in restraints that are not compatible with a single structure model, and it depends on the interpretation of the researcher whether mutually inconsistent restraints are kept, or whether a choice is made to remove some restraints.

Adding to the fuzziness of using the number violations as a measure of fit is the cutoff value used for determining these violations. For example, if the distance restraint selection criterion prevents violations larger than 0.5 Å, there still may be a significant number of violations in the 0.3–0.4 Å range. In that case, it may appear that the structure models are free of violations, whereas in reality there could be a severe problem with the restraints, certainly when there are many consistent, medium sized violations (see below). Therefore, it is very informative to also investigate the occurrences of restraint violations in lower cutoff classes, e.g. for 0.1, 0.3 and 0.5 Å, in order to get an indication of the quality of the fit of the experimental data to the structural model.

Consistent restraint violations, i.e. restraint violations that occur in more than ~75% of the structures of an ensemble, typically indicate serious problems in the dataset [29]. Analyzing these restraint violations is crucial in the re-interpretation of the data and to resolve problems in the structures.

Finally, we note that deposited structures that contain many consistent violations should be used with great caution, and without the original data available for re-interpretation, these structures are deemed unreliable. Further, we have observed that in many instances the deposited structure ensembles and accompanying restraints exhibit different numbers of restraint violations than actually reported in the corresponding literature and in the deposition information (data not shown). Even more dramatic is the finding that for some cases it appears that the deposited restraints are not those that were actually used to calculate the deposited structures.

3.1.2. RMS deviations and energies of restraints

An often-reported measure of fit of the structure ensemble to the input data is the root mean square (RMS) deviation of inter-atomic distances from their experimentally derived upper and lower bounds. The calculation of the RMS depends on the form of the restraining function, i.e. in case of a bi-harmonic potential all deviations from the target distance are included, whereas in a square-well potential with upper and lower bounds only those deviations which are outside these bounds are taken into account [31]

$$\mathrm{RMS}_{\mathrm{NOE}} = \sqrt{\frac{1}{N_{\mathrm{r}}N_{\mathrm{m}}}\sum_{k=1}^{N_{\mathrm{r}}}\sum_{l=1}^{N_{\mathrm{m}}}(\Delta_{kl})^2}$$

$$\text{with} \quad \begin{cases} d_{kl} > r_k^{\mathrm{upper}} & \Delta_{kl} = (d_{kl} - r_k^{\mathrm{upper}}) \\ r_k^{\mathrm{lower}} < d_{kl} \le r_k^{\mathrm{upper}} & \Delta_{kl} = 0 \\ d_{kl} < r_k^{\mathrm{lower}} & \Delta_{kl} = (r_k^{\mathrm{lower}} - d_{kl}) \end{cases} \tag{1}$$

with d_{kl} the actual distance for restraint k in model l, r_k^{upper} the upper bound of restraint k and r_k^{lower} the lower bound of restraint k. The sum is calculated over all N_{r} distance restraints and N_{m} structural models. An identical expression can be used to calculate the RMS for torsion angle restraints deviations (accounting for the circular nature of this property), with d_{kl} the actual torsion angle measured in model l, and r_k^{upper} and r_k^{lower}, the upper and lower limit of the torsion angle range allowed for by restraint k, respectively.

In general, the lower the RMS deviation, the better the fit of the structure to the experimental data. However, reported RMS deviation values are not easily interpretable, i.e. they are only useful as a relative reference within a dataset and choice of the restraining function.

A third measure of fit reports the overall energies of structures, which contain large contributions of the experimental restraints. Again, the lower the restraint energy, the better the fit of the structure to the data. However, restraint energies depend on the potential energy function used, the size of the dataset, and the force field, and therefore the values are again only meaningful in a relative fashion.

3.1.3. NMR R-factors and cross-validation

One of the most important quality indicators and measure of fit in X-ray structure determination is the so-called *R*-factor, or residual. The *R*-factor measures the difference between the experimentally observed data and those back-calculated from the structure model. Thus, the lower the *R*-factor, the better the model represents the experimental data. Although the *R*-factor is generally used in X-ray crystallography as an indicator of the quality and accuracy of the atomic model, it was shown that diffraction data can be 'overfit' or 'misfit' such that incorrect models can be refined to reasonably good *R*-values [32]. Therefore, it was proposed to use the 'free *R*-factor' or cross-validated *R*-factor. Calculation of the free *R*-factor requires splitting of the diffraction data into two datasets, a 'working' and a 'test' dataset. The working dataset, which comprises ~90% randomly chosen reflections of the full data set, is used to calculate the structure model. The test dataset consists of the remaining ~10% of the reflections and is used to validate the structure model. The *R*-factor calculated for the test dataset using the structure model built on basis of the working dataset, is called the free *R*-factor. A good model should allow for a good prediction of the test dataset. Thus, the closer the free *R*-factor is to the *R*-factor, the better the model. The free *R*-factor has the advantage over the *R*-factor that it can detect overfitting of the data: overfitting of the working dataset will decrease the *R*-factor, but increase the free *R*-factor.

In NMR, NOE-intensities have been used for constructing similar measures [33–35]. An NMR *R*-factor definition, analogous to that used in X-ray crystallography, is

$$R = \frac{\sum_{i,j} W_{ij}(\tau_{\mathrm{mix}})\left|A_{ij}^{\mathrm{calc}}(\tau_{\mathrm{mix}})^p - A_{ij}^{\mathrm{exp}}(\tau_{\mathrm{mix}})^p\right|}{\sum_{i,j} W_{ij}(\tau_{\mathrm{mix}})A_{ij}^{\mathrm{exp}}(\tau_{\mathrm{mix}})^p} \tag{2}$$

with A_{ij}^{calc} and A_{ij}^{exp} the calculated and experimental cross-peak intensities resulting from the interacting spins i and j, respectively, $W_{ij}(\tau_{\mathrm{mix}})$ a mixing time dependent weighting factor, and the power p can simply be 1 or, for example, 1/6 to reflect the distance dependent behavior of the NOE. The weighting factor should account for the uncertainties arising from both the experimental and computational procedures. A major difference between the X-ray and NMR *R*-factors is that each reflection in an X-ray experiment contains information relating to the complete structure of the molecule. In contrast, the information contained in an NOE is much more specific, providing only local information involving two atoms.

NOE *R*-factors are seldom used in NMR structure calculation because accurate NOE intensities can be difficult to obtain due to practical problems such as dynamics, overlap and baseline problems. More accurate fits to the NOE intensities could in principle be obtained from direct refinement against NOE intensities, but this procedure is somewhat problematic from a theoretical and computational standpoint [36]. As an alternative, complete cross-validation has been proposed in order to calculate a free *R*-factor for NMR structures from the set of distance restraints that are derived from the NOE intensities [37]. Because distance restraints contain only information on the relative position of a few atoms and not on the overall structure of the molecule, a single test set is not sufficient to calculate a reliable free *R*-factor. Therefore, a number of randomly chosen test-sets of equal size are chosen to calculate an average free *R*-factor.

Recently, Clore and Garrett have developed an *R*-factor using residual dipolar couplings [38]

$$R_{\mathrm{dip}} = \left[\frac{5\langle(D_{\mathrm{obs}} - D_{\mathrm{calc}})^2\rangle}{2(D_{\mathrm{a}}^{\mathrm{AB}})^2(4+3\eta^2)}\right]^{1/2} \quad (3)$$

in which D_{obs} and D_{calc} are the observed and calculated residual dipolar couplings, $D_{\mathrm{a}}^{\mathrm{AB}}$ and η are the axial component and the rhombicity of the dipolar coupling tensor, respectively. This dipolar coupling *R*-factor is easily calculated, simple to interpret, and can be used for complete cross-validation. As such, the residual dipolar coupling *R*-factor provides a more practical way of assessing the fit of a structure model to experimental data.

3.1.4. Independent validation and Q-factors

In independent validation of structures a specific type or subset of experimental data is completely omitted from the structure calculation process, and only used to validate the result. X-ray and NMR structures of the protein ubiquitin were validated in this fashion using anisotropic carbonyl chemical shifts [14]. Changes in the chemical shifts of carbonyl carbons were measured in a liquid crystalline phase and back-calculated from structure models that were determined without the inclusion of the chemical shifts as structural restraints. Structures calculated with residual dipolar couplings as orientational restraints showed an improved fit to the carbonyl chemical shift data as compared to structures calculated without residual dipolar couplings. The agreement between the structures and the observed property, in this case the changes in carbonyl chemical shift, was expressed using the quality or *Q*-factor [14], defined as

$$Q = \mathrm{RMS}(\Delta\delta^{\mathrm{meas}} - \Delta\delta^{\mathrm{pred}})/\mathrm{RMS}(\Delta\delta^{\mathrm{meas}}) \quad (4)$$

in which $\Delta\delta^{\mathrm{meas}}$ and $\Delta\delta^{\mathrm{pred}}$ are the measured and predicted changes in the chemical shift, respectively. As for the NMR *R*-factors, the lower the *Q*-factor the better the agreement between model and the observed independent data. A similar *Q*-factor has been used as a measure of fit for residual dipolar couplings [39,40], and is equivalent to $\sqrt{2}$ times the residual dipolar coupling *R*-factor [38] (cf. Eq. (3)).

As a second example of independent validation, we mention the use of the residual dipolar coupling *Q*-factors to corroborate improvements in refinement schemes and force fields [27,41]. Structures of the protein ubiquitin were calculated and refined using NOE based distance restraints and torsion angle restraints, and validated independently against six different types of residual dipolar couplings. Structures refined in explicit solvent (see Section 6) showed significant decreases in the residual dipolar coupling *Q*-factors as compared to those that were not refined in solvent.

3.2. Information content in experimental restraints

3.2.1. Number of restraints, completeness and redundancy

An often-used quality measure for the structure calculation input data is simply the total number of restraints, following the simple philosophy ‘the more restraints, the better the structure’. A more sophisticated indicator is the so-called completeness of the data set defined as the ratio between the number of observed and the number expected restraints [42].

For restraints derived from chemical shifts and through–bond interactions, such as *J*-couplings and residual dipolar couplings, the total number is a priori known from the primary structure of the molecule. Hence, the total number of these measured restraints is indicative of the quality of the resulting structure. Exceptions to this are hydrogen-bond restraints derived from ^{3h}J and ^{2h}J-coupling data, for which the expected number cannot be known from the primary sequence alone.

In case of restraints derived from through–space interactions (e.g. NOE restraints) the situation is more complicated. The total number of observable through–space interactions depends on the size and the shape of the molecule, and is therefore not very meaningful as a parameter to compare different molecules and datasets. The size and shape dependence is, in theory, absent in the completeness for NOE restraints. However, it is not known a priori how many NOEs it is possible to measure, and therefore this number can only be derived from the calculated structure ensemble. Therefore, the completeness depends on the calculated structure, and the calculated structure depends on the completeness. Only a very weak correlation between NOE completeness and Ramachandran plot quality, which is generally considered as an independent measure of quality (*vide infra*) has been observed [43,44]. This lack of correlation underscores the problematic nature of completeness as a measure of quality [42].

A second complicating factor is the presence of internal dynamics in biomolecules. Based on a simple static structure model certain NOEs may be expected, but they might not be detectable due to dynamics. Further, overlap can be a severe problem, resulting in the presence of unresolved, unassigned or ambiguous NOEs. Both of these complicating factors are a much greater problem for restraints derived from through–space interaction than for those derived from through–bond interactions.

Occasionally, the information contained in a dataset is described using the number of restraints per residue. This measure corrects to some extent for the size of the molecule and is typically more indicative of structure quality than the total number of restraints. Plots of the number of restraints per residue as a function of the sequence can be quite useful for identifying regions in a structure that are problematic, dynamic or under-determined. However, this measure also has to be interpreted with caution, as it strongly depends on the residue type: glycines will typically display much fewer

NOEs than the larger leucines, and charged, solvent exposed and dynamic residues will on average show fewer NOEs than buried, hydrophobic residues of similar size [42].

Finally, it must be noted that some restraints do not add any information to the structure and thus are in fact redundant. Redundant restraints are normally referred to as those restraints that are always fulfilled because of the geometric restrictions imposed by the molecular topology. For example, an intra-residual distance restraint from H^{α} to H^{N} of 0–3.5 Å in a protein is always fulfilled because the covalent structure already restricts this distance to a maximum of ~3.0 Å. Often datasets contain significant numbers of redundant restraints, which renders the total number of restraints not very meaningful as a quality indicator. Further, double occurrences of restraints are commonly found in deposited datasets. Statistics on the number of restraints, completeness and redundancy can easily be calculated using the program AQUA [45], although analysis of ambiguous restraints has not been implemented.

3.2.2. Quantitative evaluation of experimental NMR restraints

We recently developed a method named quantitative evaluation of experimental NMR restraints (QUEEN), to evaluate the amount of information contained in distance and angular restraints [46]. The method calculates a measure for the information contained in a single distance restraint, given all other restraints, and thus provides a way of quantifying the contribution of each restraint relative to the total dataset.

QUEEN can be used for a number of applications. First, it allows the identification of the most important distance restraints, i.e. those restraints that contain the most information on the molecular structure. Traditionally, a classification of restraints in long-range, medium-range, sequential and intra-residual restraints is used to provide a qualitative measure of the information content in NMR derived distance restraints. In general, long-range interactions have a large impact on the determination of the fold of the structure, and thus are considered to contain more information. To illustrate this concept, Fig. 3 shows the comparison of the information content calculated using QUEEN and the number of NMR distance restraints in the PDZ2-AS data set for each of the aforementioned restraint classes.

It is clear from this comparison that the number of restraints in the different classes is not a good measure for the relative information content. Results of structure calculations where intra-residual restraints were omitted yielded structures in which only 3% of these intra-residual restraints were violated (data not shown). This result confirms that intra-residual restraints, as expected, contain only little information on the fold of the structure, although they typically make up a major part of the dataset. As such, observable short-range interactions are probably more useful for confirming assignments and adding to the completeness of datasets, rather than as an important source of structural information.

Fig. 4 shows an example of the QUEEN analysis for the YBOX protein YB1 [47] in which the ten restraints that contain the most information are indicated. Clearly, these ten most important restraints are all long-range restraints, which largely define the β-barrel structure of YB1.

A second feature of QUEEN is its ability to identify the ‘uniqueness’ of a restraint, i.e. the degree of support of a restraint by other restraints in the dataset. This feature can be used in error analysis: very unique, and thus less supported, restraints may indicate false assignment(s) in the dataset. As an example, Fig. 4 shows the identification of a unique restraint between residue A26 and the aromatic ring of residue Y45 in 1GB1 [48]. This restraint positioned the aromatic ring perpendicular to the orientation found in

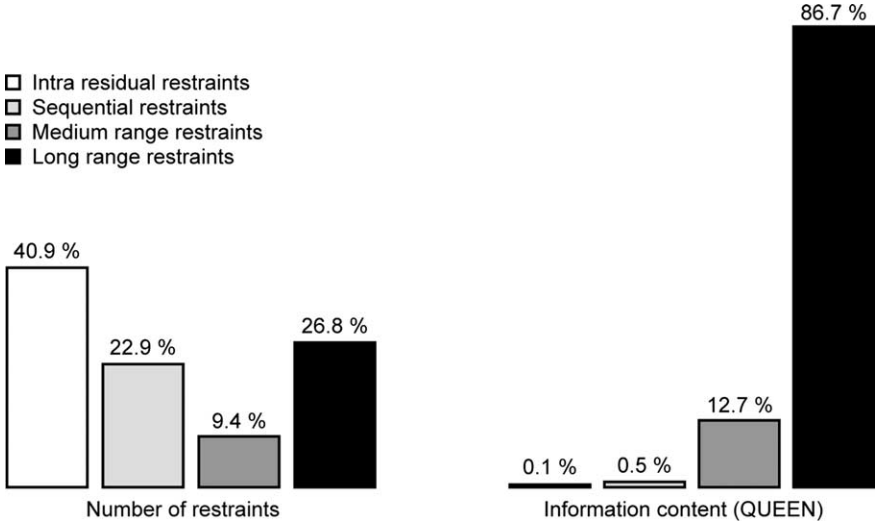

Fig. 3. Comparison of the number of restraints (left panel) to the information content determined by QUEEN (right panel) in long-range, medium-range, sequential and intra-residual restraint classes of the PDZ2-AS (cf. Table 1) data set.

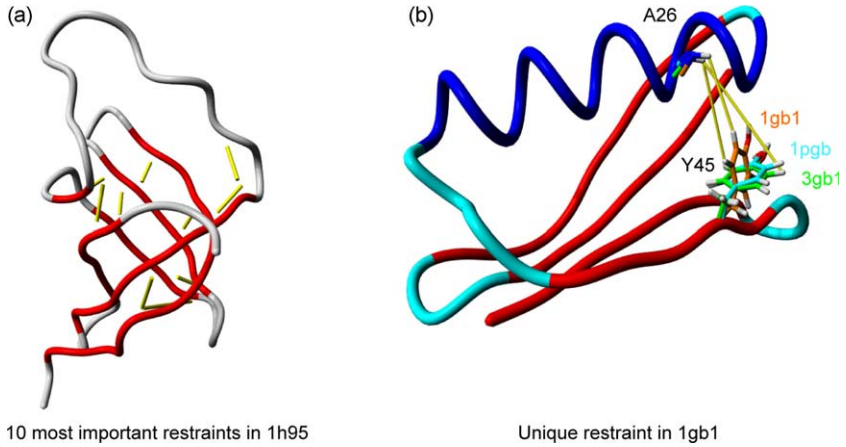

Fig. 4. Illustration of important and unique restraints in NMR structures (restraints are depicted in yellow). (a) The 10 most important restraints in the human Y-Box protein YB1 (PDB entry 1H95). (b) The immunoglobulin binding domain of protein G. Side chain conformations of A26 and Y45 are depicted for PDB entry 1GB1 (NMR, orange), 3GB1 (NMR, green) and 1PGB (X-ray, cyan). The indicated NOE distance restraint involves the H^N proton in A26 and $H^{\varepsilon 1}$ in 1GB1; the H^N proton in A26 and both $H^{\varepsilon 1}$ and $H^{\varepsilon 2}$?in 3GB1. Figure generated with YASARA.

the X-ray structure 1PGB [49], because it involved only one of the two normally indistinguishable H^{ε}-protons of the tyrosine ring. In a later NMR structure of this protein (3GB1) [50], this restraint was changed to involve both H^{ε} protons of the tyrosine ring. The modified restraint in 3GB1 is much more supported by the remainder of the dataset, and resulted in a conformation of Y45 similar to that observed for the X-ray structure.

The third feature of QUEEN worth mentioning is its ability to identify redundant restraints. The QUEEN redundancy check is based on structural redundancy and therefore is more generally applicable than the classical redundancy analysis that only uses the covalent information to identify redundant restraints. Any restraint that, in the presence of all other restraints, does not add information is a redundant restraint, irrespective of its traditional classification as intra-residual, sequential, medium- or long-range. Table 1 lists the number of non-redundant restraints for the PDZ2-AS dataset. Compared to the total number of restraints, the non-redundant encompasses only 565 restraints (40%), with a large overrepresentation of the long-range category. However, we do not advocate the removal of the redundant restraints from datasets, as they contain valuable confirmative information, allowing for automated validation procedures such as QUEEN. Removal of redundant restraints would, however, be useful for reporting more sensible restraint counts.

4. Precision and accuracy of NMR structure ensembles

An important aspect of any structure determination is the final positional uncertainty in the molecular coordinates. In X-ray crystallography the measure used to describe this uncertainty is the *B*-factor or temperature factor, which depends on aspects such as the quality of the crystal, internal dynamics, and disorder.

In most NMR structure studies, the uncertainty in the molecular coordinates is described using the coordinate RMS deviation, or RMSD. Often this parameter is used as a measure of quality of the structure determination, i.e. it is assumed that lower RMSD values indicate more precise and thus better results. However, the more relevant question to ask is how accurately the ensemble of structures represents the true molecular conformation(s) and the variance therein. The variance in the derived coordinates should also reflect the internal dynamics of the molecule. In addition, for proper statistics the precision of the structure model(s) should also reflect uncertainties arising from experimental, data-conversion and computational procedures.

4.1. Precision versus accuracy

To understand the meaning of coordinate RMSDs it is important that precision and accuracy are clearly differentiated. The precision of an observation is generally understood as the measure of reproducibility of repeated measurements and can be expressed as the variance of the measured value (x) around its average value ($\langle x \rangle$). Accuracy, on the other hand is the measure of the closeness of the measurement to the true value [51,52]. The difference between precision and accuracy is illustrated in Fig. 5 for an experiment designed to determine the x- and y-coordinates of a given atom. Assuming that the true coordinates are known for this atom and have a fixed value, these x,y-values are represented by the filled circles in the center of the coordinate frames in Fig. 5a–d. The data points obtained from the different measurements are expected to spread and their values are represented by the open circles. Four different situations can now be distinguished that illustrate the difference between the accuracy and the precision of the measurement:

1. Fig. 5a depicts a *precise*, but *inaccurate* measurement: the narrow clustering of the open circles indicates that

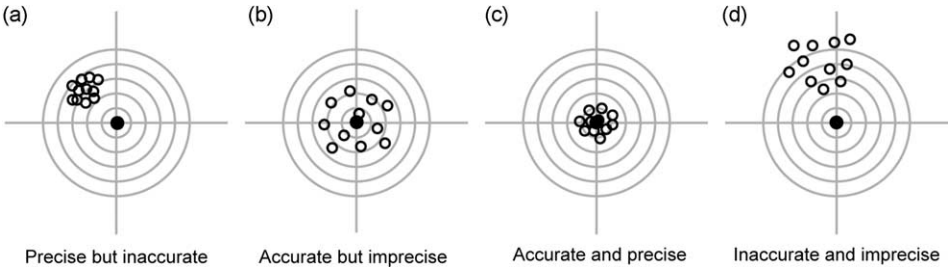

Fig. 5. Illustration of precision and accuracy. The true x,y-coordinate of a given atom is indicated by a filled circle, sampled values of the coordinates are indicated by open circles. (a) Precise but inaccurate sampling; (b) accurate but imprecise sampling; (c) accurate and precise sampling; (d) inaccurate and imprecise sampling.

the measurement is very reproducible and thus precise, however, the values do not cluster around the true values of x and y.

2. Fig. 5b depicts an *accurate*, but *imprecise* measurement: the large variation in the open circles indicates that the measurement is not very reproducible and not precise, however, the values cluster around the true values of x and y.
3. Following similar reasoning the situations depicted in Fig. 5c can be described as an *accurate* and *precise* measurement.
4. Likewise, the situation shown in Fig. 5d represents an *imprecise* and *inaccurate* measurement.

For NMR structures determined in solution it is to be expected that, due to internal dynamics, structures will always display a variation around an average conformation. Assuming this average conformation can be determined, the true spread in the coordinates x, y of an atom may be represented by the filled circles in Fig. 6. Assuming further that the measurement is accurate, i.e. the values obtained from the experiment are spread around the average of the true x- and y-coordinates, three situations can be distinguished:

1. The variance in the coordinates is smaller than the physical variance (the values that are obtained are too precise and precision *underestimates* the physical variance).
2. The variance in the coordinates *equals* the physical variance (precision equals the physical variance, the ideal situation).
3. The variance in the coordinates is larger than the physical variance (the values that are obtained are not precise enough and precision *overestimates* the physical variance).

Practical aspects severely influence estimates of the precision of NMR structures. First, the average conformation cannot be determined without bias because of problems in defining the superposition of structures prior to averaging (*vide infra*). Second, the desired outcome from a structure determination is the situation where the precision equals the physical variance (corresponding to an accurate representation). However, in theory this can only be achieved when the experiments, conversion of experimental data into structural restraints, and computational procedures are perfect and do not introduce additional errors. Since all steps in NMR structure determination involve experimental errors, assumptions and simplifications, the ideal situation can never be achieved in practice. The cumulative effect of these errors in the different steps is expected to result in the situation where the precision overestimates the true variance, as shown in Fig. 6.

In the NMR community there is a trend to minimize the RMSD of an ensemble in order to suggest a precise ensemble of structures, and it has often been shown that precision overestimates accuracy [36,37,52–55]. Consequently, this potentially results in the undesirable situation where the precision underestimates the true variance and the ensemble does not contain any information on some of the truly accessible conformations. For this reason we recently proposed that the RMSD should actually be maximized, only limited by the accordance with the experimental data and geometry restrictions [56]. We also showed that structure ensembles can be improved significantly in terms of almost all validation criteria, while simultaneously increasing the positional RMSD values by a factor 2–3. Although this neither guarantees that the physically available conformational space is sampled completely nor that the so-called ‘resampled’ ensemble represents an accurate representation, the procedure does increase the chances that information on all conformations consistent with the experimental data is present within the ensemble.

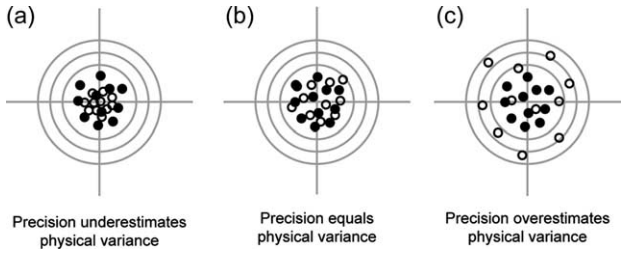

Fig. 6. Illustration of observed precision and true physical variance. It is assumed that the values obtained from the experiment (open circles) are spread around the average of the true (filled circles) x- and y-coordinates.

Maximizing the RMSD values also provides for a more reliable error estimate, which is essential in order to have a good understanding of the available data. This error estimate should include all factors that contribute to the error: experimental errors, dynamics, conversion of data and computational procedures.

In the discussion above, it was assumed that the true conformations are known and that average conformations can be determined. The true value of the parameter under investigation, or the so-called 'gold standard', is required to assess the accuracy of experimental results. Unfortunately, under practical circumstances such a gold standard does not exist and accuracy can only be assessed by using simulated datasets. Although certainly not a gold-standard, it often can be useful to compare the NMR derived results with a relevant X-ray structure, provided that overall indicators suggest such a comparison to be sensible. Bearing obvious differences resulting from mobility, different conditions or crystal contacts in mind, such a comparison may point to regions that warrant careful examination.

Assessing the precision or RMSD of an ensemble of structures is, like the accuracy, not trivial because many different factors influence the final result. First of all, the calculation of meaningful positional RMSD values requires superposition of the ensemble of structures. Superposition requires the definition of regions and atoms that should be superimposed, and many different procedures are employed. A measure for the uncertainty in the coordinates that does not require superposition is circular variance [57], which is calculated from the variation in backbone and side chain dihedral angles. Interestingly, high coordinate RMSD values can coincide with low circular variance. An example is given in Fig. 7 for the hinge-like behavior often observed for loop regions. The figure shows both the positional RMSD and circular variance values for the loop regions (residues 32–46) of the PDZ2-AS protein. Whereas RMSD values are all larger for this region when compared to the remainder of the protein, circular variance data, in particular for the ϕ–ψ combination, indicate the presence of two hinges (residues 36–37 and 43–44, respectively) connected by a relatively structured loop. Unfortunately, circular variance cannot be translated into an uncertainty in Cartesian coordinates.

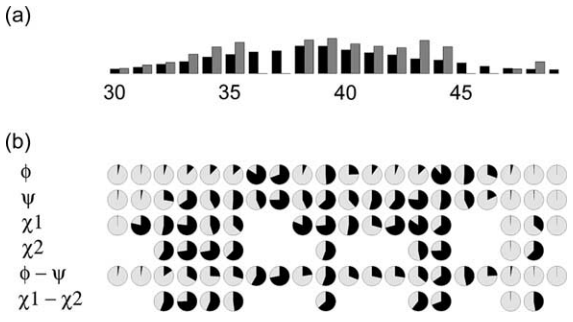

Fig. 7. Plots of (a) the backbone (black) and side chain (gray) RMSD and (b) circular variance for the loop region residues 30–48 in the 22 PDZ2-AS structures from Fig. 2. Figure generated with PROCHECK_NMR.

The second major problem in calculating positional RMSD values relates to the criteria used for selecting the final ensemble. As discussed before, there exists no unbiased procedure and different RMSDs will be obtained for ensembles selected based on different criteria. Finally, the uncertainty in the coordinates will depend to some extent on the force field used to calculate and refine the ensemble of structures. Force fields can differ significantly in the variation that is allowed for bond lengths, bond angles and, more importantly, the backbone ω-angles. The allowed variation in the ω-angles especially can have a significant influence on the RMSD of the backbone [56]. In conclusion, it is clear that the positional RMSD values should be interpreted with great caution when using these as measures for precision or accuracy.

5. Validation of geometric quality

Checking the geometric quality of structure models is an important aspect of NMR structure validation. In high-resolution X-ray structures the geometric quality is determined largely by the data and does not rely heavily on the use of force fields for optimization of the model. In contrast, structure determination by NMR spectroscopy is much more dependent on the use of force fields and refinement procedures, since the information contained in the NMR data alone is not enough to determine a molecular structure. It is therefore particularly important for NMR structures to be checked carefully for abnormalities or errors in the geometry. In this section the various geometric properties that are used to validate protein structure models will be discussed, along with the procedures and measures of quality implemented in the structure validation software packages WHAT IF [58], PROCHECK [57] and PROCHECK_NMR [45]. As an example, the results of structure validation performed on the PDZ2-AS data set are listed in Table 2.

5.1. Z-scores and RMS Z-scores

Most of the checks that are carried out make use of reference values in order to judge whether a property of the structure is to be considered good or bad, normal or abnormal. The program WHAT IF uses so-called *Z*-scores and RMS *Z*-scores as quality indicators for the various structure properties (see also [27,44] and http://www.cmbi.ru.nl/whatif/checkhelp). A *Z*-score relates the value x of a parameter to a normalized Gaussian distribution derived from a database, i.e.

$$Z = \frac{x - \langle x_{db} \rangle}{\sigma(x_{db})} \tag{5}$$

Table 2
Structure quality indicators for the PDZ2-AS data set

Structure Z-scores	
First generation packing quality	-1.6 ± 0.3
Second generation packing quality	-1.3 ± 0.4
Ramachandran plot appearance	-3.3 ± 0.5 (79.4%/16.8%/2.3%/1.5%)[a]
$\chi1/\chi2$ rotamer normality	-1.5 ± 0.5
Backbone conformation	-4.5 ± 0.7
RMS Z-scores	
Bond lengths	0.90 ± 0.04
Bond angles	0.84 ± 0.03
Omega angle restraints	0.83 ± 0.07
Side chain planarity	0.9 ± 0.2
Improper dihedral distribution	0.89 ± 0.05
Inside/outside distribution	1.01 ± 0.02
Number of bumps per 100 residues	5 ± 2
Unsatisfied buried hydrogen donors	11 ± 3
Unsatisfied buried hydrogen acceptors	0.1 ± 0.3
RMSD (backbone/heavy atom)[b]	$0.7\pm0.1/1.5\pm0.1$

The analysis was performed on 50 structures calculated using the full PDZ2-AS data set (cf. Table 1, b3), including a refinement in explicit solvent.

[a] The percentages in, respectively, the favored, allowed, generously allowed and disallowed regions of the Ramachandran plot from PROCHECK are given in parenthesis. All residues were included in the PROCHECK analysis.

[b] RMSD values were calculated after superimposing the ensembles on the secondary structure elements, comprising residues 13–19, 27–31, 47–53, 58–61, 68–73, 76–77, 83–93, 96–103 [96].

where $\langle x_{db}\rangle$ and $\sigma(x_{db})$ denote the database-derived average and standard deviation, respectively. *Z*-scores can be interpreted using the basic rules for normal distributions that ~68% of the observations are expected to be found within ±1 standard deviation from the average ($Z=\pm1$), 95% within ±2 standard deviations ($Z=\pm2$), etc. *Z*-scores smaller than -4 or larger than $+4$ are considered outliers, which does not necessarily mean that they are in error, but rather that they are just very unlikely to occur, or in other words, to be correct. It is, however, absolutely normal to find a few outliers. For example, in the case of 20,000 Gaussian distributed observations one expects ~20 outliers, which means that finding 10 outliers is equally bad as finding 30 outliers.

In order to check whether a distribution of values has more (or less) outliers than expected from the reference distribution, it is useful to calculate the RMS of the population of *Z*-scores, or RMS *Z*-score

$$\mathrm{RMSZ}=\sqrt{\frac{\sum_{j=1}^{N}Z_j^2}{N}} \tag{6}$$

where Z_j is the *Z*-score as defined in Eq. (5) for observation j, and N is the total number of observations. In the case where the distribution of a parameter has an identical average and standard deviation as the reference distribution, the RMS *Z*-score is ~1.0. For narrower and wider distributions, the RMS *Z*-scores will be <1.0 and >1.0, respectively (cf. Fig. 8a). It must be stressed here that in order to calculate a meaningful RMS *Z*-score the averages of the reference distribution and the model distribution have to be equal [27]. As will be shown below, RMS *Z*-scores of geometric parameters are very sensitive indicators of problems pertaining to local geometry.

5.2. Bonded geometry

5.2.1. Bond lengths and angles

Bond lengths and angles in proteins and nucleic acids are accurately known from analyses of structures of small molecules taken from the Cambridge structural database (CSD) [59,60]. They are usually referred to as the so-called Engh and Huber parameters and form the reference parameters for the WHAT IF, PROCHECK and PROCHECK_NMR programs. Although constraining bond lengths and angles according to these parameters seems trivial, structures showing significant deviations were found in many cases [41,61,62]. Further, bond lengths and angles are often tightly constrained in NMR structure calculations, resulting in structures having the proper target value, but with too tight distributions [41]. Outliers in bond lengths

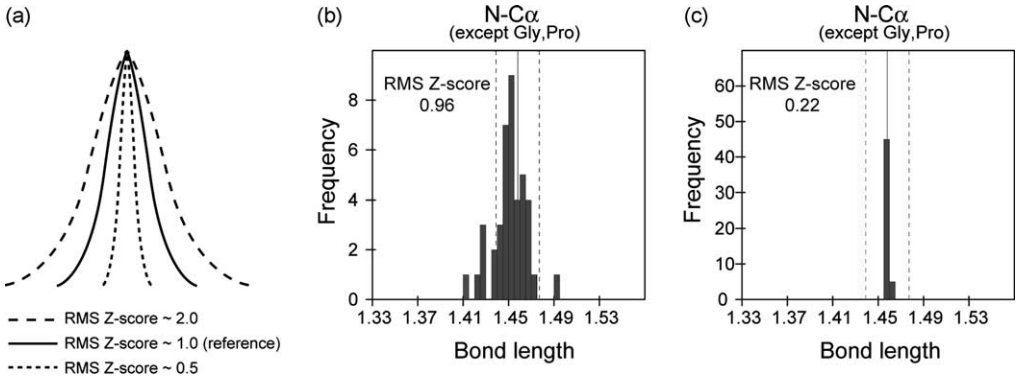

Fig. 8. (a) Illustration of RMS *Z*-scores and corresponding normal distributions. (b) Distribution of N–C^{α} bond lengths in a high resolution X-ray structure (RMS *Z*-score = 0.96). (c) Distribution of N–C^{α} bond lengths in an NMR structure with tight constraints on the bond lengths (RMS *Z*-score = 0.22). Figure generated with PROCHECK, RMS *Z*-scores calculated with WHAT IF. Figures (b) and (c) generated with PROCHECK.

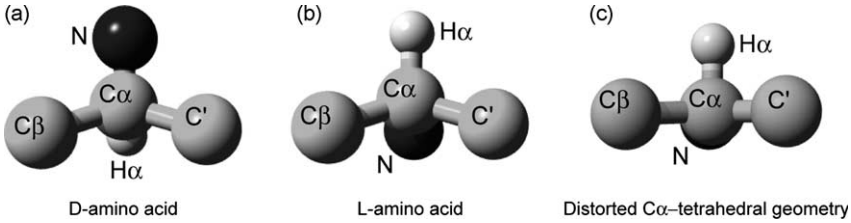

Fig. 9. Illustration of chirality and deviations from tetrahedral geometry in proteins. (a) D-amino acid; (b) L-amino acid; (c) deviation from ideal tetrahedral geometry in a L-amino acid. Figure generated with YASARA.

and angles can sometimes also be caused by non-bonded interactions or strain, possibly due to erroneous NMR restraints.

Fig. 8 illustrates the distributions of N–C^{α} bond lengths for an X-ray structure with an RMS *Z*-score for the bond lengths of ~1.0, and an NMR structure that exhibits a typically low variation in the bond lengths (RMS *Z*-score ~0.2). The low variation in this parameter observed for the NMR-derived structure is caused by the force field and protocol used to calculate and refine the structures that tightly restrain the local geometry. In our opinion force fields should allow for the experimentally observed variations to be reflected in the structure models.

5.2.2. Chirality and tetrahedral geometry

Proteins and nucleic acids contain a number of chiral centers, for example the C^{α} atoms in amino acids (cf. Fig. 9). In the case of NMR structures the correct chirality has to be constrained by the applied force field, since the NMR data do not contain information on the bonded geometry. In the early days of NMR structure determination global inversions of chirality sometimes occurred, and erroneous occurrences of D-amino acids in NMR structures are still occasionally found [61].

Deviation from ideal tetrahedral geometry can occur to some extent [63]; however, large deviations are unlikely and may point to errors in the process of modeling of the structure [61,64]. The geometry of all atoms that have three bonds to non-hydrogen atoms is routinely checked by the programs WHAT IF and PROCHECK. Deviations that are too large, i.e. $|Z|>4$, are typically reported.

5.2.3. Side chain planarity

Planar side chain groups are present in nine of the 20 natural amino acids. Their deviations from planarity have been analyzed in small molecules from the CSD and in high-resolution X-ray structures [62]. The resulting reference parameters describing the planarity of the different side chains are now used to check new protein structures. In several recent studies it was shown that the planarity of the amino acid side chains in NMR structures is often modeled incorrectly [41,65]. As an example, Fig. 10 shows a common problem for arginine side chains where the planarity of the guanidinium group is seriously compromised. This non-planar conformation is clearly an error in the structure calculation/model-building process.

5.2.4. Side chain rotamers

Side chain rotamers have a preference for staggered conformations with dihedral angles around −60, 60 and 180°. This preference can be explained by differences in the energies of the eclipsed (high energy) states and staggered (low energy) states. Both states can occur since transitions are possible through rotations about the intervening bond.

The relative populations of the three side chain rotameric states vary among the amino acids, the different secondary structure elements and their environments. These relative populations are known from analysis of the reference databases of high-resolution X-ray structures. Both PROCHECK and WHAT IF use this knowledge for validation of side-chain rotamers (cf. Fig. 11). In NMR structure calculations the force field terms for the dihedral angles are often switched off, resulting in rather poor rotamer distributions with high numbers of eclipsed states. Including these terms in the structure calculation or final refinement of the structures significantly improves the quality of the rotamer distributions.

Methods have been developed that use database derived side-chain rotamer preferences directly in the structure calculation protocols in order to improve NMR structures [66–70]. It should be clear that applying these database potentials in the structure calculation and refinement renders a validation based on a similar reference database rather useless. As will be discussed below, we want to point out that these methods should be used with caution, since they

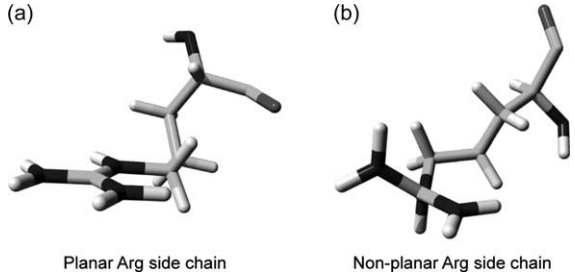

Fig. 10. Deviations from planarity of arginine side chains in NMR structures. (a) A properly constrained arginine side chain; (b) a typical deviation from side chain planarity that is often observed in NMR structures due to lack of geometric constraints. Figure generated with YASARA.

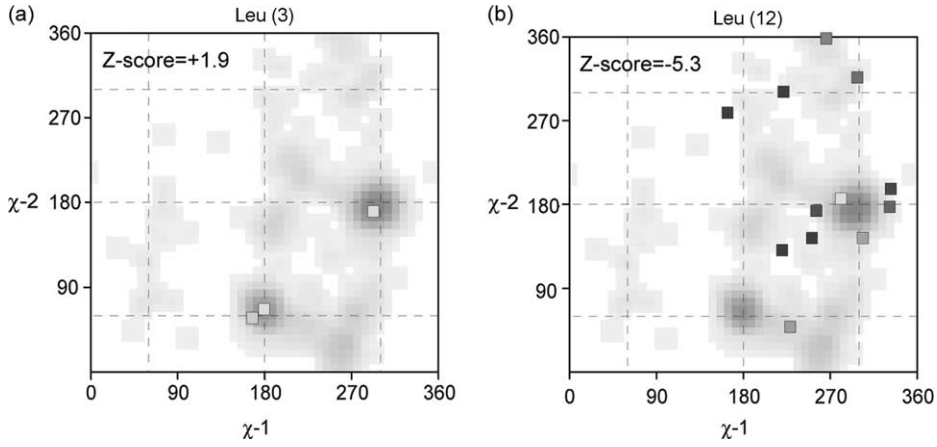

Fig. 11. Illustration of $\chi1-\chi2$ side chain-rotamer distributions and the corresponding *Z*-scores for leucine residues. (a) Favorable leucine side chain rotamers; (b) unfavorable leucines side chain rotamers. Areas that correspond to preferred conformations are shaded gray. Figure generated with PROCHECK, *Z*-scores calculated with WHAT IF.

may mask truly occurring deviations from preferred conformations.

5.2.5. Backbone conformation

The backbone conformation of a protein is determined by the dihedral angles ϕ, ψ and ω (cf. Fig. 12) of each residue, and typically contains a number of regular structure elements. These elements can be classified into the well-known secondary structure elements, such as α-helices, β-sheets and turns. Identification of these secondary structure elements is usually performed by the program DSSP [71], which uses hydrogen bonding patterns as a decisive criterion. Such secondary structure classifications of backbone conformations are very useful in describing and displaying the overall molecular topology. However, for validation of the quality and normality of the backbone conformation more elaborate methods are used.

5.2.5.1. Ramachandran plot. An important indicator for protein structure quality is the combination of the dihedral angles ϕ and ψ. Based on steric considerations Ramachandran et al. [72] showed that the ϕ and ψ combinations of amino acids in a polypeptide chain are restricted to certain ranges, which can be visualized in a so-called Ramachandran plot. Apart from steric restrictions, ϕ and ψ exhibit preferences that depend on residue type and secondary structure elements, resulting in more narrow residue specific ϕ/ψ ranges. The importance of the Ramachandran plot for assessing the quality of a protein structure has been shown by several studies [44]. In X-ray structures a correlation between the clustering of residues in the Ramachandran plot and both the resolution and the free *R*-factor was found [43].

There are different ways of describing the quality of the Ramachandran plot for a structure model. The most widely used is the classification of ϕ/ψ combinations in four regions of the Ramachandran plot: favored, additionally allowed, generously allowed and disallowed, as implemented in the PROCHECK program (cf. Fig. 13a). A variant on this classification scheme is the division into core and non-core regions, where the core regions contain 98% of all non-glycine residues in 400 high-resolution structures and the non-core region the remaining 2% [43].

A drawback of statistics based on divisions in different regions of the Ramachandran plots is that these regions are

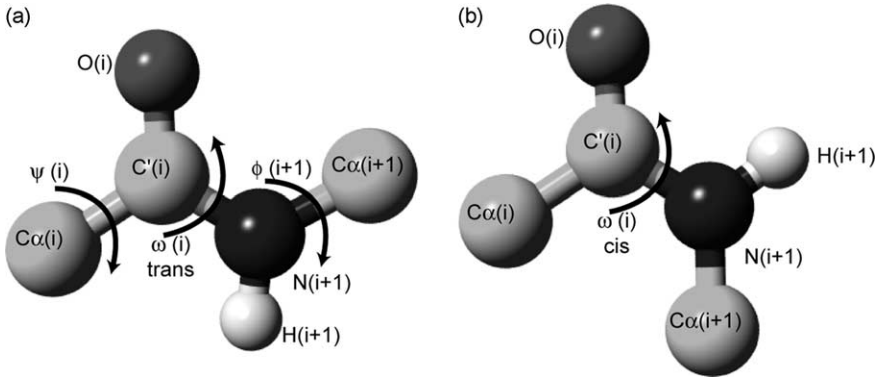

Fig. 12. (a) Atoms and bonds defining the protein backbone dihedral angles in neighboring residues (i) and ($i+1$). (b) Illustration of the protein backbone with the ω angle in the *cis*-conformation. Figure generated with YASARA.

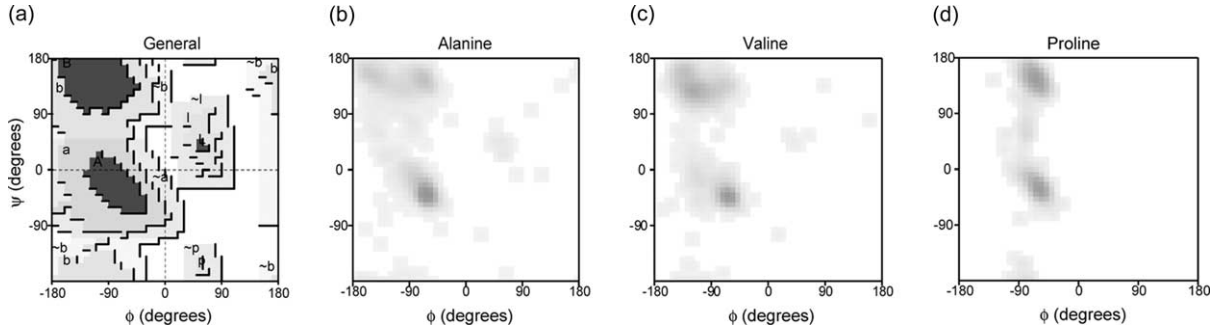

Fig. 13. (a) Illustration of the generalized Ramachandran plot preferences (favored, additionally allowed, generously allowed and disallowed regions are indicated in black, gray, light gray and white, respectively), and residue specific Ramachandran plot preferences (preferred regions are shaded) for (b) alanine, (c) valine and (d) proline. Figure generated with PROCHECK.

generalized for all residues (except glycines and prolines). As stated above, different amino acids exhibit different ϕ/ψ preferences, which is illustrated in Fig. 13b–d which shows the overall, Ala, Val and Pro ϕ/ψ distributions in the PROCHECK database. Considerable differences exist between the most populated regions. Neglecting residue specific preferences can result in classifying a residue as located in the most favored regions, whereas a residue specific analysis would result in a poorer score, warranting further investigation.

A method that does take the residue specific preferences into account was developed by Hooft et al. [44] and is implemented in WHAT IF. *Z*-scores are calculated for each residue, and a composite overall *Z*-score is calculated for the full protein. A positive correlation between the overall *Z*-score and the number of residues in the most favored regions of the Ramachandran plot was observed. Fig. 14 illustrates the correlation between the clustering of residues in the PROCHECK Ramachandran plot and the WHAT IF *Z*-score: the more positive *Z*-scores generally correspond to a tighter clustering of residues in the most favored regions, whereas more negative *Z*-scores generally correspond to a higher occurrence of residues outside the favored regions of the Ramachandran plot. However, a high percentage of residues in the most favored regions does not necessarily imply a high *Z*-score: structures that appear very regular based on the four-region analysis have been identified to be not that regular when using the WHAT IF *Z*-score [44]. In practice, the combination of *Z*-scores and visual inspection of the Ramachandran plot and structure is a good way of assessing the quality of the structure model.

How can we interpret the different Ramachandran plot statistics of a structure model? In the case where the four-region classification or the core/non-core division is used, the most interesting residues are those that occur in the disallowed or non-core regions, i.e. the outliers. It is important to question whether these outliers are truly abnormal conformations or are errors. This issue can only be resolved by careful inspection of the structure and the input data. Typical examples of non-erroneous outliers include active-site residues and D-amino acids. The latter have allowed ϕ/ψ combinations that are mirrored with respect to the L-amino acids, and hence the apparent non-core ϕ/ψ combinations can be falsely reported as errors by the validation software. On the other hand, residues that appear in the favored regions of the generalized Ramachandran plot should not a priori be regarded as good. A more

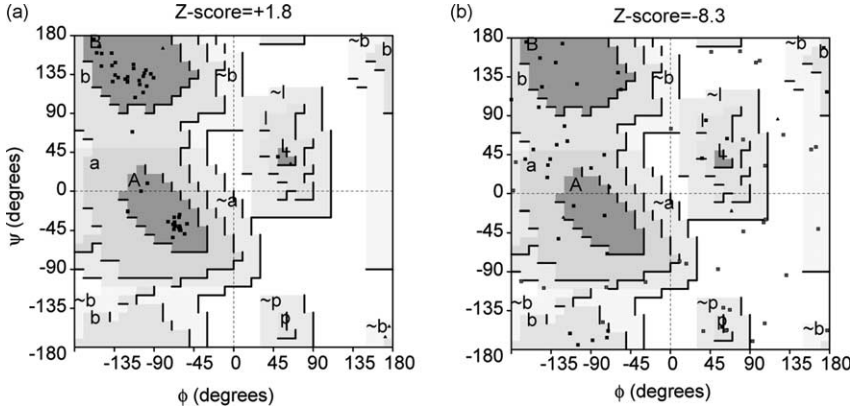

Fig. 14. Illustration of the correlation between WHAT IF Ramachandran *Z*-scores and the four-region classification implemented in PROCHECK. (a) Ramachandran plots of a structure with *Z*-score = +1.8 and (b) of a structure with *Z*-score = −8.3. Figure generated with PROCHECK.

critical inspection of individual *Z*-scores may identify unlikely conformations of these residues.

In order to get an overall indication of the quality of the Ramachandran plot the overall *Z*-score or the equivalent resolution in PROCHECK(_NMR) can be used. The equivalent resolution is based on the observation that the number of residues in the most favored regions increases with increased resolution of X-ray structures. Thus, the higher the equivalent resolution the better the model is assumed to be, although this should be interpreted not as an absolute standard for reasons given above. For *Z*-scores the situation is different. Since *Z*-scores are based on normal reference distributions, this implies that a high *Z*-score cannot be considered better than a low *Z*-score. However, as shown in Fig. 14, high *Z*-scores indicate better agreement with the most likely regions of the database-derived Ramachandran distributions. In general, structures with *Z*-scores between -2 and $+2$ are considered to be within normal ranges and thus good structures. Structures with *Z*-scores $> +2$ are exceptionally close to the ideal values for all residues which may either be a true characteristic caused by the quality and information contained in the data, or by the use of database potentials to improve the Ramachandran plot appearance specifically [66–70] (see below). Structures with *Z*-scores < -2 should be inspected further, because they have increased numbers of outliers suggesting serious problems with the structure.

In recent years methods have been developed to improve the Ramachandran characteristics of NMR derived protein structures. An obvious way is by increasing the amount of experimental data used to calculate the structures. However, substantial improvements are also possible in the modeling and refinement process of NMR structures. It has been shown that clear improvements can be achieved by a final refinement in explicit solvent using physically more realistic force fields than those used in the normal simulated annealing schemes (see below). In another approach database-derived potentials aim to improve the Ramachandran characteristics [66–70]. Although one can argue that the use of such potentials can result in improved models from which unlikely conformations, i.e. the outliers in the Ramachandran plot, are removed, the method should be used with caution. First, it is absolutely essential that the experimental data are not compromised by the use of the database potential, since the experiments are the only source of independent information. Truly abnormal conformations typically play an important role in biological functioning, but may not be consistent with the empirical database derived potentials. In such cases, trying to optimize the Ramachandran plot may be in direct conflict with physical reality. Secondly, using the database potential renders the Ramachandran plot useless as an indicator of the overall quality of the structure model. Effects of low data content, or bad structure calculation and refinement practice, may be masked by a Ramachandran plot that appears excellent and thus indicates a good structure model.

It is important to note that the four-region division can easily be abused to make the quality look acceptable, whereas in reality it is not quite that good [43]. Furthermore, the Ramachandran plot appearance can be artificially improved by selection of only parts of the structure for the analysis. Often, these are referred to as the 'well-ordered' parts or the regular secondary structure elements. This method is only acceptable when the selection criteria are clear and the structure(s) of the non-selected region(s) are essentially unknown due to dynamics or the lack of data and are not used for any analysis.

5.2.5.2. ω torsion angles. The ω angle is the torsion angle about the peptide bond, involving atoms $C^{\alpha}(i)$–$C'(i)$–$N(i+1)$–$C^{\alpha}(i+1)$ (Fig. 12) and is dependent on the residue type [73]. Because of the partial double bond character of the peptide bond its free rotation is impaired, which results in a conformation of the peptide bond that is always close to planarity. The ω angle is mostly found in the *trans*-conformation (180°), but *cis*-conformations (0°) are found as well and are thought to be important determinants of protein function. Most of the *cis*-peptide bonds occur on the N-terminal side of prolines (Xaa-Pro), but cases are known where *cis*-conformations occur for peptide bonds that do not involve prolines (Xaa-nonPro). Based on thermodynamic considerations, it is expected that about 30% of the Xaa-Pro peptide bonds should be in the *cis*-conformation and 1.5% of all the Xaa-nonPro peptide bonds, which is a much higher number than is actually found in X-ray structures deposited at the PDB [74].

Interestingly, it has been noted that there is a correlation between the occurrence of *cis*-peptide bonds and the resolution of X-ray structures, suggesting that *cis*-peptide bonds are often missed in lower resolution X-ray structures [74]. For NMR structures the situation is even worse, since in most NMR force fields ω angle *trans*-conformations are simply imposed on the structure, unless explicitly specified otherwise. Concomitantly, we find that only 0.2 and 0.04% Xaa-Pro and Xaa-nonPro *cis*-peptide bonds, respectively, occur in NMR structures deposited at the PDB. This survey excluded a few extreme outliers, where up to 20 *cis*-peptide bonds per structure were present, which is certainly caused by incorrect treatment of the ω angles.

Another problem of practical importance is the narrow range of the ω angles around their 'ideal' values. From analysis of high resolution X-ray structures of proteins it is known that the standard deviation of this torsion angle is $\sim$5.6°, which means that deviations from planarity of up to 12° are readily found [63,73,75]. However, in NMR structure calculations the ω angle is often tightly restricted to 180°, resulting in ω angle distributions that do not reflect the naturally occurring variation (cf. Fig. 15). This relatively tight restraining of the ω angle not only restricts the variation in the backbone conformation but also influences the ϕ and ψ backbone torsion angles and $\chi1$–$\chi2$ distributions [41,56,61].

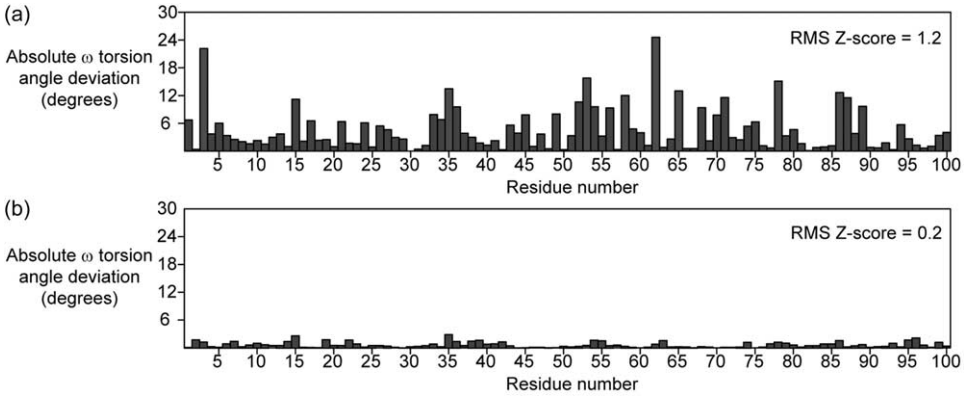

Fig. 15. Variation in ω torsion angles in an X-ray and an NMR structure of lysozyme. (a) High-resolution X-ray structure ($\omega \sim 180 \pm 6°$, RMS *Z*-score = 1.2, PDB code 3LZT). (b) NMR structure with tight constraints on the ω angle ($\omega \sim 180 \pm 1°$, RMS *Z*-score = 0.2, PDB code 1E8L). Figure generated with PROCHECK.

5.2.5.3. Backbone normality. Description of the backbones of proteins using classification in secondary structure elements is in some cases insufficient, owing to large differences in conformation within the elements, e.g. for residues at the beginning, in the middle or at the end of an α-helix. Moreover, loop conformations are known to vary significantly between different loops. In such cases, a context dependent comparison of the backbone conformation using a 3D backbone database becomes very useful. A so-called backbone normality check is implemented in WHAT IF. In this procedure the relative C^{α}-positions of five sequential residues are compared to all C^{α}-positions of five sequential residues in the reference database. The number of matching conformations in the database is taken as a measure for the acceptance of the backbone conformation around the central C^{α} nucleus. This check, which is a true normality check, can be used to identify rare backbone conformations, such as in strange loops or circular peptides (Fig. 16). However, low scores in the backbone normality may also indicate errors in the structures and can therefore be used as a validation tool as well.

5.3. Non-bonded interactions

5.3.1. Inter-atomic bumps

Inter-atomic bumps or bad contacts occur when the distance between the centers of two atoms is less than physically realistic, since the Van der Waals repulsion allows overlap of atoms only to a very limited extent. Occurrences of these bumps are analyzed by simple routines in PROCHECK and WHAT IF. PROCHECK identifies atoms as being involved in a bad contact if they are closer in space than 2.6 Å, only considering the non-hydrogen atoms and those pairs of atoms that typically do not form a hydrogen bond. WHAT IF implements a slightly more sophisticated bump analysis, reporting bumps for any pair of non-hydrogen atoms that are closer in space than the sum of their respective Van der Waals radii, −0.4 Å. WHAT IF uses more relaxed criteria for atoms in potential hydrogen bonds and atoms that are less than four bonds away in the structure.

In practice the WHAT IF check for bumps is more critical than the PROCHECK bad contact analysis and results in much higher counts of overlapping atoms. A comparison between X-ray and NMR structures using the WHAT IF bump checks revealed that overlap between atoms occurs about 10 times more often in NMR structures than in X-ray structures [56]. High occurrences of bad contacts in NMR structures have been noted previously [61], especially for structures calculated in XPLOR in which Van der Waals radii were commonly set to 0.8 times the standard values used in the CHARMM force field [76]. Molecular dynamics refinement of NMR structures in explicit solvent has been shown to be an effective method of removing bumps from structures [27,41].

5.3.2. Hydrogen bonding

Hydrogen bonds play an important role in the folding and stability of proteins. Considering that the energy difference between the folded and the denatured state is 5–15 kcal $\cdot$ mol^{-1} and that the energy of formation of a single hydrogen bond is in the order of 2–5 kcal mol^{-1} [77], the absence or presence of only 1–3 hydrogen bonds can have a profound effect on the stability of a protein structure. This general notion is confirmed by the observation that protein

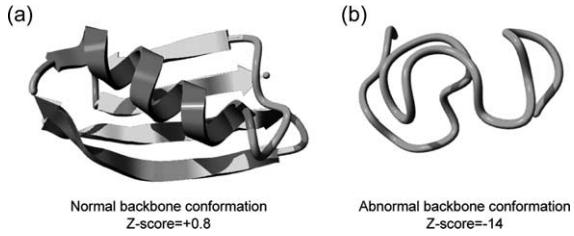

Fig. 16. Illustration of normal and abnormal backbone conformations. (a) Normal backbone conformation (*Z*-score = +0.8, PDB code 3GB1); (b) abnormal backbone conformation (*Z*-score = −14, PDB code 1HZ3). Figure generated with YASARA.

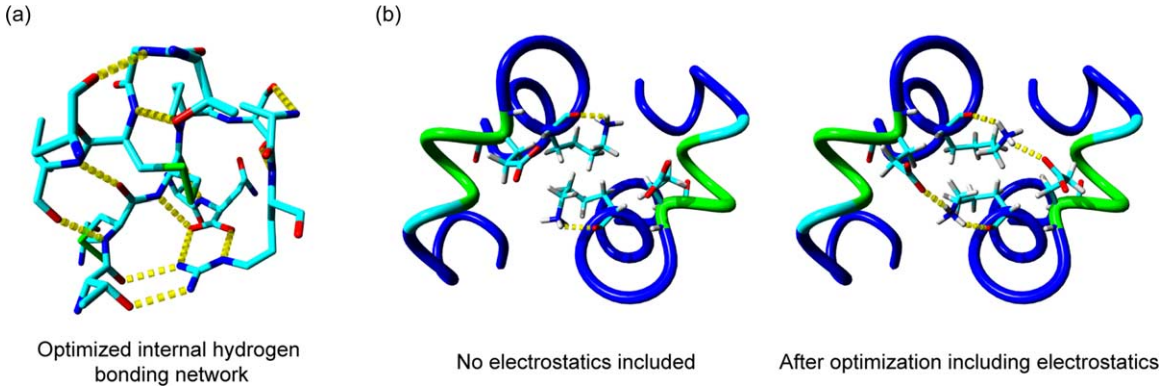

Fig. 17. (a) Illustration of an optimized hydrogen-bonding network in the interior of the protein crambin (PDB code 1CRN). Hydrogen bonds are indicated with dashed lines. (b) Example of the effect of inclusion of electrostatics in NMR structure refinement. The original dimeric structure (left panel) was refined without inclusion of electrostatic force field terms and the dimer interface exhibits non-optimal electrostatic interactions (PDB code 1KA3). Salt-bridges that stabilize the interface are formed after energy minimization including electrostatics (right panel). Figure generated with YASARA.

structures indeed have a very low number of unsatisfied hydrogen bond donors [78] (cf. Fig. 17).

When building a structure model based on X-ray crystallographic data it is not trivial to place the polar hydrogens in an optimal fashion. For example, in structures determined at less than atomic resolution, hydrogen atoms simply cannot be distinguished. Another difficulty in X-ray crystallographic data is that it is normally not possible to distinguish between nitrogen and oxygen atoms in the side chains of glutamines and asparagines, and between carbon and nitrogen atoms in histidine side chains, leading often to conformations of these side chains that are 180° flipped with respect to the most optimal conformation (side chain flips).

In NMR structure determination hydrogen atoms are the primary source of structural restraints and the resonances of many of the hydrogen atoms that are involved in hydrogen bonds can be readily observed. However, identification of hydrogen bond acceptors can only be done unambiguously in a direct measurement of the through–hydrogen bond *J*-coupling [11,12]. Most of the hydrogen bond restraints that are used in NMR structure calculations are derived indirectly from amide proton exchange rates and chemical shifts [79,80], and have to be treated with some caution due to the inherent ambiguity in the assignment of the hydrogen bond acceptor. Moreover, most of the indirect restraints are derived for hydrogen bonds in the backbone of the protein. For side chains potentially involved in hydrogen bond formation, the ambiguity in the possible hydrogen bond acceptor is usually much higher and cannot be reliably translated into structural restraints. Another limitation arises from the inability to observe the hydrogen atoms of the hydroxyl groups in serine, threonine and tyrosine under typical conditions, due to chemical exchange processes. These residues, however, are often important hydrogen bond donors when they reside in the interior of a protein structure.

Thus, both for NMR and X-ray derived structures, generating a model that displays a good hydrogen-bonding network requires proper modeling. In the case of X-ray structures the heavy atom coordinates are refined against the electron density and provide primary geometric constraints for the different positions a proton can occupy. Optimization of the hydrogen bonding network therefore does not require changes in the heavy atom coordinates, although slight modifications in bond lengths and angles may have to be made in the case of side chain flips.

For NMR structures the situation is more complicated, since there is usually much more variation allowed in the heavy atom coordinates. This can be traced back to the much lower information content in the NMR derived restraints, making the modeling of NMR structures much more dependent on computational methods and force fields. Due to the polar character of hydrogen bond donors and acceptors the inclusion of electrostatic forces in the refinement of protein NMR structures, in implicit or explicit solvent, helps to improve the hydrogen bonding network [27,41,81].

Finally, a number of methods have been developed in the past to deal with non-optimal hydrogen bonding networks in protein structures [82,83]. An improved method [84] has been implemented in WHAT IF and was used in a survey of protein structures [62], revealing many ($\sim$15%) asparagine, glutamine and histidine side chain flips.

5.3.3. Electrostatics

Like hydrogen bonding interactions, electrostatic interactions play an important role in protein stability and function. For example, salt bridges stabilize coiled-coil structures [77]. Further, active sites, such as formed by the catalytic triad in serine proteases, ligand binding sites, and metal binding sites are often highly charged (see also [85]). Therefore, calculation of electrostatic properties is an important aspect of protein engineering. It has been shown that optimization of the hydrogen bonding network has a positive effect on pK_a calculations and the results of common electrostatics calculations [86].

Like in the case of the hydrogen bonding interactions, in X-ray structures the optimization is limited to repositioning hydrogens and/or flipping side chains of some polar residues. In NMR structures, the situation is again more complicated. The most important problem is the lack of information accessible by NMR on the conformation of the electrostatically interacting groups. Structural information on the carboxyl group in negatively charged residues, such as Glu and Asp, can only be assessed through limited and indirect information from the hydrogen atoms in the remainder of the side chain or from chemical shifts. The same holds for the positively charged head-groups of Arg and Lys side chains, for which the hydrogen atom resonances are often unobservable or not useful for structural analysis due to chemical and conformational exchange. Therefore, a proper treatment of electrostatic forces in the NMR structure calculation or refinement process is essential and can be achieved by refinement using molecular dynamics simulations in explicit solvent (cf. Fig. 17). This not only removes unrealistic charge distributions or interactions, but in addition can be very useful in predicting intermolecular interactions and understanding biology.

5.3.4. The packing of residues in protein structures

Residue packing is a very important indicator for structural quality. Different types of amino acids display differential preferences for their neighboring residues, and these preferences ultimately determine the global fold of the structure. Using physico-chemical knowledge some general rules can be derived for the preferences of neighboring residues, e.g. hydrogen bond donors like to be around hydrogen bond acceptors, aromatic residues like to be stacked, and hydrophobic residues like to be around other hydrophobic residues (cf. Fig. 18). However, such general rules are not sufficiently useful to judge the packing of residues in a structure model.

A useful empirical method of checking the packing quality is the directional atomic contact analysis, expressed as the DACA score [87]. From an analysis of known protein structures solved at high-resolution, reference distributions were derived for favorable packing environments. Residues are subdivided into geometric fragments without internal degrees of freedom, and atoms classified into different chemical types. The evaluation of the packing preferences now involves counting and generation of normalized distributions of the occurrences of a contact of any given fragment with an atom of a certain chemical type. Contacts between certain atom types and residue fragments will occur more often in the reference database when they are favorable, and hence are given a positive quality score. The X-ray DACA scores were shown to be correlated to the crystallographic *R*-factor in a refinement series: the lower the *R*-factor, i.e. the better the agreement between the structure and the experimental data, the higher the packing quality index. This observation implies that the packing quality index can be used as an independent indicator for the quality of a structure model. Further, it was shown that structures that were known to be incorrect scored significantly lower on the packing quality scores than structures that were assumed to be correct.

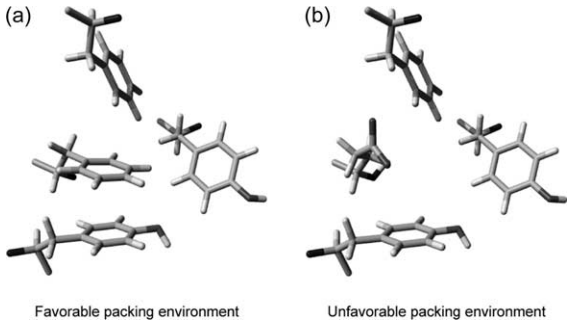

Fig. 18. Illustration of (a) favorable and (b) unfavorable packing of residues in a hydrophobic pocket. An aromatic ring located in the hydrophobic pocket is favored over the hydrophilic glutamate. Figure generated with YASARA.

Several studies showed the packing quality in NMR structures to be worse than in X-ray structures. For example, Ratnaparkhi et al. [88] analyzed 70 structures of proteins for which both NMR and X-ray structures were available, either of the same protein or within the same protein family. Using an alternative method to analyze the overall packing of the structures, expressed by the mean normalized protein-packing value [89], they found a much larger spread in packing values in the NMR structures as compared to the X-ray structures. The results of this study further suggested that this scattering of packing values is caused by the lack of information in NMR data on the packing of residues, and possibly is an artifact of the NMR structure determination process with major consequences for the accuracy of NMR structure models.

The conclusion that NMR structures generally have worse packing of residues is also confirmed by our own analysis of recent NMR structure ensembles and X-ray structures [56]. Further, our results also showed that the packing quality is heavily dependent on the refinement protocol used in the final stage of the structure determination process [27,41], confirming the suggestion that packing is not well determined by the NMR data itself. Applying the latest refinement protocols in explicit solvent using full Lennard-Jones potentials for the Van der Waals interactions shows remarkable improvements in the modeling of the packing of protein structures [27,41].

An alternative method for improving packing quality of NMR structures was developed by Clore and co-workers [50]. In this method the radius of gyration is used as a restraint in the structure calculation process. Although the method clearly improves the packing of NMR structures and addresses a major point of concern, it should be mainly regarded as a way of removing the symptoms of refinement protocols that do not include physically realistic treatments of non-bonded interactions.

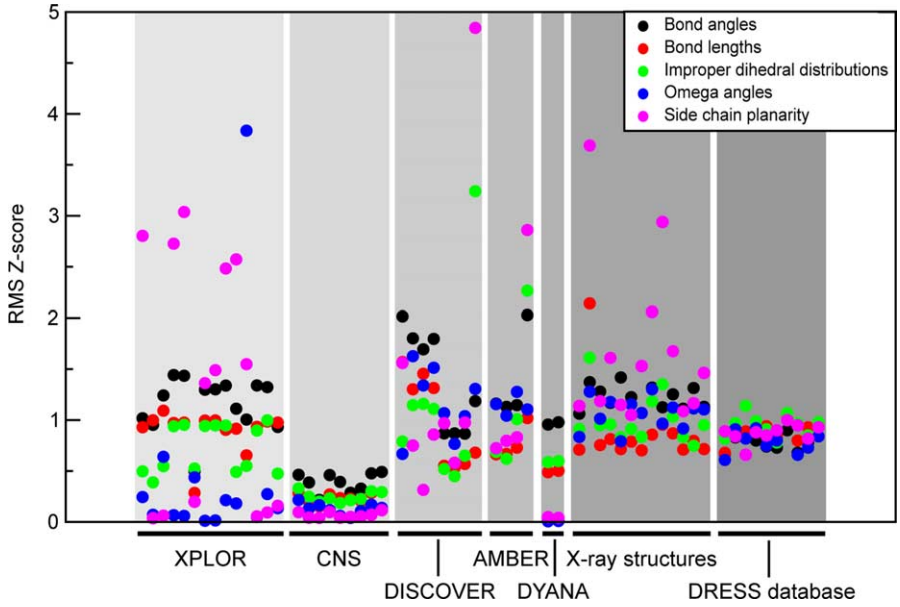

Fig. 19. RMS *Z*-scores for five parameters describing the local geometry of 37 NMR-ensembles of protein structures deposited at the Brookhaven PDB in 2001, 13 high-resolution (<1 Å) X-ray structures and 10 randomly chosen NMR-ensembles from the DRESS database [28]. On the *x*-axis we have grouped the NMR structures according to the program used for the final refinement of the structures (adapted from [41]).

5.4. NMR versus X-ray structures

It is interesting to compare NMR structures to X-ray structures from a structure validation perspective. A number of studies have shown that NMR structures in general are of lower quality than X-ray structures [8,41,56,61,65,88,90]. The difference in quality can be attributed to the much lower data content in NMR structures [42,65] and the dependence of the structure quality on the force field that was used to refine the NMR structures. Fig. 19 shows the WHAT IF RMS *Z*-scores for various parameters describing the local geometry in NMR and high-resolution X-ray structures (figure adapted from [41]). We observed large variation in RMS *Z*-scores among the NMR structures solved with different software packages. To us, this indicates the absence of a consensus in the methods for restraining the local geometry, with many structures being too tightly or too loosely restrained in one or more of the structure properties. Markedly, some extreme outliers are found, mostly for the side chain planarity, but also for one structure in which the ω angles have values up to 90°, yielding of course an exceptionally poor RMS *Z*-score ~3.9. As the figure also shows, large deviations from the expected values occur even in X-ray structures solved at 1 Å resolution, indicating that high-resolution X-ray structures should also be checked carefully for abnormalities and possible errors.

Table 3 shows how NMR structures compare to the reference database of high quality X-ray structures in terms of overall quality indicators and inter-atomic bumps. It is obvious that even recently released NMR structure ensembles score on average rather low on the main quality indicators and, in terms of structure *Z*-scores, can best be compared to X-ray structures of ~4 Å resolution (data not shown).

Table 3
Structure quality indicators of X-ray and NMR structures

	X-ray[a]	NMR[b]	DRESS[c]
Structure Z-scores			
First generation packing quality	−0.4±1.3	−3.1±2.2	−2.1±1.8 (−3.4±1.8)
Second generation packing quality	0.3±1.9	−2.8±2.2	−1.8±1.0 (−2.6±1.0)
Ramachandran plot appearance	−1.4±1.7	−4.2±1.7	−4.3±1.4 (−6.1±1.9)
$\chi1/\chi2$ rotamer normality	−1.2±1.3	−3.5±1.8	−1.6±1.4 (−5.1±2.1)
Backbone conformation	−0.6±0.9	−3.5±2.9	−3.7±2.4 (−4.5±3.1)
Inter-atomic bumps			
No. bumps/100 residues	7±6	40±52	10±8 (88±60)

[a] The analysis included 267 X-ray structures released in 2004.
[b] The analysis included 173 NMR ensembles of proteins released in 2004.
[c] The analysis included 100 NMR ensembles in the DRESS database released between 1993 and 2002 (validation scores for the original structures are given in parentheses). Note: validation scores for the DRESS database deviate from those in the original paper [28], due to the use of an updated reference database in WHAT IF.

Table 4
Average quality indicators for 100 NMR structures from the DRESS database

	Before refinement	After refinement
RMS violation of all distance restraints	0.09±0.09	0.03±0.02
RMS violation of all dihedral angle restraints	4.4±6.4	0.5±0.4
Number of consistently violating restraints[a]	8±17	0±1
PROCHECK Ramachandran Plot results		
Most favored	65.0±17.9	74.9±15.7
Allowed	27.0±12.0	19.4±10.2
Generously allowed	4.7±4.0	2.2±2.0
Disallowed	3.3±13.8	3.5±13.8

[a] Defined as those restraints that violate (>0.5 Å) in more than 50% of the members of a structure ensemble of at least 10 structures.

6. Refinement of structures in explicit solvent

As has been discussed in previous sections, many of the structural problems present in typical NMR structures appear to originate from the applied force field and structure refinement protocol. One of the most important causes of problems is that the non-bonded interactions are often severely simplified for the sake of computational speed. These simplifications result in an unrealistic treatment of the electrostatic and Van der Waals interactions and many NMR structures therefore exhibit bad packing quality, high numbers of inter-atomic bumps and non-optimal internal hydrogen bonding patterns.

From the mid-nineties several studies have demonstrated that biomolecular NMR structures can be significantly improved by the inclusion of explicit solvent molecules, or

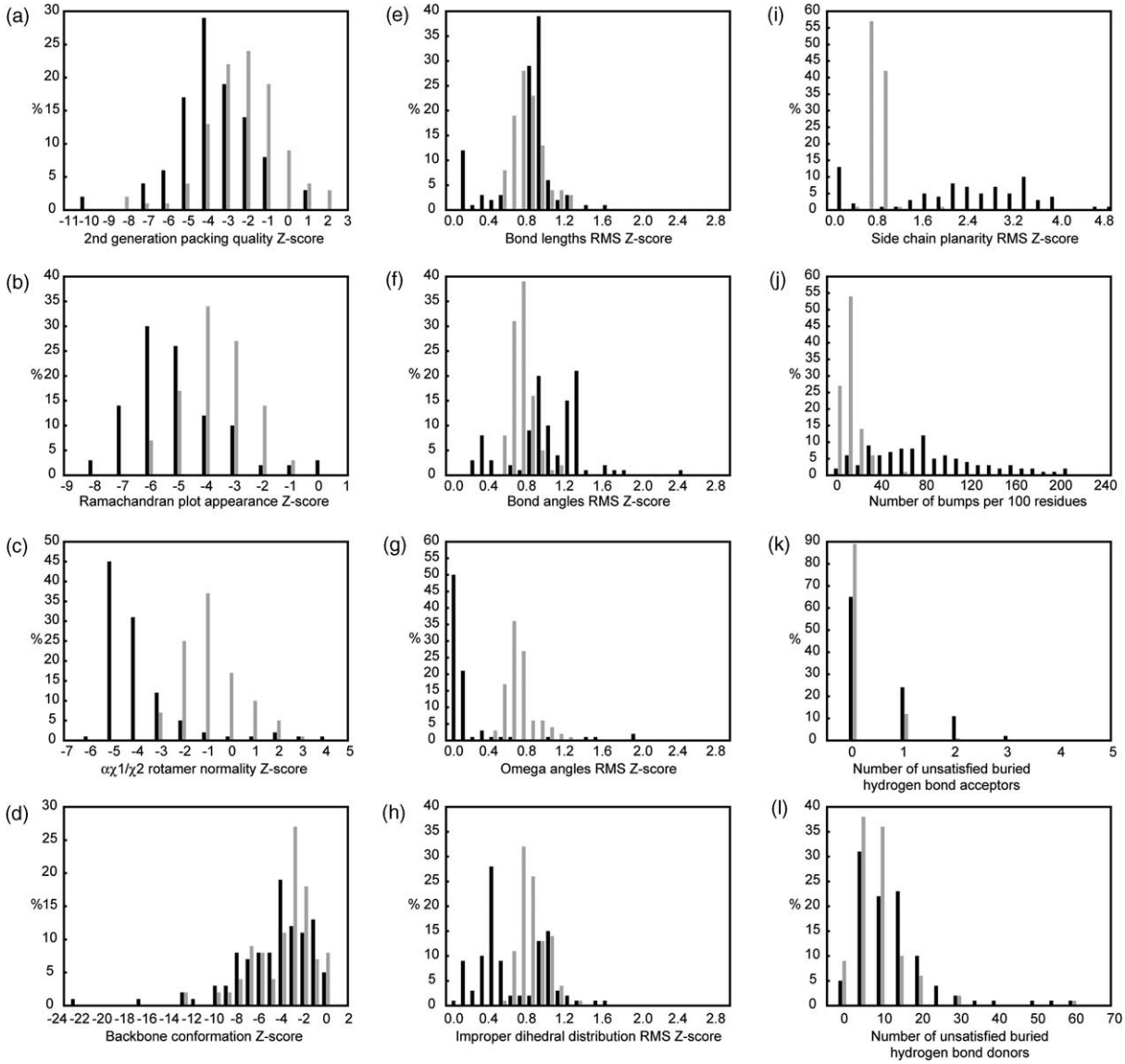

Fig. 20. Distributions of WHAT IF quality indicators for 100 NMR structures in the DRESS database before refinement in explicit solvent (black bars) and after refinement in explicit solvent (gray bars).

the use of an implicit solvent model, in restrained molecular dynamics refinement using force fields that include more realistic non-bonded interaction terms [27,41,81,91–94]. Unfortunately, not all newly deposited structures are refined using the latest protocols and many structures in the PDB are poorly refined according to today's standards. With the increased computational speed it is now possible to perform refinements in explicit solvent in a matter of minutes to hours per structure, and it has become feasible to refine NMR structure ensembles on a large scale. Recently, we have started a collaborative effort involving several European- and US-based laboratories to refine deposited NMR structure ensembles, resulting in the DRESS database of refined solution structures [28] (http://www.cmbi.ru.nl/dress). The DRESS database uses slightly modified refinement protocols from ARIA2.0 [27], resulting in significantly improved NMR structure models as compared to the originally deposited structures (cf. Table 4 and Fig. 20). Improvement on all quality indicators is so obvious, that refinement in explicit solvent should in future be considered to constitute standard practice.

7. Concluding remarks

We have attempted to give an overview of the most important aspects of validation of protein structure models derived from NMR spectroscopic data. It is of great importance for the reliability and usefulness of biomolecular structure models that they are carefully checked and optimized prior to deposition in the structure databases. Further, alongside the coordinates it is crucial that the original NMR data are also deposited in order to identify possible errors in the process of assignment and conversion to coordinate restraints, and to allow future re-interpretation and optimization of the NMR data and structures when new improved methods become available. Currently, it is possible to deposit chemical shifts and coordinate restraints at the BioMagResBank and the Protein Data Bank, and a deposition system for raw NMR spectral data is being developed in the collaborative computing project for the NMR community (CCPN) [95]. Again, we urge researchers in the NMR community to make use of the available systems for deposition of both data and structures. Only then we can aim to investigate possible causes for systematic errors in NMR derived structure models in a detailed manner, and develop methods for improvement of the structure determination process on the basis of NMR data.

Acknowledgments

The authors wish to thank Hugo van Ingen for critical reading of the manuscript and Tine Walma for providing the PDZ2-AS dataset. Further, we thank all NMR spectroscopists who deposited their experimental restraints along with the structure coordinates, without whom the presented NMR structure validation work would not have been possible.

C. Spronk and S. Nabuurs are financially supported by the Netherlands Organization for Scientific Research (NWO/CW), and the European community (5th Framework program NMRQUAL contract number QLG2-CT-2000-01313), respectively.

References

[1] J.D. Watson, F.H.C. Crick, Nature 171 (1953) 737–738.
[2] J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, D.C. Phillips, Nature 181 (1958) 662–666.
[3] G.J. Kleywegt, T.A. Jones, Structure 10 (2002) 465–472.
[4] F.C. Bernstein, T.F. Koetzle, G.J. Williams, E.F. Meyer Jr, M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, Eur. J. Biochem. 80 (1977) 319–324.
[5] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, Nucl. Acids Res. 28 (2000) 235–242.
[6] M.P. Williamson, T.F. Havel, K. Wüthrich, J. Mol. Biol. 182 (1985) 295–315.
[7] R. Kaptein, E.R. Zuiderweg, R.M. Scheek, R. Boelens, W.F. van Gunsteren, J. Mol. Biol. 182 (1985) 179–182.
[8] R.A. Laskowski, Methods Biochem. Anal. 44 (2003) 273–303.
[9] R.W. Hooft, G. Vriend, C. Sander, E. Abola, Nature 381 (1996) 272.
[10] G.J. Kleywegt, Acta Crystallogr. D 56 (2000) 249–265.
[11] A.J. Dingley, S. Grzesiek, J. Am. Chem. Soc. 120 (1998) 8293–8297.
[12] F. Cordier, S. Grzesiek, J. Am. Chem. Soc. 121 (1999) 1601–1602.
[13] N. Tjandra, A. Bax, Science 278 (1997) 1111–1114.
[14] G. Cornilescu, J.L. Marquardt, M. Ottiger, A. Bax, J. Am. Chem. Soc. 120 (1998) 6836–6837.
[15] R. Sprangers, M.J. Bottomley, J.P. Linge, J. Schultz, M. Nilges, M. Sattler, J. Biomol. NMR 16 (2000) 47–58.
[16] L. Banci, I. Bertini, G.G. Savellini, A. Romagnoli, P. Turano, M.A. Cremonini, C. Luchinat, H.B. Gray, Proteins 29 (1997) 68–76.
[17] L. Banci, I. Bertini, G. Cavallaro, A. Giachetti, C. Luchinat, G. Parigi, J. Biomol. NMR 28 (2004) 249–261.
[18] G. Cornilescu, F. Delaglio, A. Bax, J. Biomol. NMR 13 (1999) 289–302.
[19] M. Nilges, M.J. Macias, S.I. O'Donoghue, H. Oschkinat, J. Mol. Biol. 269 (1997) 408–422.
[20] J.P. Linge, M. Habeck, W. Rieping, M. Nilges, Bioinformatics 19 (2003) 315–316.
[21] T. Herrmann, P. Güntert, J. Biomol. NMR 24 (2002) 171–189.
[22] T. Herrmann, P. Güntert, K. Wüthrich, J. Mol. Biol. 319 (2002) 209–227.
[23] J. Kuszewski, C.D. Schwieters, D.S. Garrett, R.A. Byrd, N. Tjandra, G.M. Clore, J. Am. Chem. Soc. 126 (2004) 6258–6273.
[24] A. Grishaev, M. Llinas, Proc. Natl Acad. Sci. USA 99 (2002) 6707–6712.
[25] A. Grishaev, M. Llinas, Proc. Natl Acad. Sci. USA 99 (2002) 6713–6718.
[26] A. Grishaev, M. Llinas, J. Biomol. NMR 28 (2004) 1–10.
[27] J.P. Linge, M.A. Williams, C.A.E.M. Spronk, A.M.J.J. Bonvin, M. Nilges, Proteins 50 (2003) 496–506.
[28] S.B. Nabuurs, A.J. Nederveen, W. Vranken, J.F. Doreleijers, A.M. Bonvin, G.W. Vuister, G. Vriend, C.A.E.M. Spronk, Proteins 55 (2004) 483–486.
[29] P. Güntert, Q. Rev. Biophys. 31 (1998) 145–237.
[30] P. Güntert, Prog. NMR Spec. 43 (2003) 105–125.

[31] A.T. Brünger, P.D. Adams, G.M. Clore, W.L. Delano, P. Gros, R.W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, et al., Acta Crystallogr. D 54 (1998) 905–921.
[32] A.T. Brünger, Nature 355 (1992) 472–475.
[33] C. Gonzalez, J.A.C. Rullmann, A.M.J.J. Bonvin, R. Boelens, R. Kaptein, J. Magn. Reson. 91 (1991) 659–664.
[34] T.L. James, Methods Enzymol. 239 (1994) 416–439.
[35] W. Gronwald, R. Kirchhofer, A. Gorler, W. Kremer, B. Ganslmeier, K.P. Neidig, H.R. Kalbitzer, J. Biomol. NMR 17 (2000) 137–151.
[36] A.M. Gronenborn, G.M. Clore, Crit. Rev. Biochem. Mol. Biol. 30 (1995) 351–385.
[37] A.T. Brünger, G.M. Clore, A.M. Gronenborn, R. Saffrich, M. Nilges, Science 261 (1993) 328–331.
[38] G.M. Clore, D.S. Garrett, J. Am. Chem. Soc. 121 (1999) 9008–9012.
[39] J. Meiler, W. Peti, C. Griesinger, J. Biomol. NMR 17 (2000) 283–294.
[40] M. Zweckstetter, A. Bax, J. Am. Chem. Soc. 122 (2000) 3791–3792.
[41] C.A.E.M. Spronk, J.P. Linge, C.W. Hilbers, G.W. Vuister, J. Biomol. NMR 22 (2002) 281–289.
[42] J.F. Doreleijers, M.L. Raves, T. Rullmann, R. Kaptein, J. Biomol. NMR 14 (1999) 123–132.
[43] G.J. Kleywegt, T.A. Jones, Structure 4 (1996) 1395–1400.
[44] R.W. Hooft, C. Sander, G. Vriend, Comp. Appl. Biosci. 13 (1997) 425–430.
[45] R.A. Laskowski, J.A. Rullmannn, M.W. MacArthur, R. Kaptein, J.M. Thornton, J. Biomol. NMR 8 (1996) 477–486.
[46] S.B. Nabuurs, C.A.E.M. Spronk, E. Krieger, H. Maassen, G. Vriend, G.W. Vuister, J. Am. Chem. Soc. 125 (2003) 12026–12034.
[47] C.P. Kloks, C.A.E.M. Spronk, E. Lasonder, A. Hoffmann, G.W. Vuister, S. Grzesiek, C.W. Hilbers, J. Mol. Biol. 316 (2002) 317–326.
[48] A.M. Gronenborn, D.R. Filpula, N.Z. Essig, A. Achari, M. Whitlow, P.T. Wingfield, G.M. Clore, Science 253 (1991) 657–661.
[49] T. Gallagher, P. Alexander, P. Bryan, G.L. Gilliland, Biochemistry 33 (1994) 4721–4729.
[50] J. Kuszewski, A.M. Gronenborn, G.M. Clore, J. Am. Chem. Soc. 121 (1999) 2337–2338.
[51] J.C. Hoch, An Amateur Looks at Error Analysis in the Determination of Protein Structure by NMR, Plenum Press, New York, 1991.
[52] D. Zhao, O. Jardetzky, J. Mol. Biol. 239 (1994) 601–607.
[53] T.F. Havel, K. Wüthrich, J. Mol. Biol. 182 (1985) 281–294.
[54] G.M. Clore, M.A. Robien, A.M. Gronenborn, J. Mol. Biol. 231 (1993) 82–102.
[55] F.R. Chalaoux, S.I. O'Donoghue, M. Nilges, Proteins 34 (1999) 453–463.
[56] C.A.E.M. Spronk, S.B. Nabuurs, A.M. Bonvin, E. Krieger, G.W. Vuister, G. Vriend, J. Biomol. NMR 25 (2003) 225–234.
[57] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, J. Appl. Crystallogr. 26 (1993) 283–291.
[58] G. Vriend, J. Mol. Graph. 8 (1990) 52–56.
[59] F.H. Allen, O. Kennard, R. Taylor, Acc. Chem. Res. 16 (1983) 146–153.
[60] R.A. Engh, R. Huber, Acta Crystallogr. A47 (1991) 392–400.
[61] J.F. Doreleijers, J.A. Rullmann, R. Kaptein, J. Mol. Biol. 281 (1998) 149–164.
[62] R.W. Hooft, C. Sander, G. Vriend, J. Appl. Crystallogr. 29 (1996) 714–716.
[63] A.L. Morris, M.W. Macarthur, E.G. Hutchinson, J.M. Thornton, Proteins 12 (1992) 345–364.
[64] P. Schultze, J. Feigon, Nature 387 (1997) 668.
[65] J.F. Doreleijers, G. Vriend, M.L. Raves, R. Kaptein, Proteins 37 (1999) 404–416.
[66] J. Kuszewski, A.M. Gronenborn, G.M. Clore, Protein Sci. 5 (1996) 1067–1080.
[67] J. Kuszewski, A.M. Gronenborn, G.M. Clore, J. Magn. Reson. 125 (1997) 171–177.
[68] J. Kuszewski, G.M. Clore, J. Magn. Reson. 146 (2000) 249–254.
[69] J. Kuszewski, C. Schwieters, G.M. Clore, J. Am. Chem. Soc. 123 (2001) 3903–3918.
[70] I. Bertini, G. Cavallaro, C. Luchinat, I. Poli, J. Biomol. NMR 26 (2003) 355–366.
[71] W. Kabsch, C. Sander, Biopolymers 22 (1983) 2577–2637.
[72] G.N. Ramachandran, C. Ramakrishnan, V. Sasisekharan, J. Mol. Biol. 7 (1963) 95.
[73] M.W. MacArthur, J.M. Thornton, J. Mol. Biol. 264 (1996) 1180–1195.
[74] M.S. Weiss, A. Jabs, R. Hilgenfeld, Nat. Struct. Biol. 5 (1998) 676.
[75] K.S. Wilson, S. Butterworth, Z. Dauter, V.S. Lamzin, M. Walsh, S. Wodak, J. Pontius, J. Richelle, A. Vaguine, C. Sander, et al., J. Mol. Biol. 276 (1998) 417–436.
[76] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, J. Comp. Chem. 4 (1983) 187–217.
[77] C.-I. Branden, J. Tooze, Introduction to Protein Structure, second ed, Garland Publisher, New York, 1999.
[78] I.K. Mcdonald, J.M. Thornton, J. Mol. Biol. 238 (1994) 777–793.
[79] D.S. Wishart, B.D. Sykes, F.M. Richards, Biochemistry 31 (1992) 1647–1651.
[80] D.S. Wishart, B.D. Sykes, J. Biomol. NMR 4 (1994) 171–180.
[81] B. Xia, V. Tsui, D.A. Case, H.J. Dyson, P.E. Wright, J. Biomol. NMR 22 (2002) 317–331.
[82] M.B. Bass, D.F. Hopkins, W.A.N. Jaquysh, R.L. Ornstein, Proteins 12 (1992) 266–277.
[83] I.K. Mcdonald, J.M. Thornton, Protein Eng. 8 (1994) 217–224.
[84] R.W.W. Hooft, C. Sander, G. Vriend, Proteins 26 (1996) 363–376.
[85] B. Honig, A. Nicholls, Science 268 (1995) 1144–1149.
[86] J.E. Nielsen, K.V. Andersen, B. Honig, R.W. Hooft, G. Klebe, G. Vriend, R.C. Wade, Protein Eng. 12 (1999) 657–662.
[87] G. Vriend, C. Sander, J. Appl. Crystallogr. 26 (1993) 47–60.
[88] G.S. Ratnaparkhi, S. Ramachandran, J.B. Udgaonkar, R. Varadarajan, Biochemistry 37 (1998) 6958–6966.
[89] B.S. DeDecker, R. OBrien, P.J. Fleming, J.H. Geiger, S.P. Jackson, P.B. Sigler, J. Mol. Biol. 264 (1996) 1072–1084.
[90] H. Fan, A.E. Mark, Proteins 53 (2003) 111–120.
[91] M. Billeter, Y.Q. Qian, G. Otting, M. Muller, W. Gehring, K. Wüthrich, J. Mol. Biol. 234 (1993) 1084–1094.
[92] J.J. Prompers, R.H.A. Folmer, M. Nilges, P.J.M. Folkers, R.N.H. Konings, C.W. Hilbers, Eur. J. Biochem. 232 (1995) 506–514.
[93] J. Kordel, D.A. Pearlman, W.J. Chazin, J. Biomol. NMR 10 (1997) 231–243.
[94] J.P. Linge, M. Nilges, J. Biomol. NMR 13 (1999) 51–59.
[95] R. Fogh, J. Ionides, E. Ulrich, W. Boucher, W. Vranken, J.P. Linge, M. Habeck, W. Rieping, T.N. Bhat, J. Westbrook, et al., Nat. Struct. Biol. 9 (2002) 416–418.
[96] T. Walma, J. Aelen, S.B. Nabuurs, M. Oostendorp, L. van den Berk, W. Hendriks, G.W. Vuister, Structure (Camb) 12 (2004) 11–20.
[97] C.D. Schwieters, J.J. Kuszewski, N. Tjandra, G.M. Clore, J. Magn. Res. 160 (2003) 65–73.